MPLS Traffic Engineering Technology

Overview

The MPLS Traffic Engineering (TE) Technology module discusses the requirement for traffic engineering in modern networks that must attain optimal resource utilization. The traffic engineered tunnels provide a means of mapping traffic streams onto available networking resources in a way that prevents the over use of subsets of networking resources while others subsets are under-utilized. All the concepts and mechanics that support traffic engineering are presented, including the tunnel path discovery with link-state protocols and tunnel path signaling with Resource ReSerVation Protocol (RSVP). Some of the advanced features of traffic engineering such as autobandwidth and guaranteed bandwidth are introduced as well.

Upon completion of this module, the learner will be able to perform the following tasks:

- Explain the need for traffic engineering to optimize network resources
- Describe the concepts of MPLS traffic engineering
- Identify MPLS traffic engineering features
- Explain the tunnel path attributes and setup procedures
- Describe the tunnel path maintenance
- Explain the enhanced traffic engineering features such as autobandwidth or guaranteed bandwidth

Outline

The module contains the following lessons:

- Traffic Engineering Concepts

- MPLS Traffic Engineering Components
- Constraint-Based Path Computation
- Path Setup and Maintenance
- Assigning Traffic to Traffic Trunks

Traffic Engineering Concepts

Overview

This lesson describes the concepts that allow service providers to map traffic through specific routes to optimize network resources - especially the bandwidth. The traffic engineering enables backbone networks to be engineered to deliver the total subscribed capacity to service provider customers more efficiently.

Importance

This lesson is a mandatory for the students planning to improve the usage of their network resources with MPLS traffic engineering.

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- Explain the need for traffic engineering for efficient usage of network resources
- Describe the concepts of traffic engineering based on constraint-based path selections
- Explain the role of MPLS in traffic engineering

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Cisco Certified Internetwork Professional (CCIP) level of knowledge or equivalent level of IP routing and Cisco IOS knowledge as well as solid understanding of MPLS and link state protocols (OSPF or Integrated IS-IS).

Mandatory Prerequisites:

- AMVS course

Optional prerequisites:

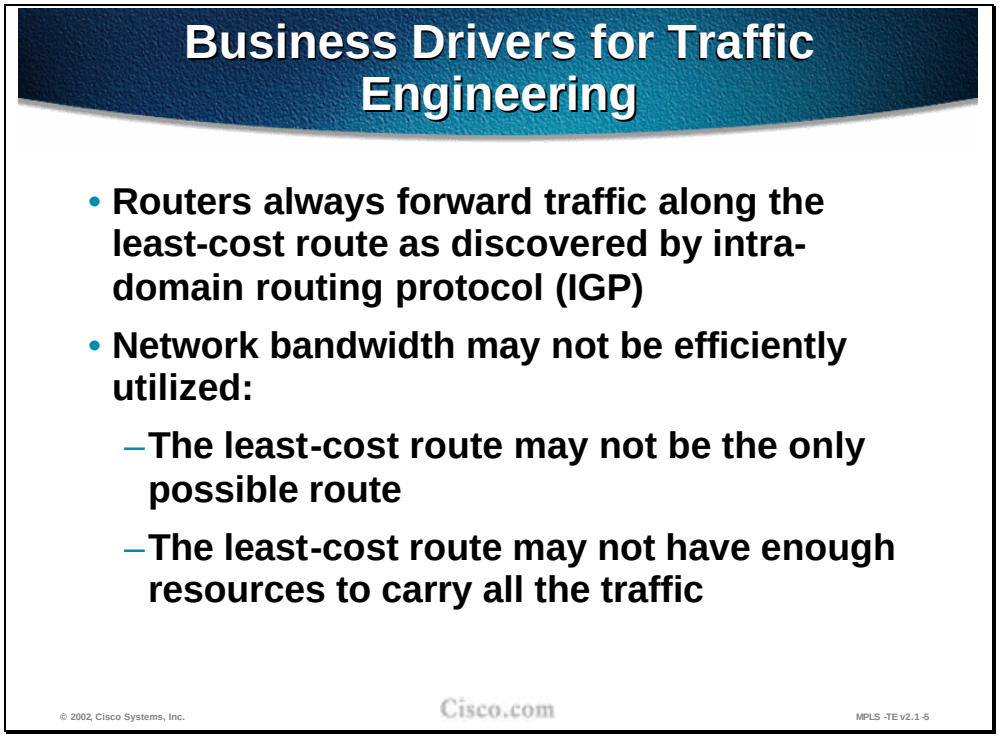
- CISIS course for students deploying MPLS TE in IS-IS environments

Outline

This lesson includes these sections:

- Overview
- Business Drivers for Traffic Engineering
- Implementing Traffic Engineering with Layer-2 Overlay Model
- Implementing Traffic Engineering with Layer-3 Model
- Using MPLS to Implement Traffic Engineering
- Summary
- Lesson Review

Business Drivers for Traffic Engineering



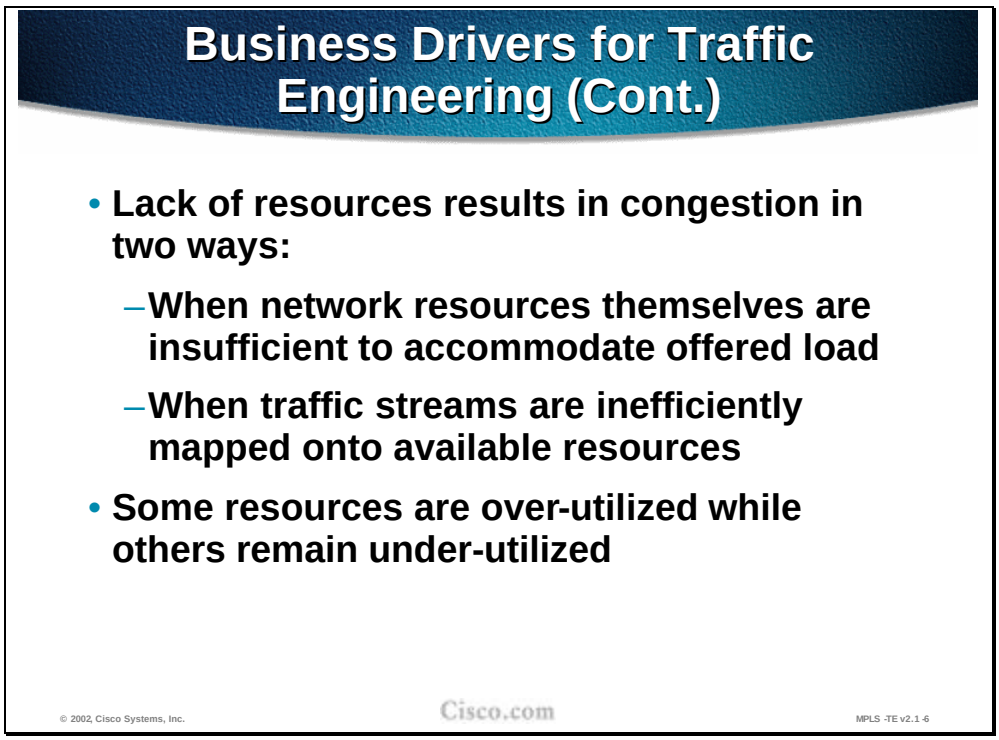
In a layer-3 routing network, packets are forwarded hop-by-hop. In each hop the destination address of the packet is used to make a routing table lookup. The routing tables are created by an interior routing protocol, IGP, which finds the least-cost route according to its metric to each destination in the network.

In many networks, this method works well. But in some networks the destination based forwarding results in the over-utilization of some links while others are under-utilized. This imbalance will be the case when there are several possible routes to reach a certain destination and the IGP selects one of them as the best and uses only that. In the extreme case, the best path may have to carry so large a volume of traffic that packets are dropped while the non-best path is almost idle.

One solution to the problem would be to adjust the link bandwidths to more appropriate values. Reduce the under utilized link and increase the over-utilized one. However, this adjustment is not always possible. The alternate path is a backup path. In the case of a primary link failure, the backup must be able to forward at least the major part of the traffic volume normally forwarded by the primary. Therefore it may not be possible to reduce the bandwidth. Without a cost saving, the budget may not allow an increase to the primary link bandwidth.

In order to provide better network performance within budget, network administrators move a portion of the traffic volume from the over-utilized link to the under-utilized link. During normal operations, this move results in less packet drops and quicker throughput. In the case of a failure to any of the links, all traffic is forwarded over the remaining link, which then of course becomes over-utilized.

Moving portions of the traffic volume cannot be achieved by traditional hop-by-hop routing using an IGP for path determination.

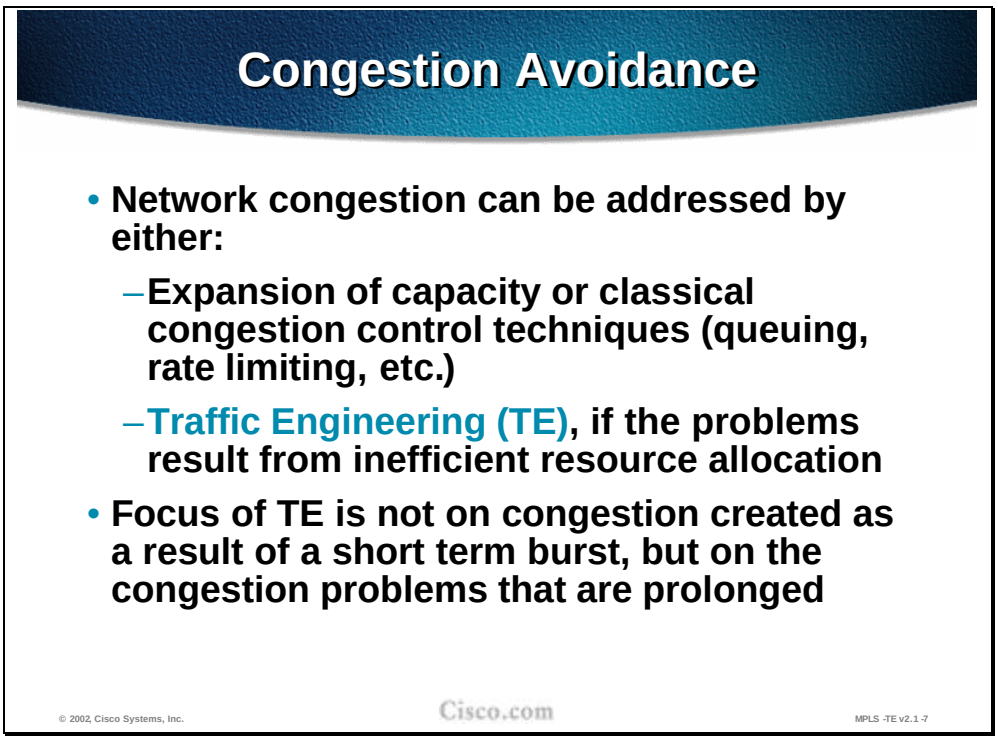


Network congestion caused by too much traffic and not enough network resources cannot be solved by moving portions of the traffic between different links. Moving the traffic will only help in the case where some resources are over-utilized while others are under-utilized. The traffic streams in normal layer-3 routing are inefficiently mapped onto the available resources.

Good mapping of the traffic streams onto the resources creates a better use of the invested money.

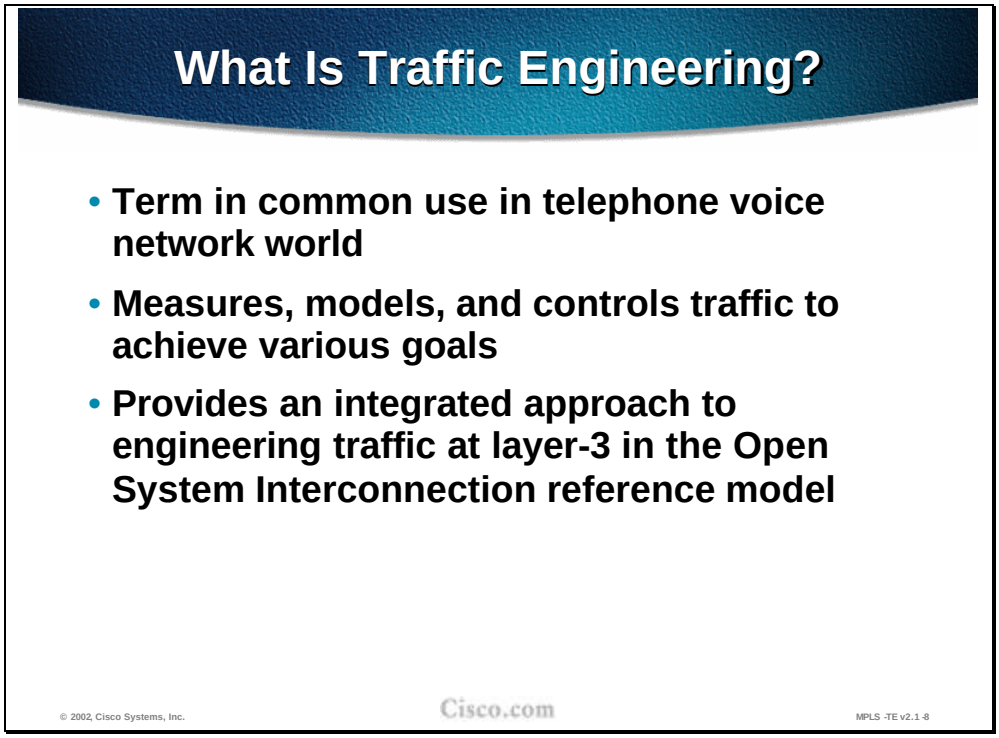
Cost savings that result in a more efficient use of bandwidth resources helps to reduce the overall cost of operations. This reduction in turn helps service providers gain an advantage over their competitors. This advantage becomes more and more important as the service provider market gets more and more competitive.

A more efficient use of bandwidth resources means that a provider could avoid a situation where some parts of its network are congested, while other parts are underutilized.



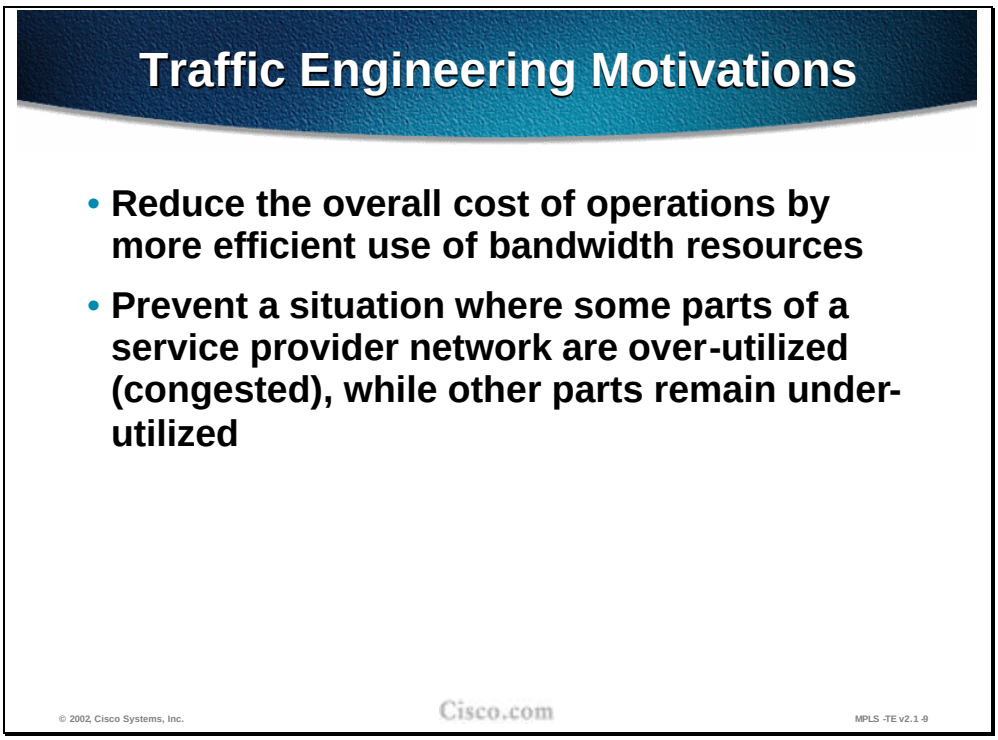
Traffic engineering does not solve temporary network congestion caused by bursty traffic. This type of problem is better handled by an expansion of capacity or by classical techniques such as various queuing algorithms, rate limiting and intelligent packet dropping.

Traffic Engineering (TE) is used when the problems result from inefficient mapping of traffic streams onto the network resources. In such networks, one part of the network suffers from congestion during long periods of time, possibly continuously, while other parts of the network have spare capacity.



The term Traffic Engineering (TE) is widely used in the telephone voice world. TE means that the traffic is measured and analyzed. Then a statistical model is applied to the traffic pattern to make a prognosis and estimations. If the anticipated traffic pattern does not match well with the network resources, the network administrators remodels the traffic pattern. Such decisions can be made to achieve a more optimum use of their own resources or to reduce costs by selecting a cheaper transit carrier.

In the data communications world, traffic engineering provides an integrated approach to engineering traffic at layer-3 in the OSI model. The integrated approach means that routers are configured to divert from destination based forwarding to move the traffic load from congested parts of the network to non-congested parts. Traditionally, this diversion was done using overlay networks where routers use carefully engineered ATM or Frame Relay PVCs to distribute the traffic load on layer-2.



Cost reduction is the main motivation for Traffic Engineering.

A cost savings that result from a more efficient use of resources help to reduce the overall cost of operations.

Additionally, more efficient use of bandwidth resources means that a provider could avoid a situation where some parts of its network are congested, while other parts are under-utilized.

Practice

Q1) What are the reasons for introducing Traffic Engineering? (Choose two.)

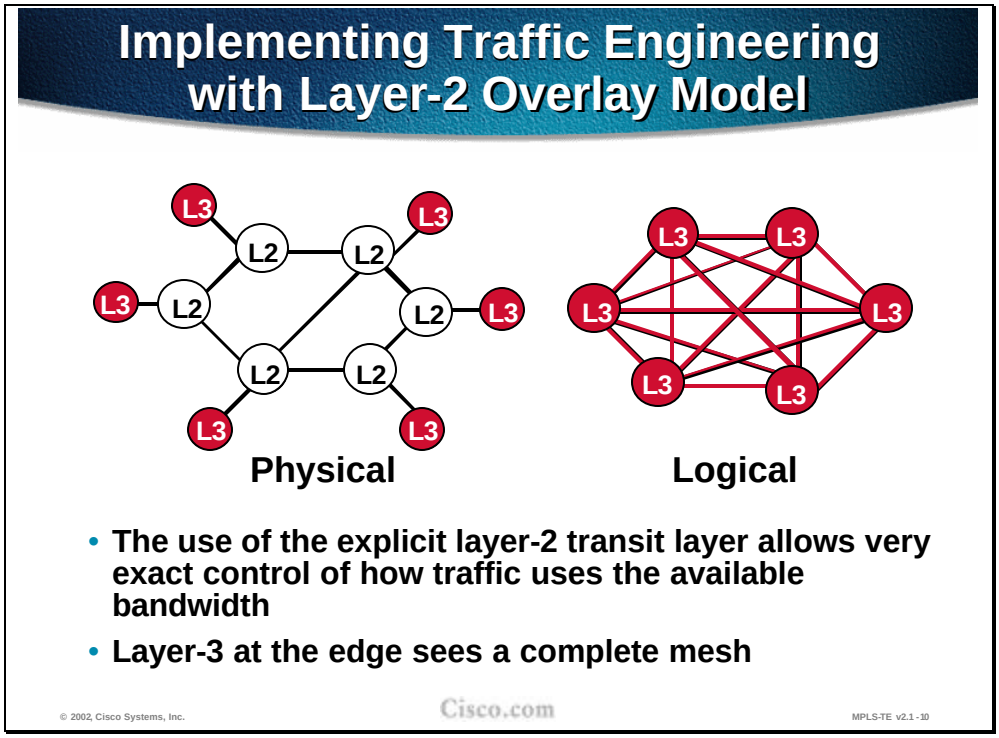
A) Traffic Engineering deals with the inefficient mapping of traffic streams onto the network resources.

B) Cost reduction is the main motivation for Traffic Engineering.

C) Traffic Engineering provides an integrated approach to engineering traffic at layer-2 in the OSI model.

D) Traffic Engineering can solve the problems of having constantly congested links.

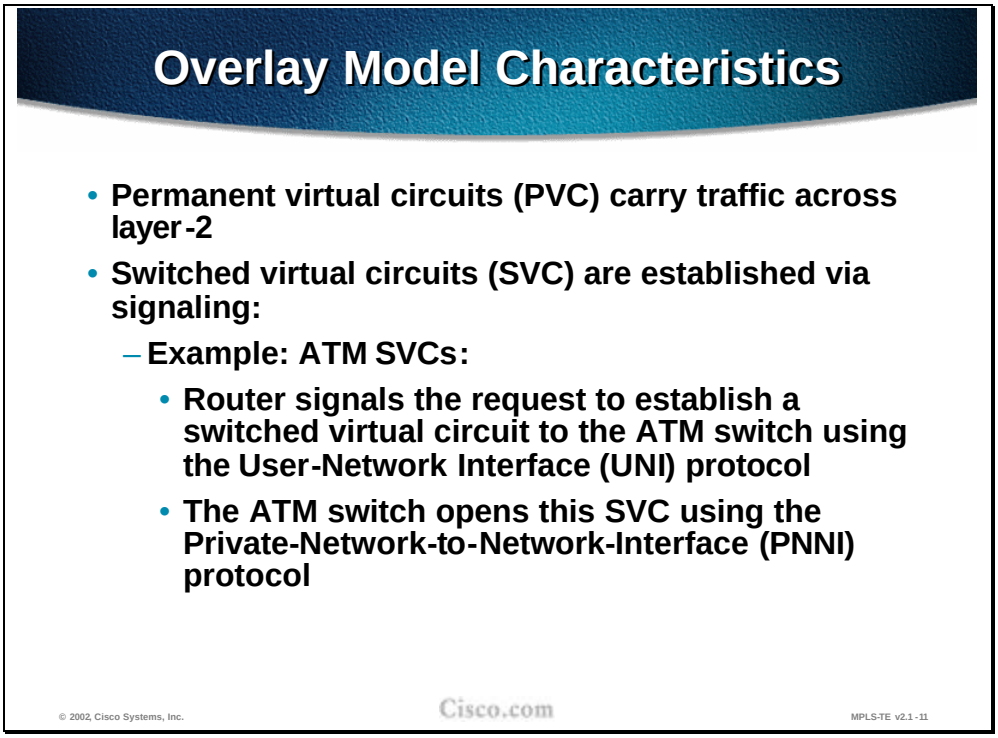
Implementing Traffic Engineering with Layer-2 Overlay Model



In the overlay model, the routers (layer-3 devices) are not aware of the physical structure and the bandwidth available on the links. The IGP views the PVCs or SVCs as point to point links and makes its forwarding decisions accordingly.

Instead all engineering is done at layer-2. PVCs are carefully engineered across the network, normally using an off-line management system. SVCs are automatically established using signaling and their way across the layer-2 network is controlled by an integrated path determination such as the PNNI protocol.

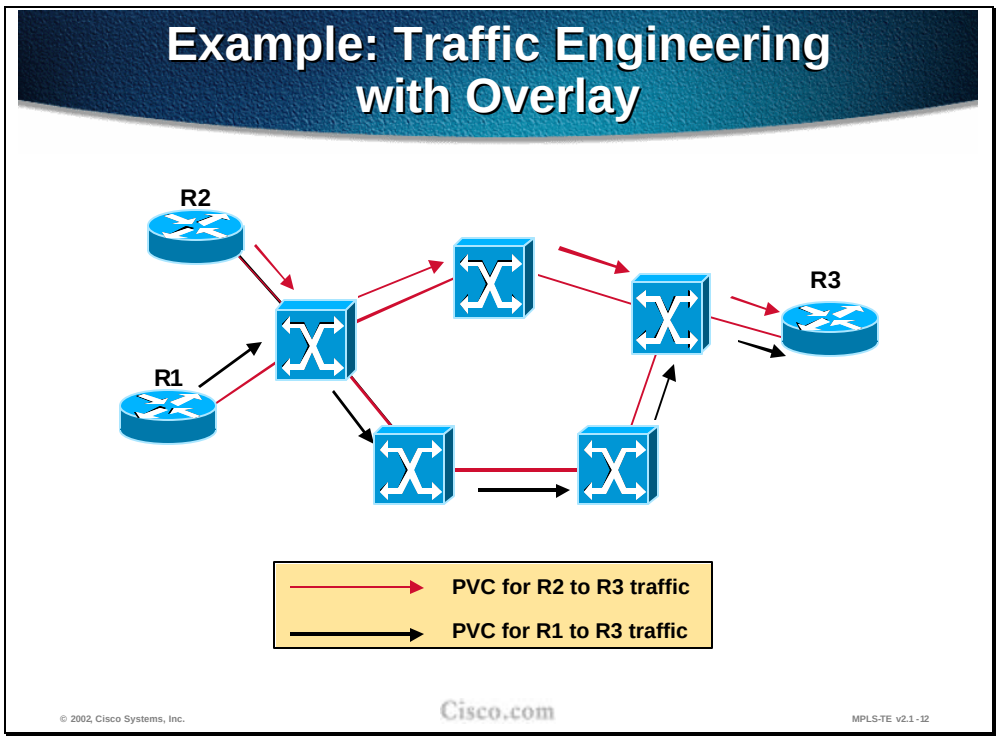
If the layer-2 network provides a full mesh between all routers, the layer-3 IGP sees all the other routers as directly connected, and, most likely, uses the direct logical link whenever forwarding a packet to another router. The full mesh gives the layer-2 full control of the traffic load distribution. Manual engineering of PVCs and/or the configuration of PNNI parameters are the tools that allow a very exact control of how the traffic uses the available bandwidth.



In the overlay model, PVCs or SVCs carry the traffic across the network.

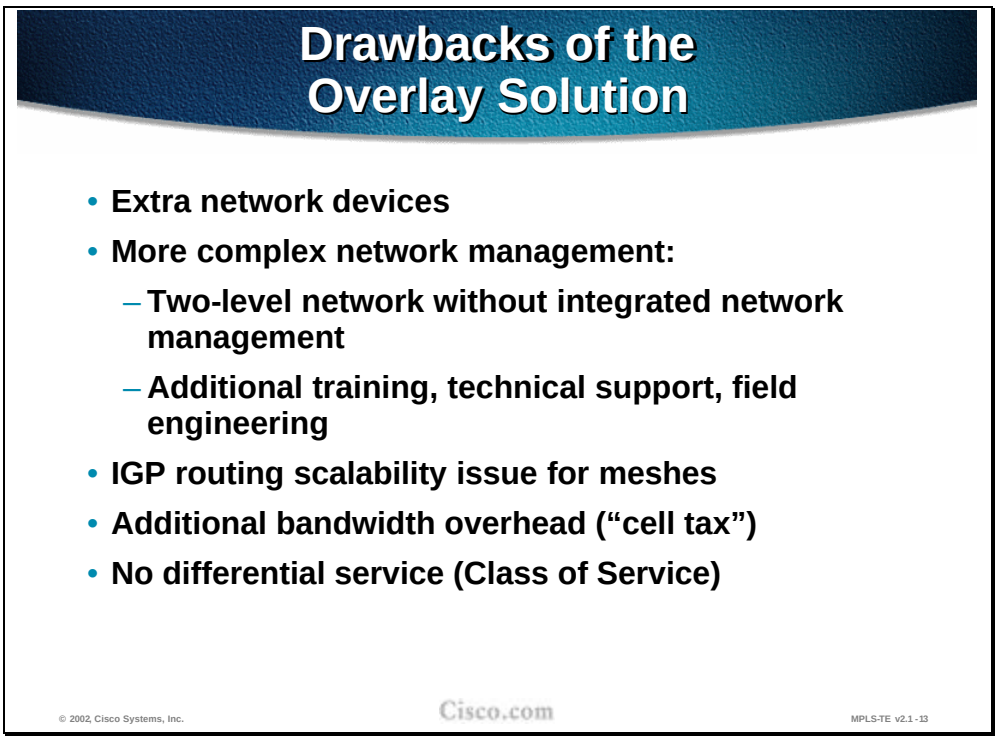
In the case of a Frame Relay network, a PVC setup is most often made using a management tool, which helps the network administrator calculate the optimum path across the layer-2 network with respect to available bandwidth and other constraints that may be applied on individual links.

ATM uses either the same type of tools as Frame Relay for PVC establishment or it may use the SVC approach where routers use a signaling protocol to dynamically establish a switched virtual circuit. When SVCs are used, the router merely asks for an SVC with certain attributes to the other router using the ATM Forum specified signaling protocol. The layer-2 network then opens this SVC internally using the PNNI (Private-Network-to-Network-Interface) protocol. PNNI, in the head end ATM switch, uses link-state information to pre-calculate a Designated Transit List (DTL), which describes the suggested total path across the ATM network. This suggested path is then validated across the ATM network by each hop switch, which then provide the SVC.



Traffic engineering in layer-2 using the overlay model, allows for detailed decisions regarding which link should be used to carry different traffic patterns.

In the example in the figure, traffic from R2 to R3 uses a PVC, which takes the shortest path using the upper transit switch. However, traffic from R1 to R3 uses a PVC, which does not take the shortest path. Traffic Engineering on layer-2 is applied to let the PVC use links that would otherwise have been under-utilized and thereby avoids over-utilization of the upper path.



Using the overlay model has several drawbacks:

- The routers are not physically connected to other routers. The layer-2 network introduces the need for an additional device, the ATM or Frame Relay switch.
- Two networks must be managed. The layer-2 network requires its own management tools, which among several other tasks support the traffic engineering as well. At the same time, the router network (layer-3) with its IGP and tuning parameters must be managed. Both these management tasks require trained staff for technical support and in the field.
- The layer-3 network must be highly meshed in order to take advantage of the benefits provided by the layer-2 network. The highly meshed network may cause scalability problems for the IGP because of the large number of neighbors.
- Overlay networks always require an extra layer of encapsulation. A Frame-Relay header must be added to the IP packets, or, when ATM is used, the IP packet must be segmented into cells, each of which must have its own header. The extra layer of encapsulation causes bandwidth overhead.
- The layer-2 devices do not have any layer-3 knowledge. Once the router has transmitted the IP packet across the physical link to the first switch, all IP knowledge is lost. When congestion does occur in the layer-2 network, the switches have no ability to selectively discard IP packets or to re-queue them due to prioritization. No IP differentiated services can be used within the layer 2 switch network.

Practice

Q1) What are the drawbacks of using overlay networks? (Choose four.)

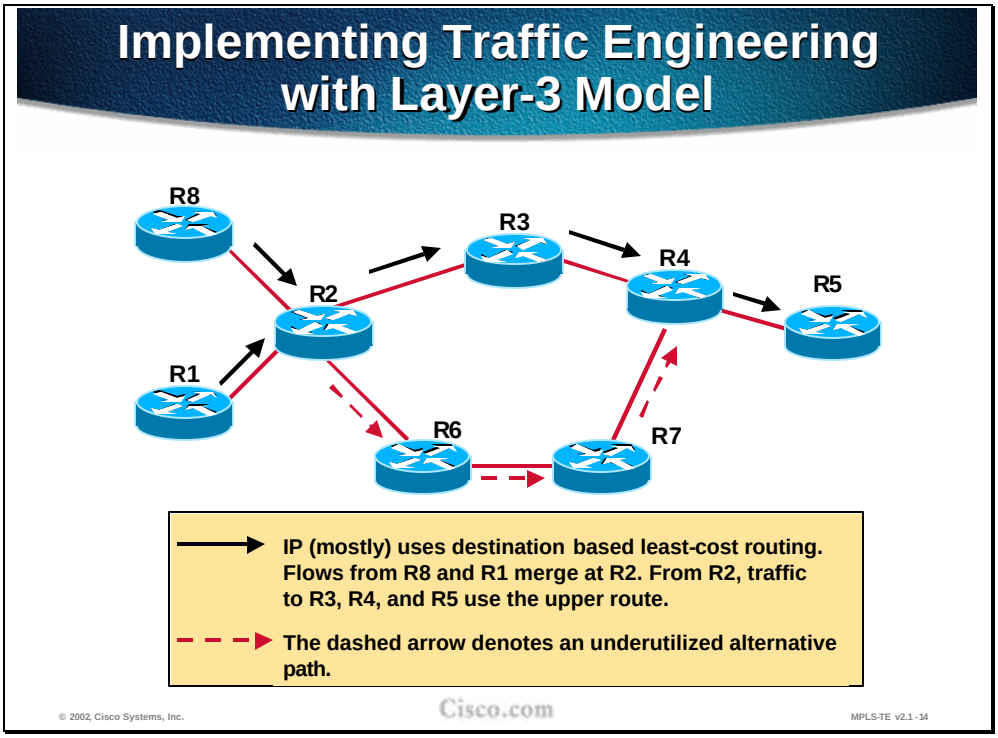
A) The layer-2 devices do not have any layer-3 knowledge for intelligent queuing and dropping.

B) The layer-2 and layer-3 network must be highly meshed.

C) Two networks must be managed.

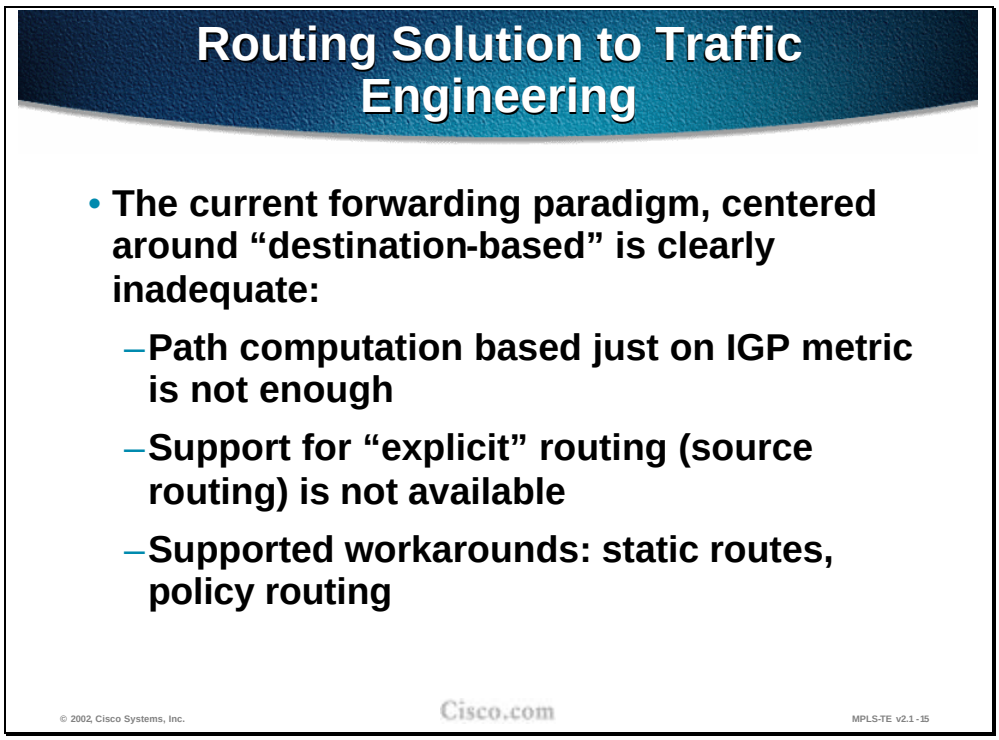
D) The layer-2 and layer-3 must be fully meshed.

Implementing Traffic Engineering with Layer-3 Model



If the same network topology is created using routers (layer-3 devices), traffic engineering must be performed differently.

- If no traffic engineering is applied to this network, traffic from both R8 and R1 towards R5 will use the least cost path (the upper path). This flow may result in the over-utilization of the path R2, R3, R4, R5 while the path R2, R6, R7, R4, R5 is under-utilized.



The destination-based forwarding paradigm currently used in layer-3 networks cannot handle the problem with over-utilization of one path while the alternate path is under utilized.

The IGP uses its metric to compute a single best way to reach each destination. Alternate routes with a higher metric are not used at all.

IP source routing could be used to override the IGP created routing table in each of the intermediate routers. However in a service provider network, source routing is most often prohibited. The source routing would also require the host to create the IP packets to request source routing. The conclusion is that source routing is not an available tool for traffic engineering.

Static routing, which overrides the IGP can be used to direct traffic to take a different path than traffic towards other destinations. However, static routing does not make it possible to discriminate between different traffic flows based on the source. Static routing also implies restrictions in how redundancy in the network can be used.

Policy based routing is able to discriminate packet flows based on the source, but suffers from low scalability and the same static routing restrictions as to how redundancy.

Practice

Q1) Why does traditional IP packet forwarding not distribute the load over all links?

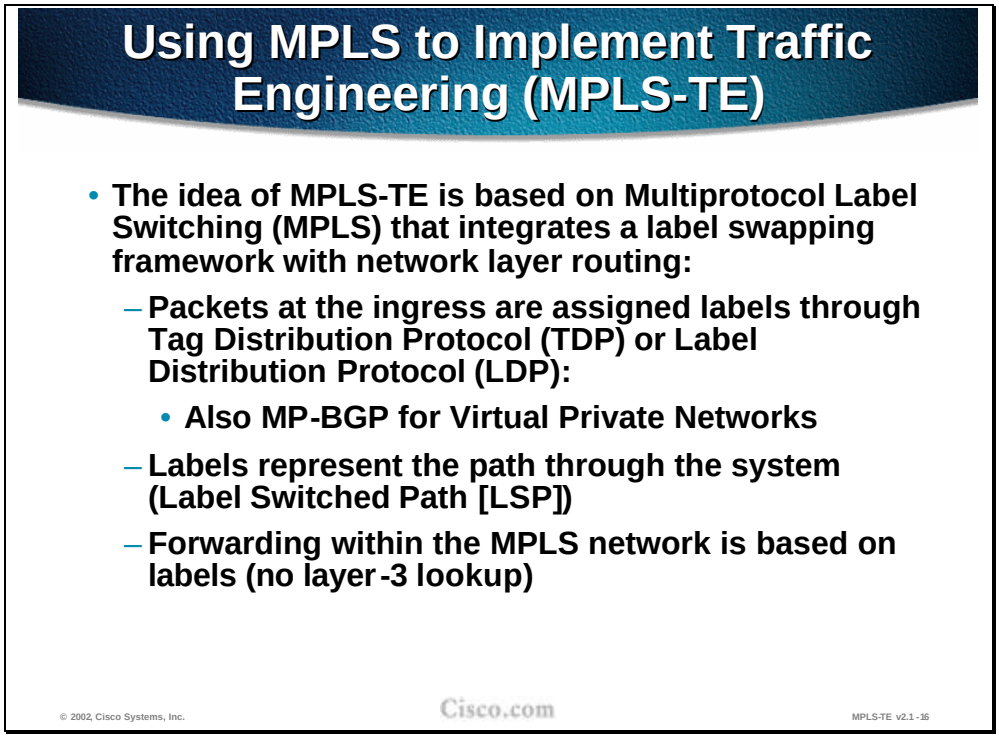
A) It can, but it requires special switching code.

B) The IGP makes one decision as to how to reach any destination - with the exception of load balancing over equal paths. Alternative routes with a higher metric are not used.

C) The IGP will always make only one decision as to how to reach any destination. Then all traffic towards that destination follows that route.

D) All routes are used and forwarding is proportional to the total cost.

Using MPLS to Implement Traffic Engineering

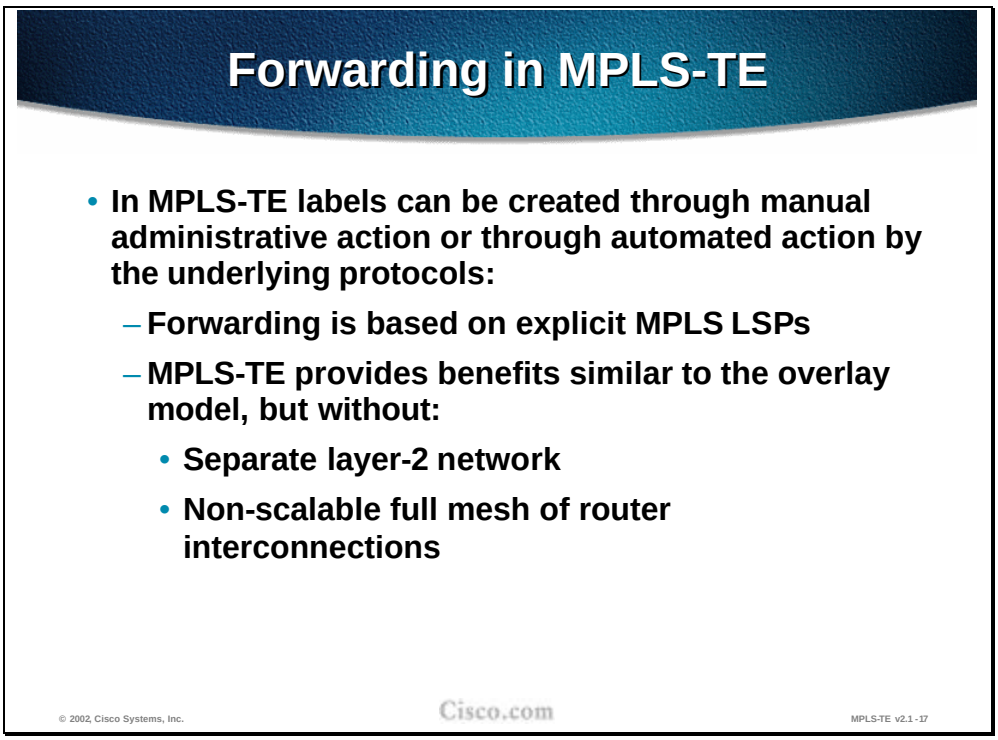


MPLS Traffic Engineering (MPLS-TE) means that the routers use the MPLS label-switching/tag-switching paradigm. Labels are assigned and distributed between routers using the Label Distribution Protocol (LDP) or the Tag Distribution Protocol (TDP). Packets are assigned labels by the ingress router, and the packet is then forwarded across the network using label switching based solely on the label, and not on the IP header information. At the egress router, the label is removed and the packet is again forwarded as an IP packet.

When full label information is exchanged, any router can reach any other router within the MPLS domain using label switching. In other words, a Label Switching Path (LSP) exists between all routers.

The existing LSPs or newly created ones between the routers are used by MPLS applications such as Virtual Private Networks (MPLS-VPN) and Traffic Engineering (MPLS-TE). A stack of two labels is imposed to the IP packet by the ingress router. The top-most label value is used to let the packet traverse the desired LSP to the router at the other end. The next label is then used by that router to indicate further actions.

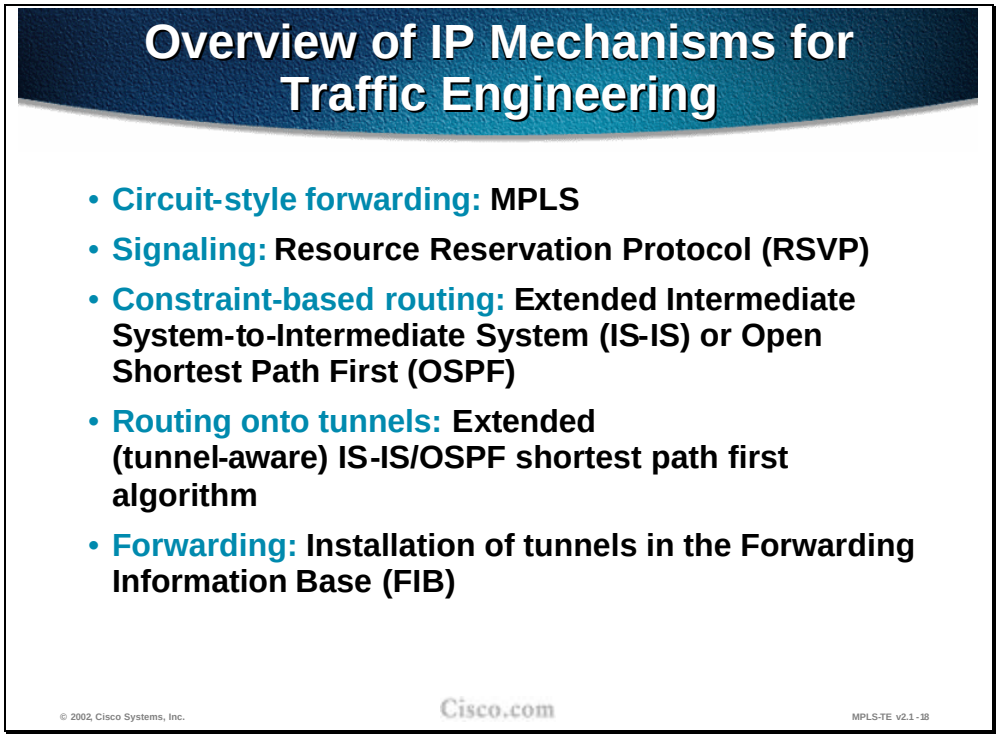
In MPLS-VPN, Multi-Protocol-BGP (MP-BGP) is used to distribute the second label in the stack used for telling the egress PE router how to forward the incoming VPN packet.



For MPLS-TE, manual assignment and configuration of the labels can be used to create LSPs to tunnel the packets across the network on the desired path. However, to increase scalability, the ReSource Reservation Protocol (RSVP) is used to automate the procedure.

The packets forwarded according to MPLS-TE have a stack of two labels (imposed by the ingress router). The top-most label identifies a specific LSP to use to reach another router at the other end of the tunnel. The second label indicates what the router at the far end of the LSP should do with the packet.

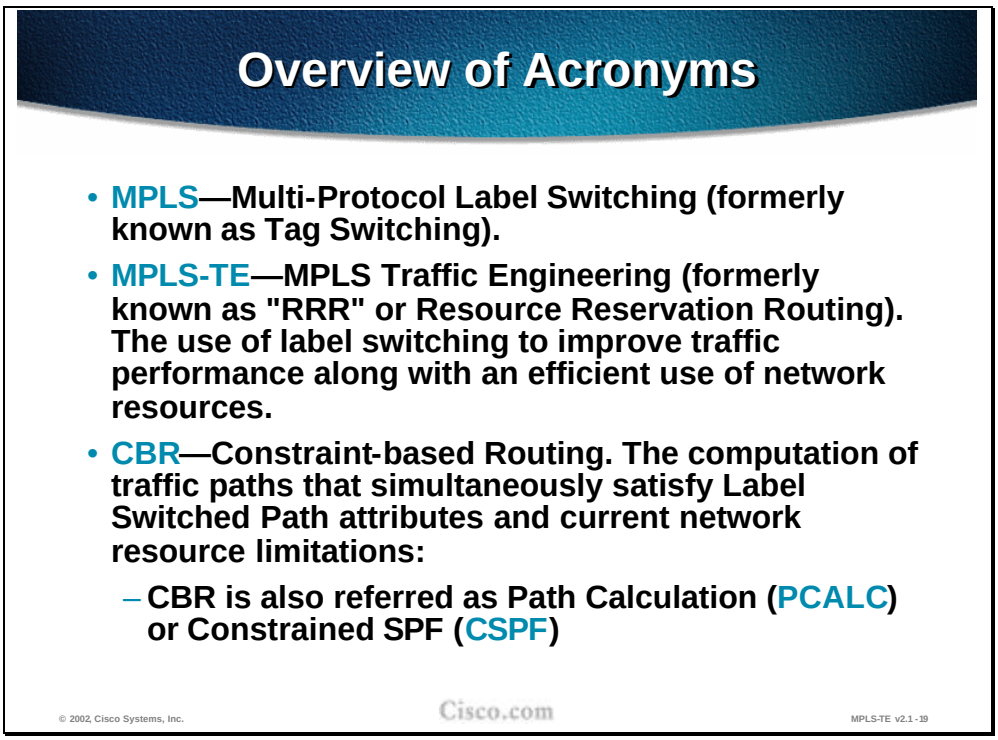
By selecting the appropriate LSP, traffic can be directed via explicitly indicated routers. The explicit path across identified routers provides similar benefits to the overlay model without introducing a layer-2 network and also without the risk of running into IGP scalability problems due to the many neighbors existing in a full mesh of routers.



MPLS-TE provides equivalent mechanisms to those described on previous slides in the overlay network. For circuit-style forwarding, instead of using ATM or Frame Relay virtual circuits, MPLS TE tunnel is used. For signaling, RSVP is used with various extensions to set up the MPLS-TE tunnels.

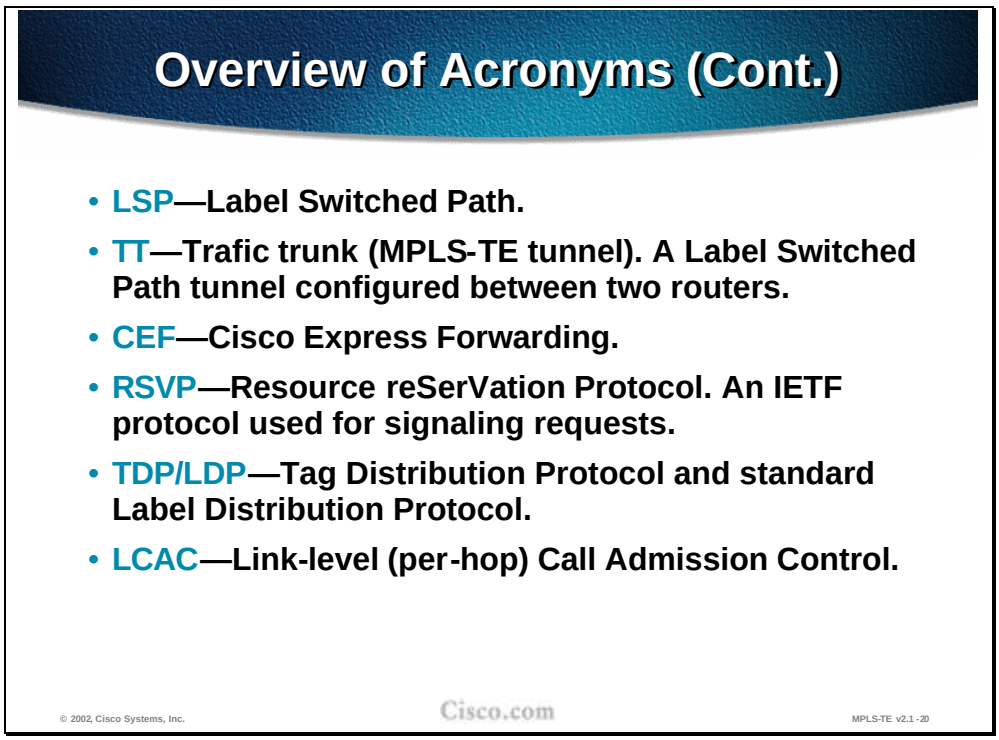
For constraint-based routing, either IS-IS or OSPF with extensions is used to carry resource information like available bandwidth on the link. Both link-state protocols use new attributes to describe the nature of each link with respect to the constraints. A link that does not have the required resource is not included in the LSP, which constitutes the MPLS-TE tunnel.

To actually direct the traffic onto the MPLS-TE tunnels, extensions to IS-IS and OSPF are needed. Directing the traffic onto tunnels results in the adding of entries in the Forwarding Information Base (FIB), the CEF-cache. The IP packets are directed into the MPLS-TE tunnel by imposing the correct label stack.



The following is a list of acronyms that is commonly used with MPLS Traffic Engineering:

- **MPLS:** Multi-Protocol Label Switching (formerly known as Tag Switching).
- **MPLS-TE:** MPLS Traffic Engineering (formerly known as “RRR” or Resource Reservation Routing). The use of label switching to improve traffic performance along with an efficient use of network resources.
- **CBR:** Constraint-based Routing. The computation of traffic paths that simultaneously satisfy Label Switched Path attributes and current network resource limitations.



The following MPLS-TE acronyms are also used very often:

- **LSP:** Label Switched Path. The path between two systems encoded with a sequence of MPLS labels.
- **TT:** Traffic trunk (MPLS-TE tunnel). A Label Switched Path tunnel configured between two routers.
- **CEF:** Cisco Express Forwarding.
- **RSVP:** Resource reSerVation Protocol. An IETF protocol used for signaling requests.
- **TDP/LDP:** Tag Distribution Protocol and standard Label Distribution Protocol.
- **LCAC:** Link-level (per-hop) Call Admission Control.

Most of the terminology used throughout this document can be found in the following documents:

- RSVP-TE: Extensions to RSVP for LSP Tunnels, RFC-3209, IETF Network Working Group, December 2001
- MPLS Traffic Engineering, RFC-2702, IETF Network Working Group, September 1999

Practice

Q1) What does MPLS provide that allows for Traffic Engineering?

A) The separation of forwarding and switching decisions.

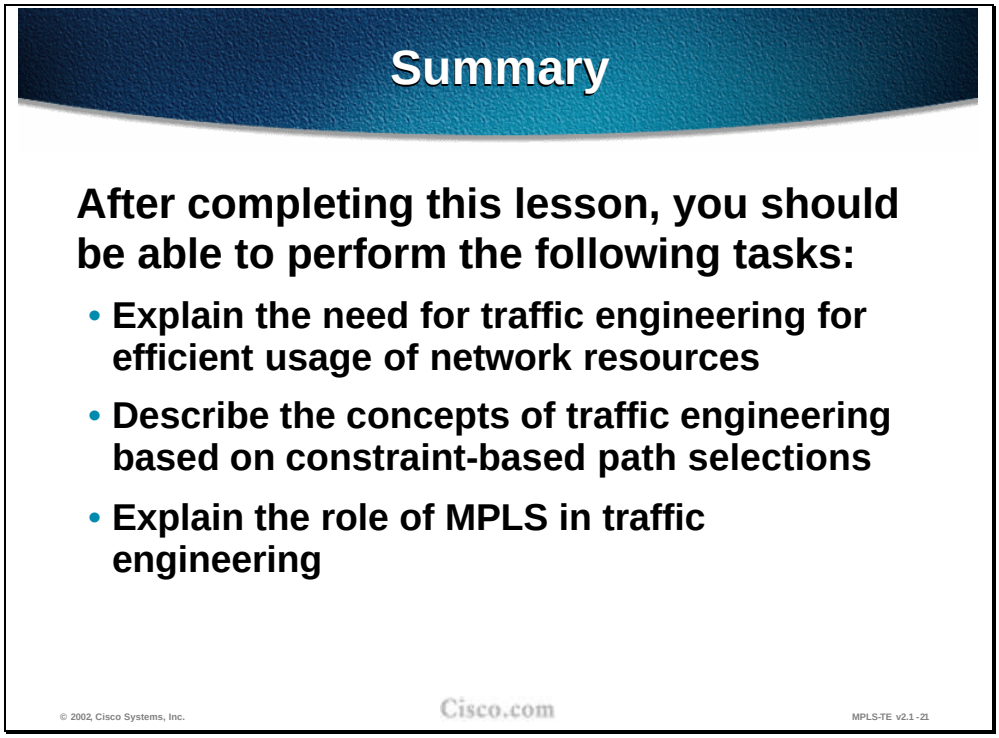
B) A separate routing table containing only Traffic Engineering addresses.

C) Packet forwarding based on labels and not based on IP destination addresses.

D) Packet forwarding based on source and destination label addresses.

Summary

This section summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- MPLS Traffic Engineering Components

Lesson Review

Instructions

Answer the following questions:

1. How can an overlay network provide Traffic Engineering?
2. What are the drawbacks of using overlay networks?
3. Why does traditional IP packet forwarding not distribute the load over all links?
4. Can IP source-routing be used to overcome the problems of overlay networks?
5. Can policy-based routing be used to overcome the problems of overlay networks?
6. What does MPLS provide that allows for Traffic Engineering?
7. Which IGPs can be used to calculate an LSP for an MPLS-TE tunnel?
8. How is the MPLS-TE created?

MPLS Traffic Engineering Components

Overview

This lesson explains the components of MPLS traffic engineering such as traffic trunks along with associated attributes, the tunnel path discovery based on link-state protocols, and the tunnel setup signaling with Resource Reservation Protocol (RSVP).

Importance

This lesson is a mandatory for the students planning to improve the usage of their network resources with MPLS traffic engineering.

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- List the components of MPLS traffic engineering
- Explain the tunnel and link attributes
- Describe the constraint-based path computation
- Describe the role of RSVP in path setup procedures
- Describe the forwarding table modification mechanisms

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Cisco Certified Internetwork Professional (CCIP) level of knowledge or equivalent level of IP routing and Cisco IOS knowledge as well as solid understanding of MPLS and link state protocols (OSPF or Integrated IS-IS).

Mandatory Prerequisites:

- AMVS course

Optional prerequisites:

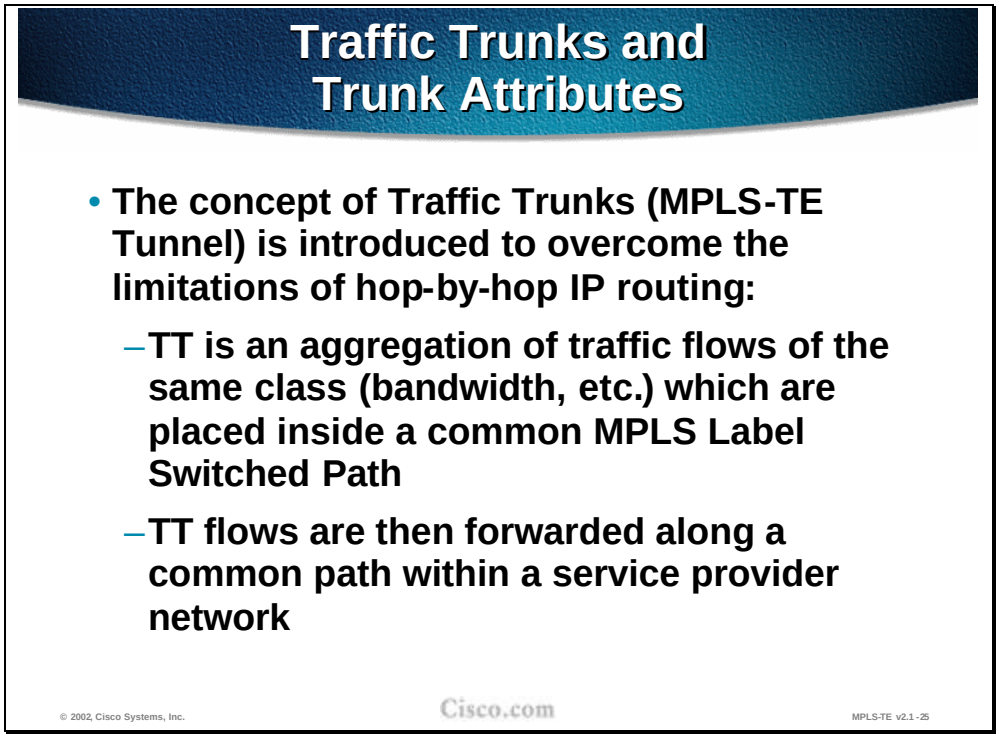
- CISIS course for students deploying MPLS TE in IS-IS environments

Outline

This lesson includes these sections:

- Overview
- Traffic Trunks and Trunk Attributes
- Network Links and Link Attributes
- Constraint-Based Path Computation
- Path Setup with RSVP Signaling
- Forwarding Table Modifications
- Summary
- Lesson Review

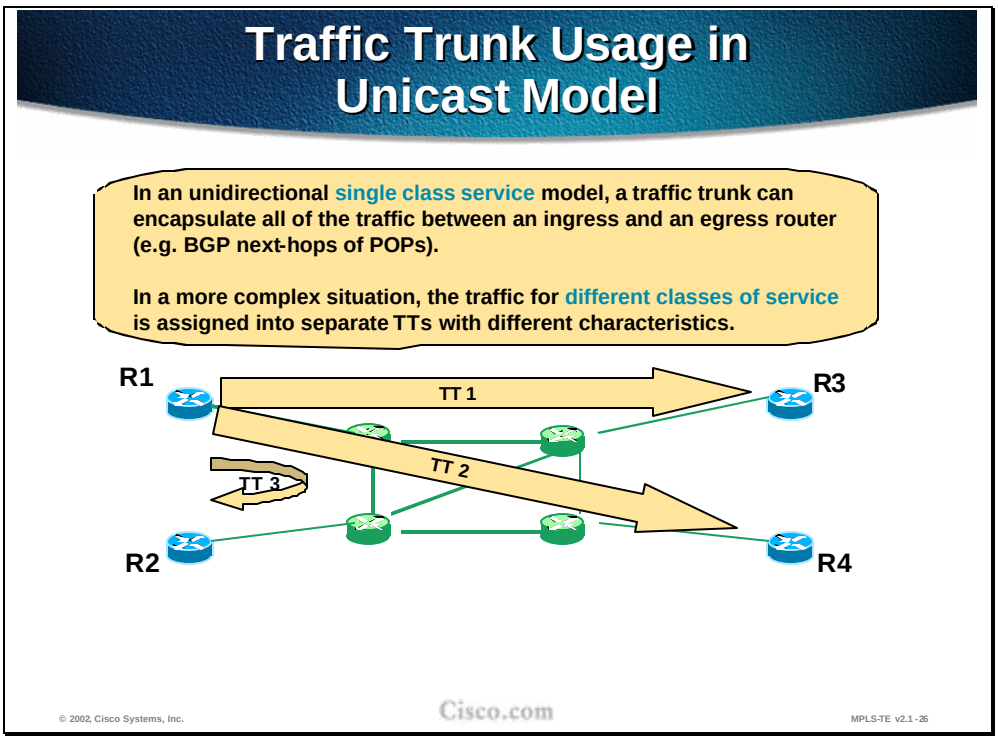
Traffic Trunks and Trunk Attributes



The aim of Traffic Engineering is to control the paths along which data flows, rather than relying simply on 'normal' destination-based routing. To fulfill this aim, the concept of a 'Traffic Trunk' must be introduced.

A Traffic Trunk is simply a collection of data flows, which share some common attribute:

- Most simply, this attribute might be traffic sharing the same entry point to the network and the same exit point. A case of this in practice would be an Internet Service Provider network, where there is a definable data flow from the Points of Presence (POP), where the customers attach to the ISP network, to the Internet eXchange points (IX), where their data typically leaves this ISP network to traverse the internet.
- In a more complex situation, this attribute could be augmented by defining separate trunks for different classes of service. For example, in an ISP model, leased-line corporate customers could be given a preferential throughput (greater guaranteed bandwidth or lower latency/higher precedence) over the dial-in home users. Even though the traffic enters and leaves the ISP network at the same points, different characteristics may be assigned to these types of users by defining separate Traffic Trunks for their data.

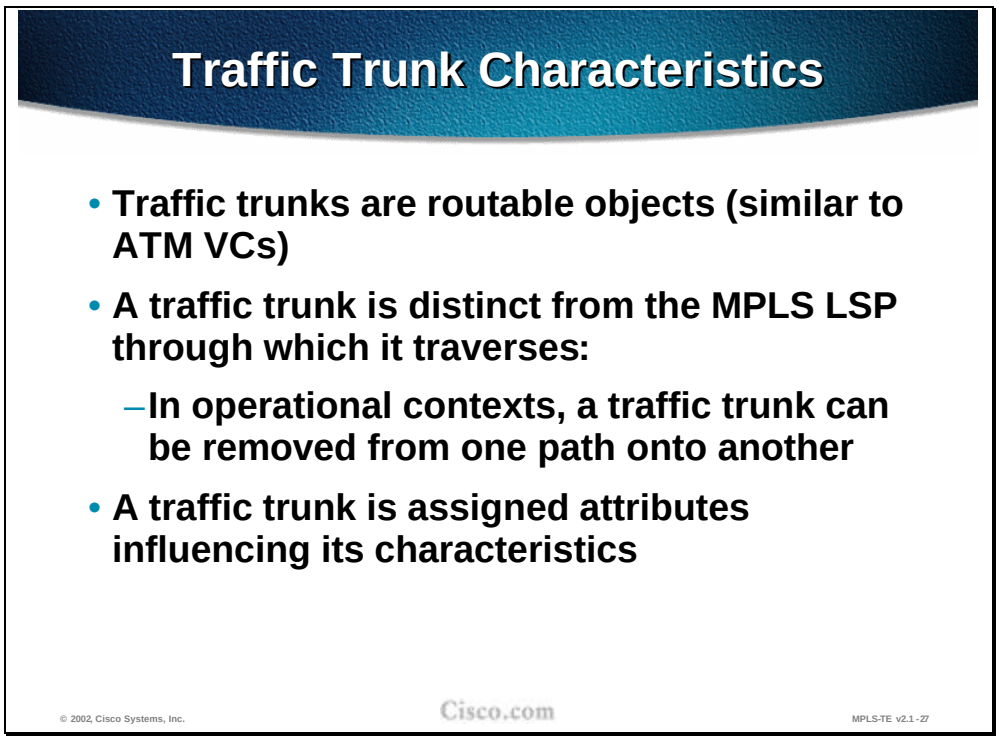


Defining the Traffic Trunks requires an understanding of the traffic flows in the network. From the understanding of the ingress and corresponding egress points, a picture of the traffic flows in the network can be produced.

In the example shown, there are Traffic Trunks (TT1, TT2 and TT3) defined for data from R1 to R2, R3 and R4. These trunks are uni-directional; they identify the traffic flows *from* R1. In practice, there are probably similar trunks operating in the opposite direction *to* R1.

There may also be trunks defined from all the other routers to each other. - Defining trunks from every router in the network to every other one might sound like an administrative nightmare: However, this is not usually the case:

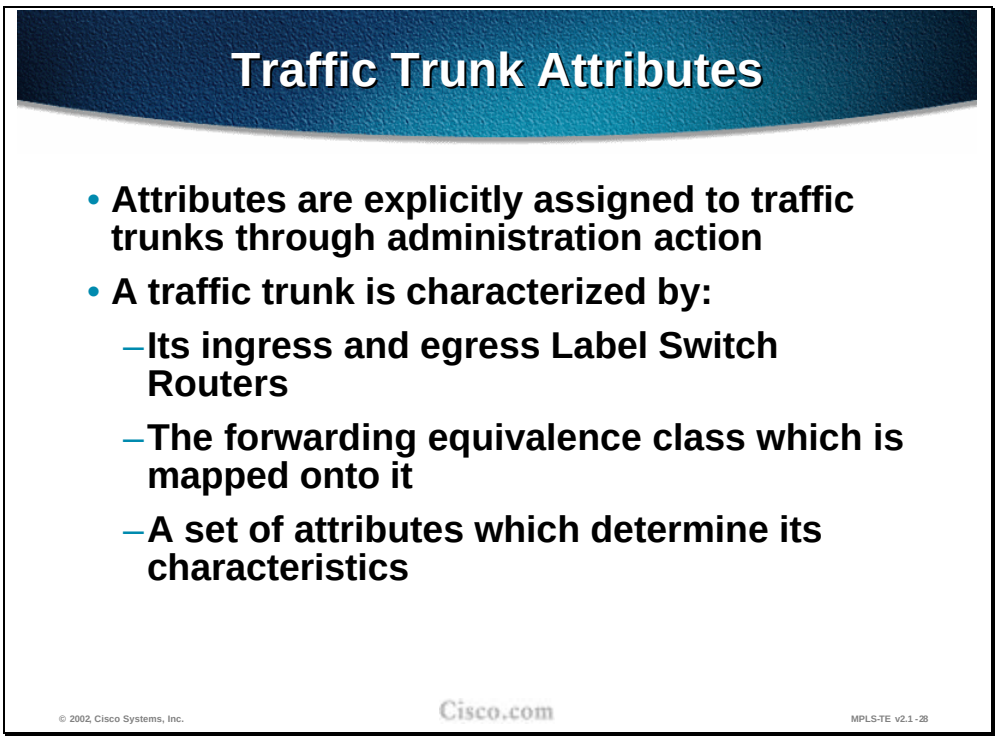
- The routers identified are on the edge of the network. The traffic trunks link these routers across the core of the network (colored green)
- In most networks it is relatively easy to identify the traffic flows and they rarely form a complete 'any-to-any' mesh.
- For example, in ISP networks, the traffic trunks would generally form a number of 'star' formations with their centers at the Internet Exchange points and the other points at the POPs. Traffic in an ISP network generally flows from the customers connected at the POPs to the rest of the Internet (reached via the IX points). A star-like formation could also exist in many networks centering on the Data-Center: both for ISP networks (providing web-hosting services) and enterprises.



Once the data flows, and therefore the Traffic Trunks are defined, the technology they use to send the data across the network is MPLS. Data entering a Traffic Trunk is assigned an MPLS Label-Switch-Path, which defines the route taken through the network. However, Traffic Trunks are distinct from the MPLS LSPs they use in two key ways:

- There is not necessarily a one-to-one mapping of Traffic Trunks on to MPLS LSPs. For administrative reasons, two Trunks may be defined between two points and may happen to pick the same path through the network. Therefore they both have the same MPLS label.
- Also, Traffic Trunks are not necessarily bound to a particular path through the network. As resources change in the core, or perhaps links fail, the Traffic Trunk may re-route, picking up a new MPLS LSP as it does.

The configuration of the Traffic Trunks includes defining the characteristics and attributes it requires. Defining the Traffic Trunks characteristics and attributes is probably the most important aspect of Traffic Engineering. Without specifying the requirements of the data in this Traffic Trunk, the data may as well be left to route ‘normally’ based on destination information only over the least cost path.

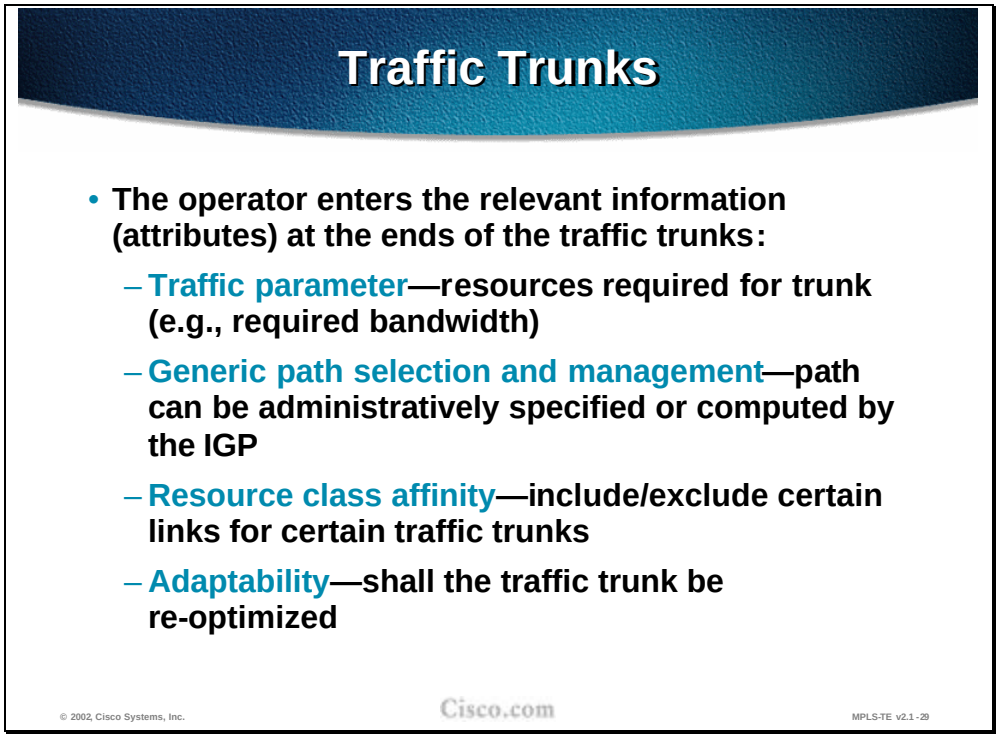


A Traffic Trunk is a set of data flows sharing some common feature, attribute or requirement. If there is no characteristic in the data flow to make it common with some other flow, there is nothing to define that data as part of a flow or group of flows.

Therefore, the Traffic Trunk, in its very definition, must include the definition of those attributes which define the commonality between the data flows making up the Trunk. The attributes that characterize a Traffic Trunk includes:

- Most fundamentally, the ingress and egress points: the routers at the ends of the Trunk. This is the most basic level of commonality between data flows; they start in the same place and end in the same place.
- More complex characteristics of the data flows, such as bandwidth and latency/precedence requirements.
- The class of data: what data is 'part of' this Trunk and what is not (which in itself is a combination of the above)

The attributes of a Traffic Trunk are defined by the network administrator when the Trunk is defined, however, some of them are in part influenced by the underlying network and protocols.

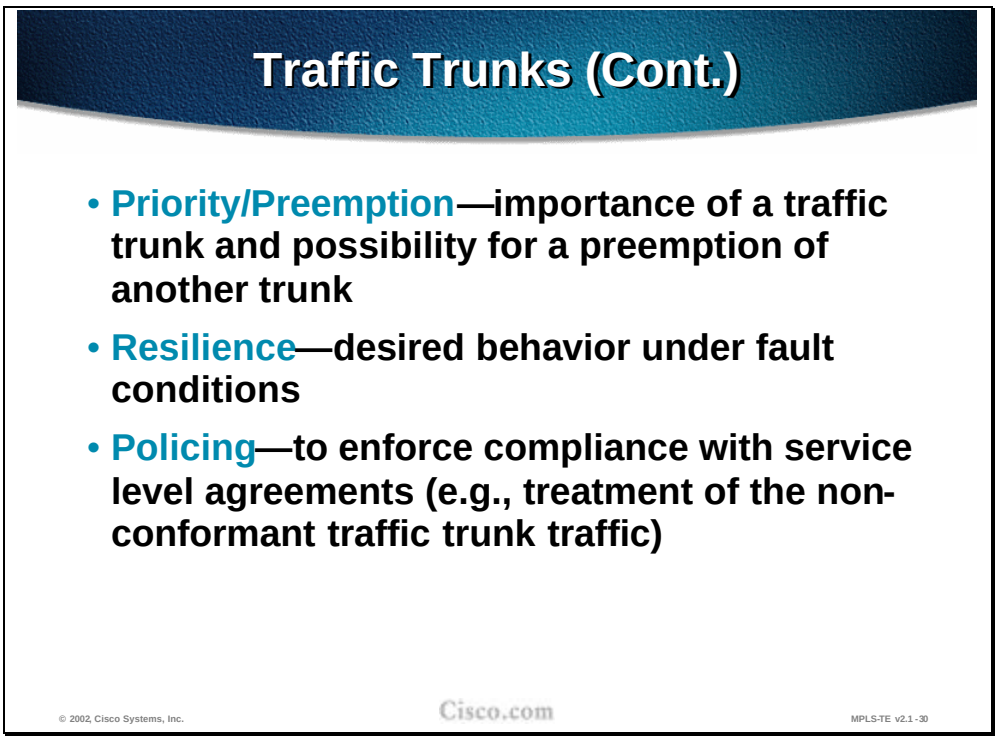


The characteristics that define the trunk are configured by the network operator include some or all of the following:

- Traffic Parameters: the resources required by the trunk, such as the minimum required bandwidth.
- Generic Path Selection and Management: the path selection criteria. The actual path chosen through the network could be statically configured by the operator or could be assigned dynamically by the network based on information from the IGP (IS-IS or OSPF).
- Resource Class Affinity: restricting the choice of paths by allowing the dynamic path to choose only certain links in the network rather than being allowed to use any link.

Note Alternatively this can be done by using the IP address exclusion feature.

- Adaptability: the ability of the path to re-route on failure or to optimize on recovery/discovery of the ‘better’ path.



Continuing the list of Traffic Trunk parameters:

- Priority/Pre-emption: Traffic Trunks can be assigned a priority (0 to 7) signifying their 'importance'. When setting up a new trunk or re-routing, a higher priority trunk can tear down (preempt) a lower priority trunk; or a new trunk of lower priority may fail to set up because some trunks of a higher priority already exist occupying the required bandwidth of the lower priority trunk.
- Resilience: What happens to a Traffic Trunk in the event of a failure in the network. Does it attempt to re-route around failures or not?
- Policing: How the trunk enforces compliance to the service-level (bandwidth, precedence) and what it does with traffic, which exceeds the service-level (examples, drop non-conforming data or send it as 'best effort').

Practice

Q1) What are the characteristics of a traffic trunk? (Choose two.)

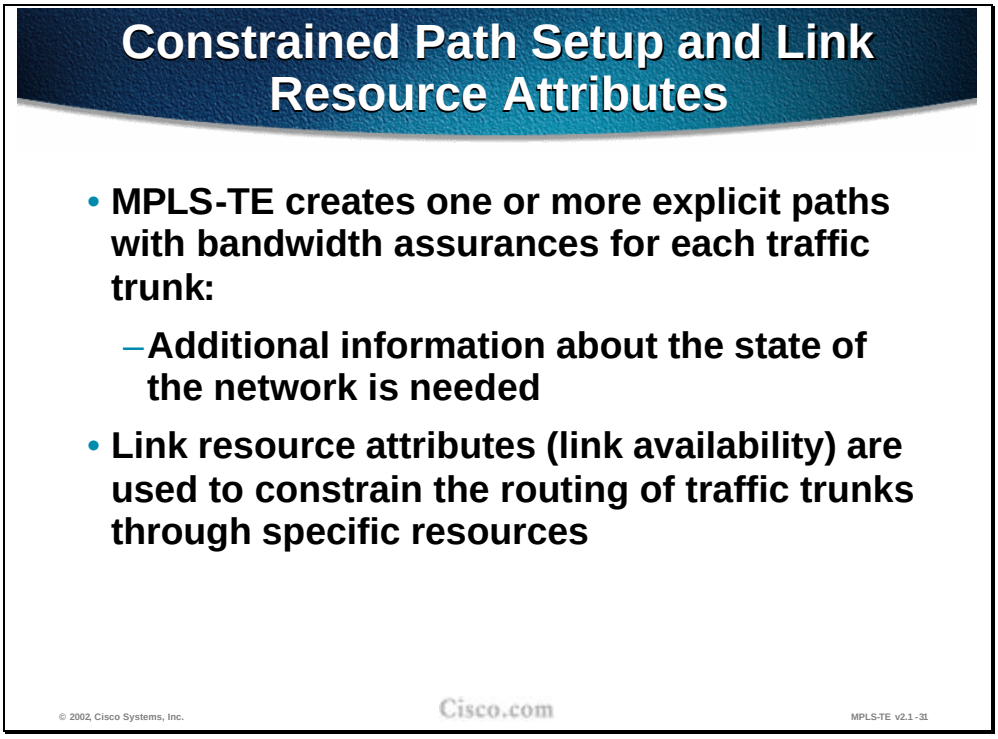
A) A traffic trunk is distinct from the MPLS LSP through which it traverses.

B) A Traffic Trunk represents a tunnel between two end-point using GRE encapsulation.

C) Once the path for the Traffic Trunk is established, it cannot be removed from one LSP path onto another.

D) A routable object characterized with ingress and egress LSR routers (head-end and tail-end), its forward equivalence class and a set of attributes.

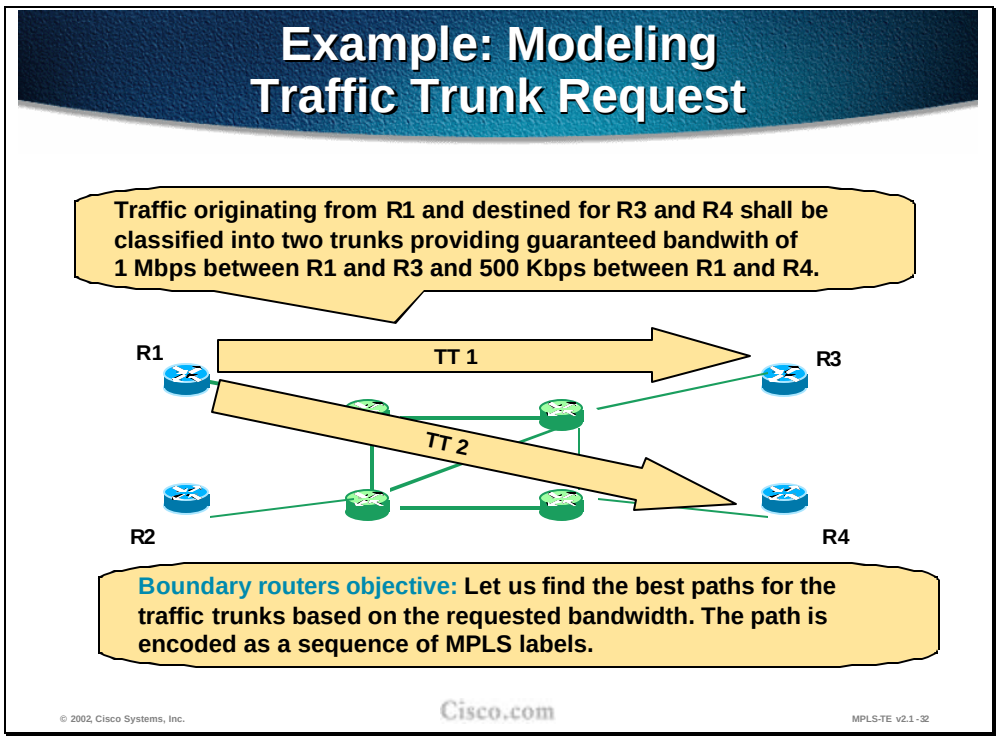
Network Links and Link Attributes



At the heart of MPLS Traffic Engineering is the ability to define trunks through the network, each with an assured amount of bandwidth.

Information must be given to the MPLS processes to create and define the Label Switched Path through the network. This information may come from an explicit configuration (manually defining a fixed LSP) or from a dynamic path-assignment process.

In order to dynamically provide the MPLS LSP that provides a guaranteed bandwidth, information must be gathered from around the network about the state of the network and the bandwidth available on the individual links in the network. Therefore link resource information must be sent to the routers terminating the Traffic Trunks so they can calculate a LSP that will provide the level of bandwidth required.



An example network is shown in the figure.

In this example R1 is carrying traffic destined for the other side of the network. Specifically, some traffic is destined for R3 and some for R4.

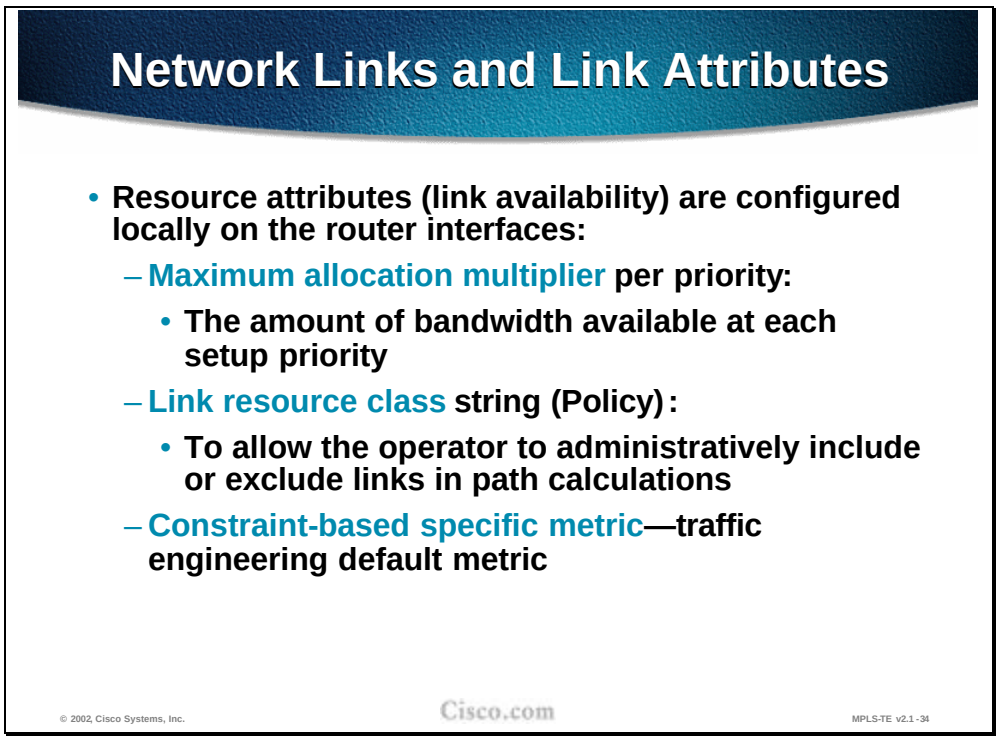
The traffic profiles identified have shown a requirement for a minimum bandwidth of 1Mbps from R1 to R3 and 500kbps from R1 to R4.

In order to carry this defined traffic across the network, two Traffic Trunks are required. R1, which is the head-end router, must create these two trunks. In order to do this, R1 must define the LSP for each trunk through the core of the network and assign the appropriate MPLS labels to the trunks (and therefore to the data using those trunks).

R1 must collate information about the network and then issue the request for building the trunks over the appropriate LSPs.

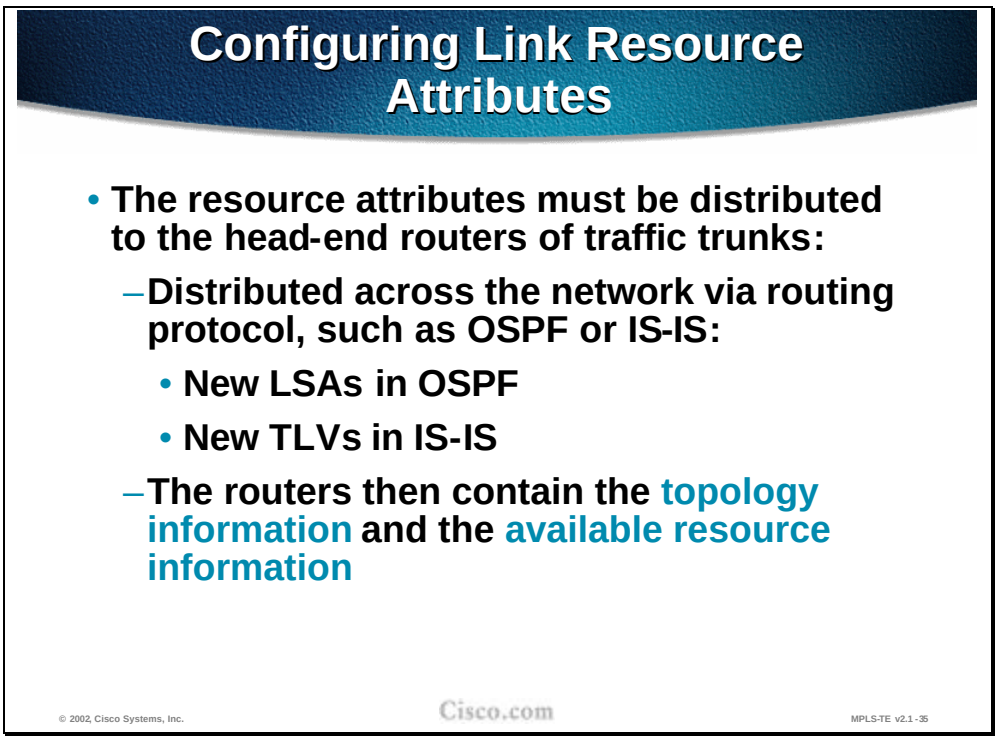
There are various processes, which may occur in the lifetime of a Traffic Trunk:

- Establish: Creating a Traffic Trunk by deciding on the LSP through the network, assigning MPLS labels and, most importantly, assigning resources to the Trunk.
- Activate: Causing data to start to use the Traffic Trunk by using some routing function, which directs traffic into the Trunk.
- Deactivate: Stopping data from using the Traffic Trunk by again using a routing function to cease the direction of data into the Trunk.
- Modify Attributes: Changing the characteristics of the Traffic Trunk (such as its available bandwidth).
- Re-route: Choosing a new path for the Traffic Trunk (most probably because of some failure in the network, or a recovery from a failure).
- Destroy: Removing the Traffic Trunk completely by reclaiming the resources allocated and perhaps the MPLS labels.



In order for the Trunk to dynamically discover its path through the network, the head-end router must be provided with information on which to base this calculation. Specifically it needs to be provided with:

- The amount of bandwidth available on each link in the network (Maximum Allocation Multiplier). Because there are priority levels for Traffic Trunks, the availability information must be sent for each priority level for each link. Including priority levels means the path decision mechanism is given the opportunity to choose a link with some bandwidth already allocated to a lower priority Trunk, forcing that lower priority trunk to be 'bounced' off the link.
- For administrative reasons, the network operator may decide some Trunks are not permitted to use certain links. To accomplish this, for each link, a 'Link Resource Class" must be defined and advertised.. The definition of the Trunk may include a reference to particular 'Affinity bits'. The Trunk Affinity bits is matched against the Link Resource Class to determine if a link may or may not be used as part of the LSP.
- Each link has a cost or metric for calculating routes in the normal operation of the IGP. It may be that, when calculating the LSP for Traffic Trunks, the link should use a different metric. Hence a 'Constraint-Based Specific Metric' may be specified.



The router at the headend of the Trunk, which is the router initiating the Trunk, must be provided with resource information for each link in the network. This headend router could potentially pick any path through the network and must know the status of every link in the network.

This knowledge is achieved only through the use of a Link-State protocol such as Integrated IS-IS or OSPF, as only this type of protocol floods information about all links to all routers.

- IS-IS has new Type-Length-Value (Type 22 TLV) fields to append this information to it's Link-State PDU advertisements
- OSPF has new Link-State Advertisement (Type 10 LSA) definitions to distribute this information

Once this information is included in the IGP advertisements and those advertisements are received by the head-end router, that router has information about the *network topology* (as it would have had in normal IGP routing) but also about the *available network resource* information, which is needed to calculate paths satisfying its Trunk requirements.

Practice

Q1) What is communicated with the Link Resource Attributes?

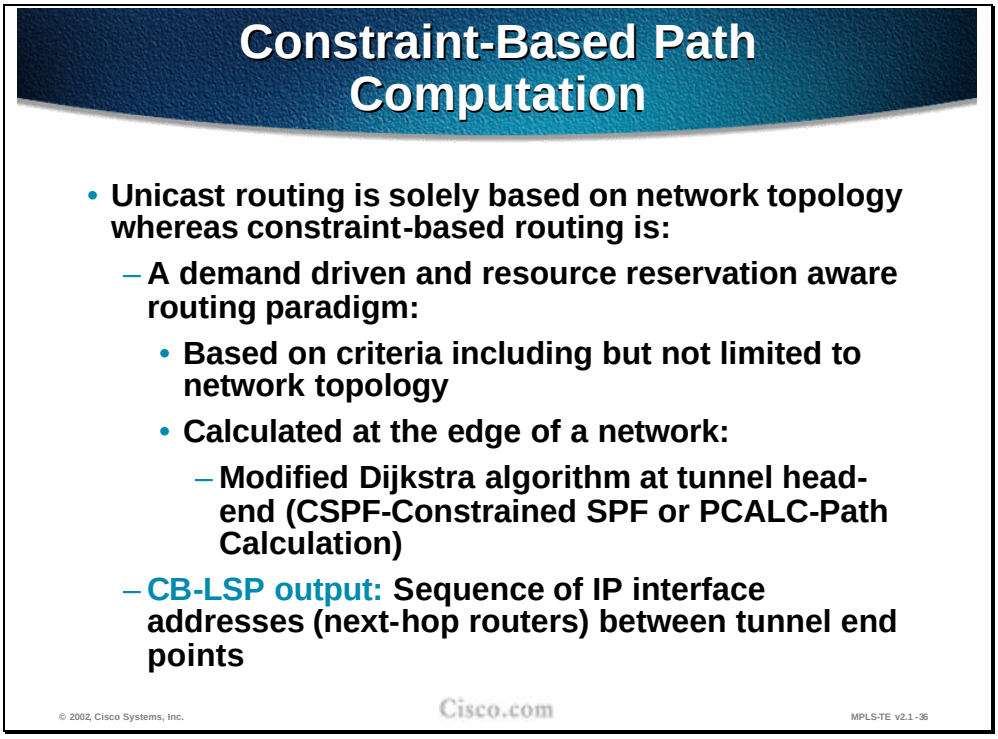
A) Link Resource information replaces the old and inferior IGP Link-State attributes.

B) The new extended metric for best-path calculation.

C) The routers initiating the Traffic Trunk request must be provided with the information on the available resources in the network.

D) Link Resource information is sent to the neighboring routers to calculate the best path for the routed IP traffic based also on the currently available bandwidth.

Constraint-Based Path Computation



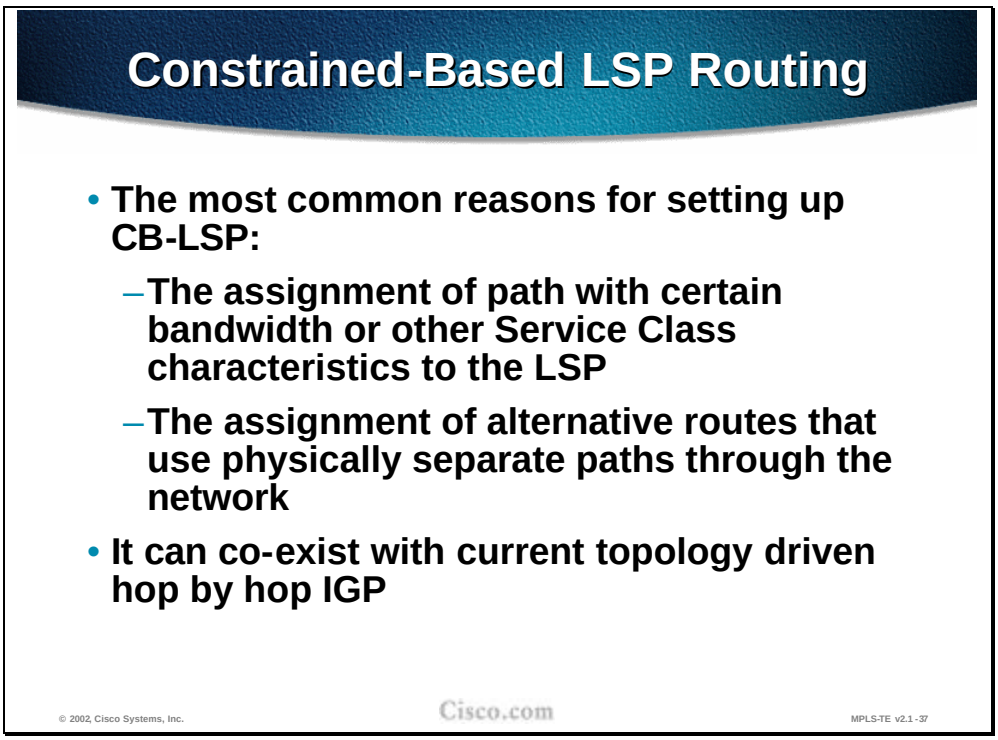
In traditional networks, the IGP calculates paths through the network based on the network topology alone. Routing is destination-based and all traffic to a given destination from a given source will use the same path through the network. That path is determined based simply on what the IGP regards as the 'least cost' between the two points (source and destination).

A Constraint-Based routing as the most often used term is in some situations also referred as a Constrained SPF (CSPF) calculation or a Path Calculation (PCALC).

Constraint-Based routing:

- Augments the use of link 'cost' by also considering other factors such as bandwidth availability or link latency when choosing the path to a destination.
- Tends to be carried out at the edge of the network, discovering a path across the core to some destination elsewhere at the other edge of the network. Typically this discovery uses the Constrained SPF (CSPF) calculation (a version of the 'usual' SPF used by IS-IS and OSPF, but considering other factors besides cost such as bandwidth availability.)
- Produces a sequence of IP addresses corresponding to the routers used as the path to the destination; the next-hop addresses for each stage of the path.

The consequence of Constraint-Based routing is that, from one source to one destination, many different paths could be used through the network depending on the requirements of those data flows.

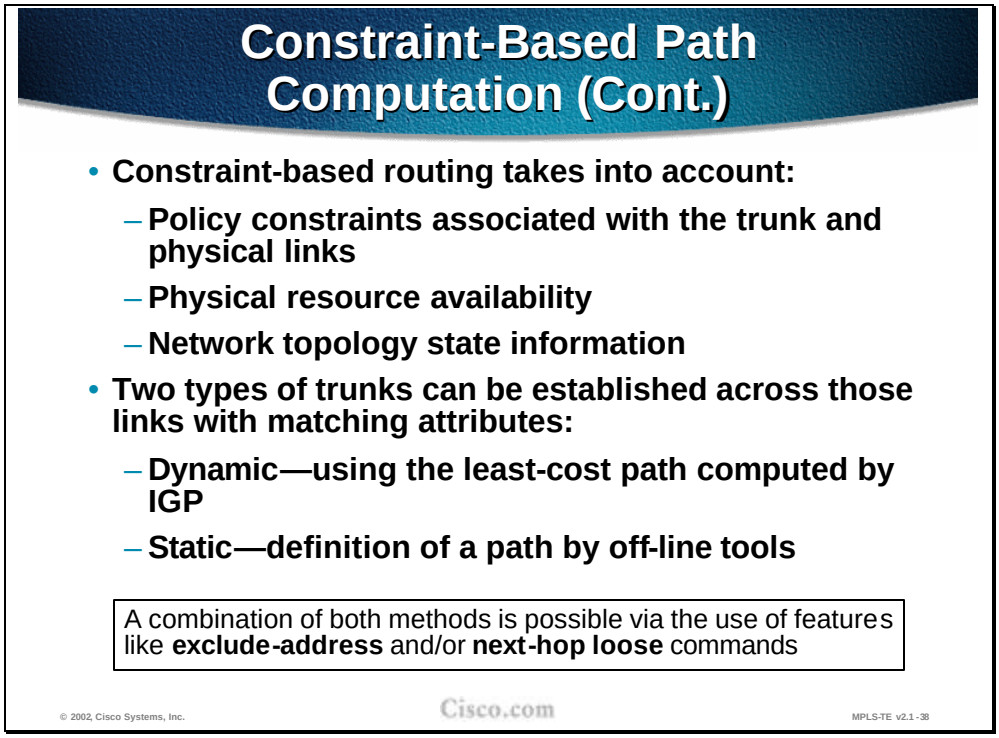


Constraint-Based routing is used typically:

- To allow the network to assign particular paths for particular data flows, assigning many different paths from one source to one destination, based on the requirements of those data flows.
- To allow the network to create physically separate paths through the network in order to provide resilient or alternate routes.

Of course the information to calculate these paths is provided *in addition to* the 'normal' link costs/metrics so that Constraint-Based and Destination-Based (hop-by-hop) routing can co-exist happily on the same network.

Constraint-Based routing requires a Link-State protocol (IS-IS or OSPF) so information about all links is flooded to all routers in the network.

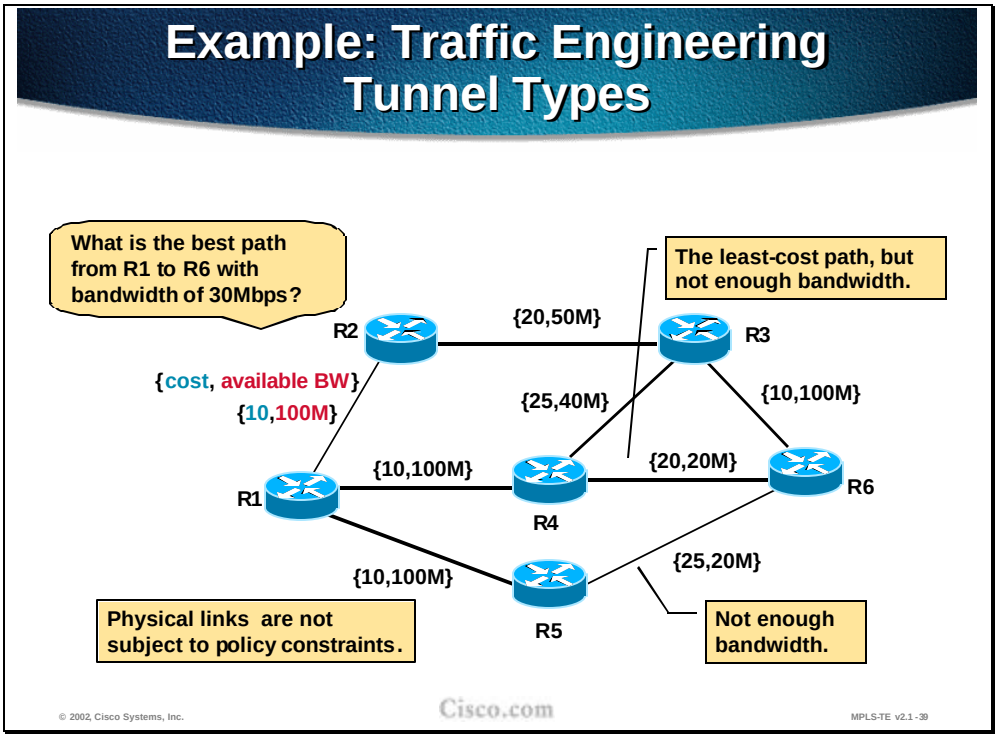


When choosing paths through the network, the Constraint Based routing system takes account of:

- The topology of the network, including information about the state of the links (the same information used by normal hop-by-hop routing)
- The resources available in the network, such as the bandwidth not already allocated on each link and at each of 8 priority levels (priority 0 to 7).
- The requirements placed on the Constraint-Based calculation defining the policy or the characteristics of this Traffic Trunk

Of course Constraint-Based routing is a dynamic process, responding to a request to create a path and calculating (or re-calculating) the path based on the status of the network at that time. Alternatively, the path taken by a Traffic Trunk can be defined statically by the operator.

By using commands like **exclude-address** or **next-hop loose** in the explicit path configuration, the operator can mix static and dynamic computation.

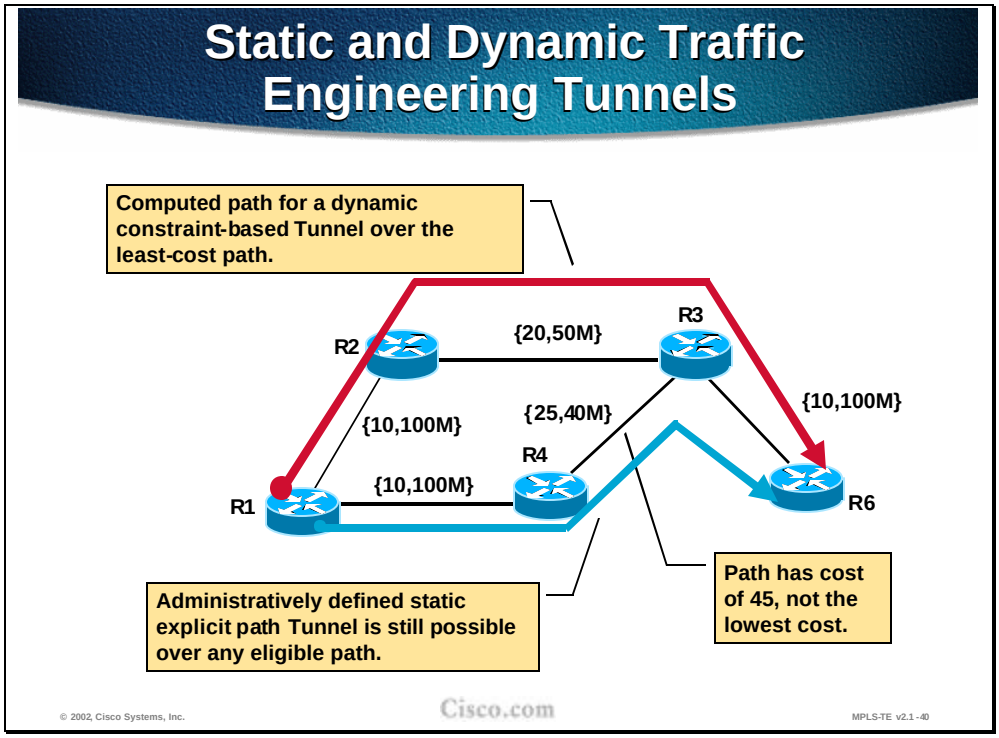


An example network is shown in the figure. Each link specifies a link cost for metric calculation and a bandwidth available for reservation, such as a metric of 10 and an available bandwidth of 100Mbps for the link between R1 and R2. Other than these criteria, no links are subject to any policy restriction disallowing their use for creating Traffic Trunks.

The requirement is to create a Trunk from R1 to R6 with a bandwidth of 30Mbps.

Based simply on the link costs, the least cost path from R1 to R6 is R1-R4-R6 with a cost of 30. However the link from R4 to R6 has only 20Mbps of bandwidth available for reservation and therefore cannot fulfill the requirements of the Trunk.

Similarly, the link R5-R6 has only 20Mbps available so no paths can be allocated via R5.



The diagram now shows only those links, which can satisfy the requirement for 30Mpbs of available bandwidth.

Over this topology, two Trunk paths are shown:

- The path colored blue (R1-R4-R3-R6) has been defined statically by the administrator. Had the administrator attempted to define a path that did not have the required free bandwidth, the trunk establishment would have failed. This trunk does indeed fulfill the minimum bandwidth requirement. However, adding the link costs gives a total of 45, which is not the lowest cost possible.
- The red (upper) path shows the result of a dynamic Constraint-Based path calculation. The calculation has ignored any links which do not satisfy the bandwidth requirement (those from the last diagram not shown in this diagram, such as the connections to R5) and then run a Contrainted Shortest-Path-First (CSPF) calculation on what remains. This calculation has yielded the path R1-R2-R3-R6 with a path cost of 40.

Practice

Q1) What is a result of a Constraint-based path calculation?

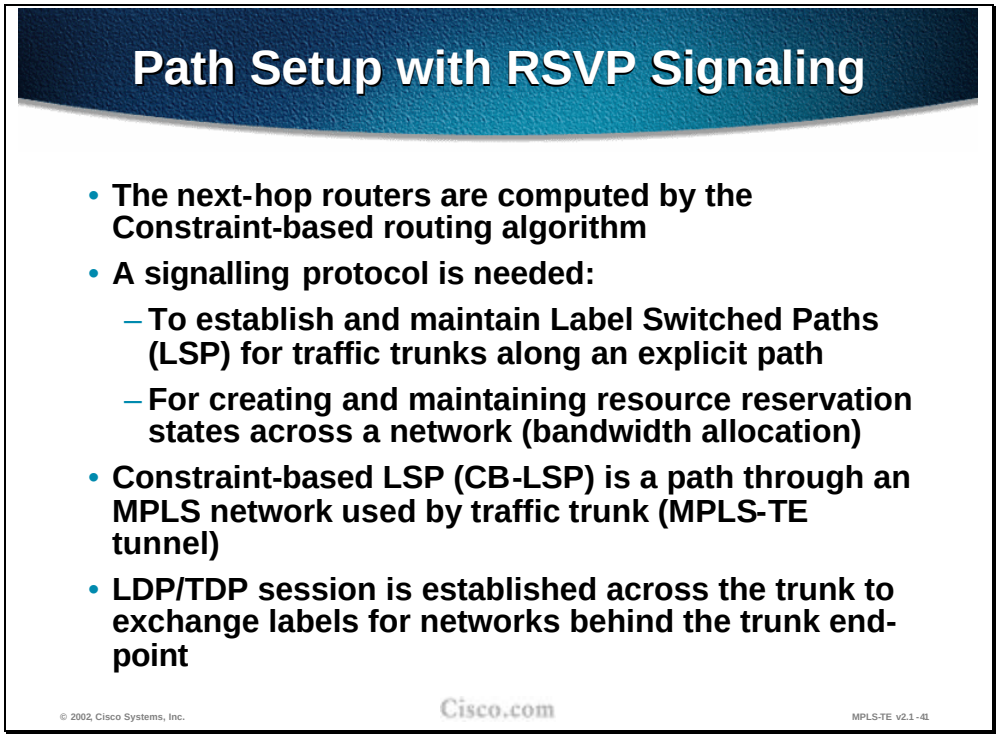
A) The result is a list of IP next-hop address with associated MPLS labels between the tunnel endpoints.

B) The LSP is specified with the list of IP addresses (source-addresses) between the tunnel endpoints.

C) The result is a list of MPLS labels between the tunnel endpoints.

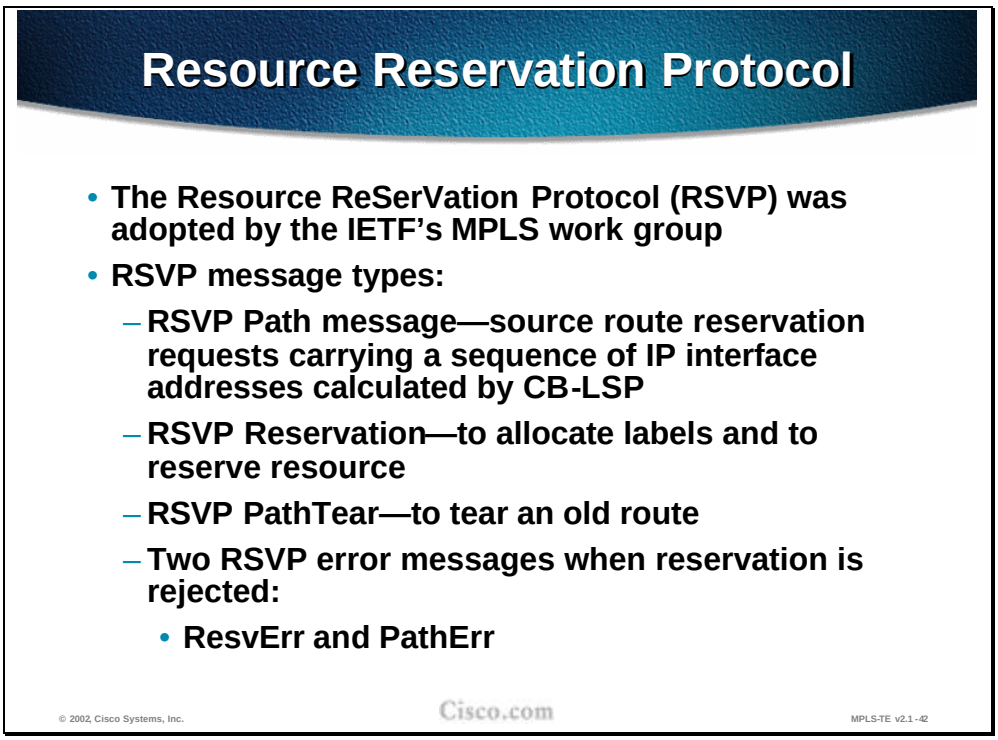
D) The LSP is specified with the list of IP addresses (next-hops) between the tunnel endpoints.

Path Setup with RSVP Signaling



The result of the Constraint-Based calculation is a list of routers, which form the path to the destination. The path is a list of IP addresses identifying each next-hop along the path.

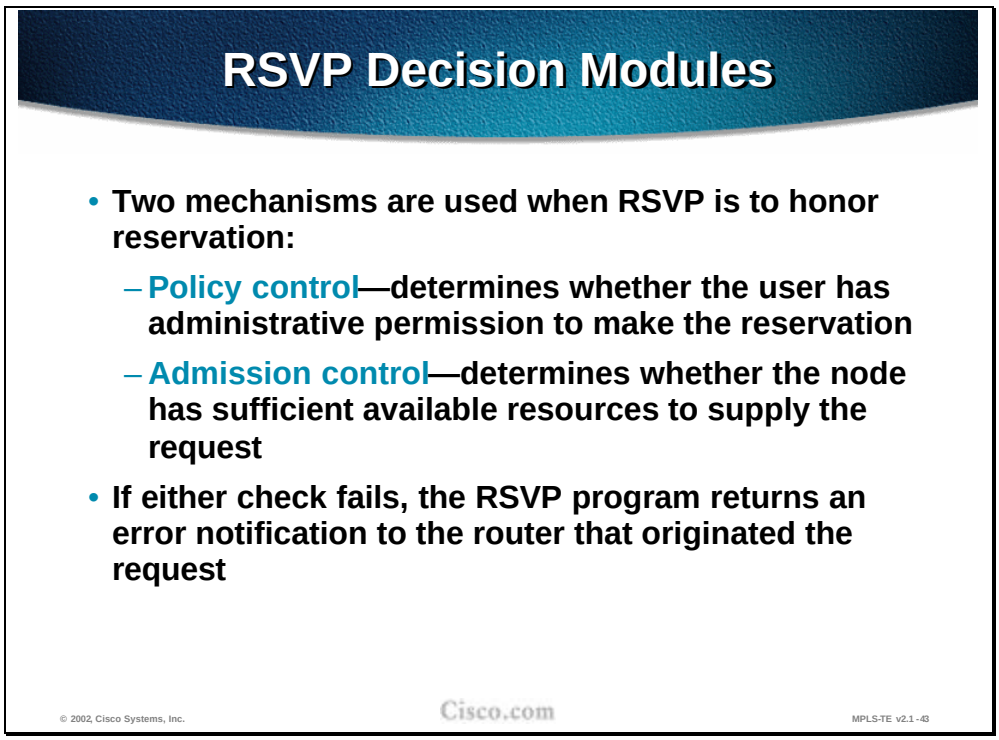
However, this list of routers is known only to the router at the head-end of the trunk attempting to build the tunnel. Somehow, this now explicit path must be communicated to the intermediate routers. It is not up to the intermediate routers to make their own Constrained SPF calculations: they merely abide by the path provided to them by the head-end router. Therefore some signaling protocol is required to confirm the path, check and apply the bandwidth reservations and finally to apply the MPLS labels to form the MPLS Label-Switched-Path through the routers. RSVP is used to confirm and reserve the path and LDP/TDP is used to apply the labels.



The Resource reSerVation Protocol (RSVP) is specifically designed to allow applications to reserve bandwidth in a network. Therefore it is an obvious candidate to perform the path confirmation and reservation in MPLS Traffic Engineering and has been adopted as such by the MPLS working group of the IETF.

RSVP operates by using the following messages:

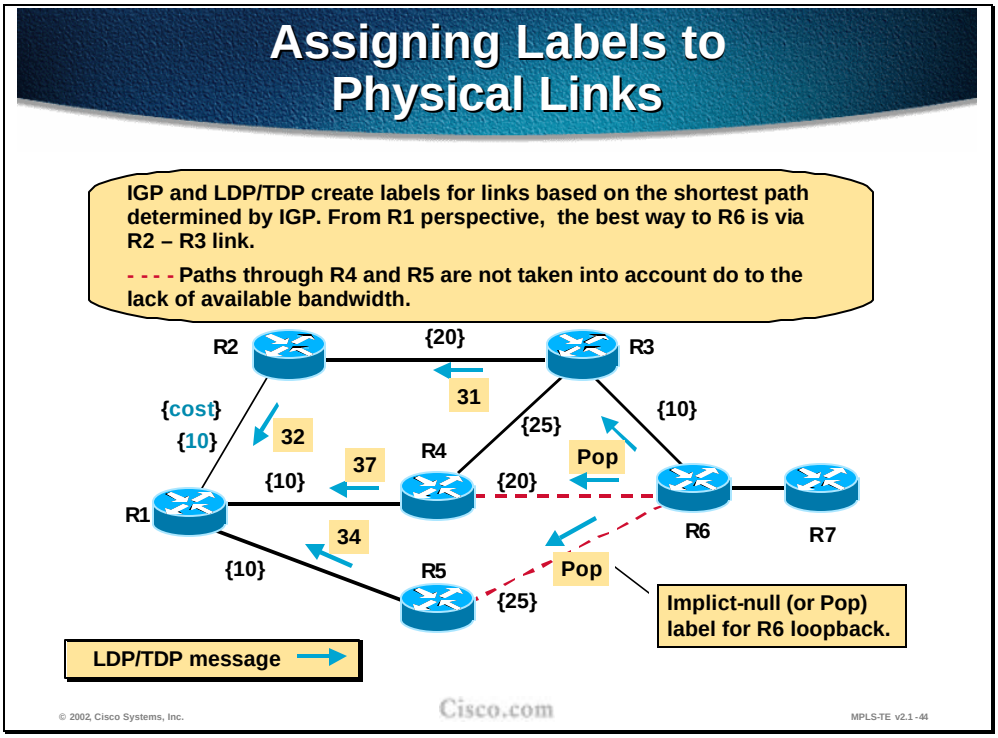
- RSVP PATH message is used to trace the path through the network, checking the resource availability at each stage and storing the path as it goes.
- RSVP RESV (RESerVation) message is sent (by the far end router) in reply to a PATH message to confirm the path and reserve the bandwidth on each router in the path.
- RSVP PATH_TEAR message tears down a reservation and releases the bandwidth allocation so it can be used again.
- During the PATH/RESV stage, the reservation could fail and lead to a PATH_ERR or RESV_ERR message being generated.



Part of the process of RSVP is to confirm whether the reservation is acceptable at each router along the path. This task is completed with the following checks:

- Policy Control: Checks whether the initiator of the RSVP request has the administrative privilege to make the reservation. This is more specific to generic RSVP where a request may be made by a host system (typically a multimedia application such as video or audio streaming). In the case of MPLS Traffic Engineering, the request should be arriving from the head-end router. .
- Admission Control: Checks whether the resources are still available to satisfy the reservation request. This is where the reason for the Constraint-Based path calculation becomes clear. Because the available resources have, in effect, been checked in advance (by the Constraint-Based path calculation), the reservation should be successful on this count.
- The reservation may not be successful due to the batched link-state routing advertisements, Some resources that are being just reserved by other traffic trunk might still be available to the router initiating a traffic trunk request.

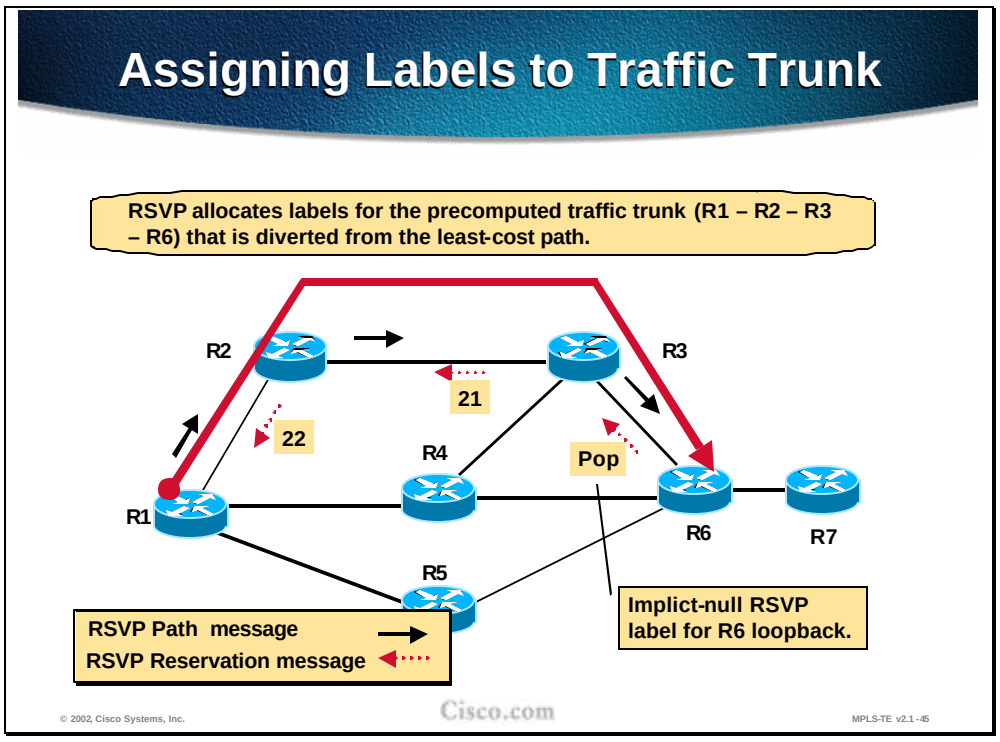
If either check fails then the reservation will be refused. A PATH_ERR would be sent if the reservation failed while the PATH part of the process was in process (because the request cannot be satisfied by one of the routers in the path). In theory, as the PATH message checks that the resource is available to be reserved on the way out, the RESV message should be accepted automatically on the way back. However, situations can arise where the RESV is the part that fails, in which case a RESV_ERR message is generated. A PATH_TEAR message follows a PATH_ERR or RESV_ERR message to tear down any remaining parts of the path.



The diagram in the figure shows a sample network based on the earlier example. This time, only the link costs (as per the IGP) are shown for each link.

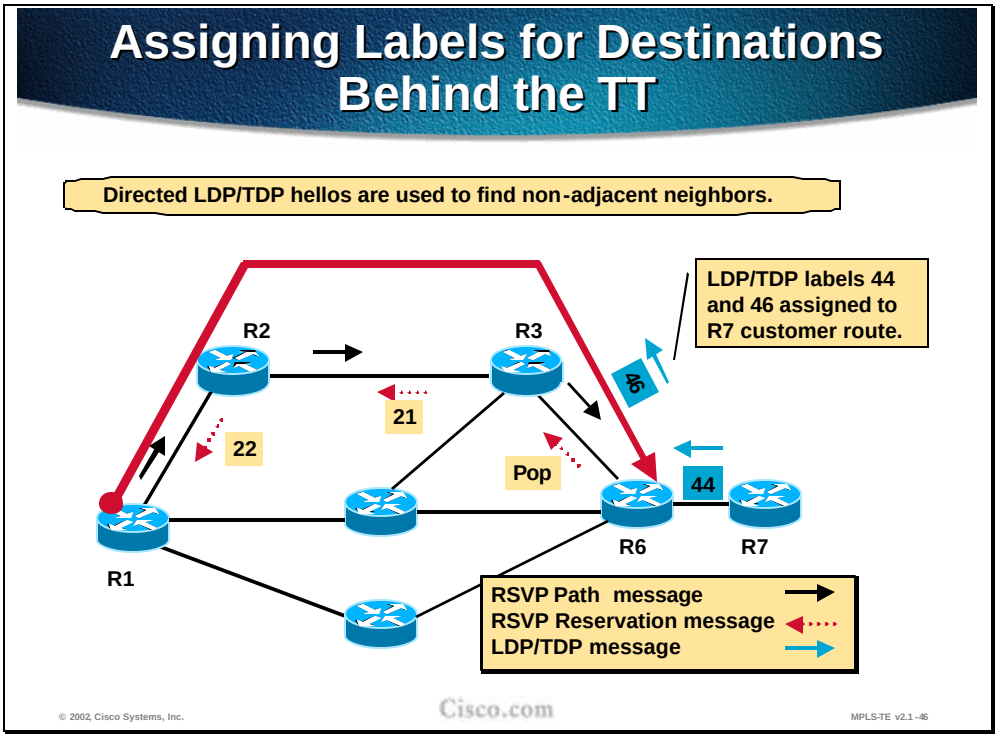
The diagram shows the interaction between the IGP and the Label/Tag Distribution Protocol. Using information from the IGP, LDP/TDP messages are sent from R6 to R1, assigning labels as they go. At R1, the least-cost path and the labels corresponding to that path are selected.

One interesting ‘label’ shown is ‘Pop’. ‘Pop’ signifies that the next router in the path is the end of this particular MPLS Label-Switch Path and that the packet should ‘pop’ back up from the MPLS layer to the routing layer.(‘Pop’ is a programming term used to ‘pop’ items off a stack of stored items. Here it is used to ‘pop’ one set of MPLS information off the MPLS label stack, and in this case leaving no labels on the stack, therefore returning the packet to the routing layer)



RSVP works by sending out PATH messages to establish the path through the network. In the case of MPLS Traffic Engineering, that path is included in the RSVP Path message either by manually configuring an explicit path or by dynamically calculating the path via CB-LSP. Therefore it is expected that the PATH message will succeed in traversing the network without being rejected along the way. While the RESV message returns along the path, it interacts with MPLS to assign labels as it goes. Again, the last label in the path (the first label allocated by the returning RESV message) is the implicit 'Pop' label to signify this is the destination router for the MPLS-encapsulated packets.

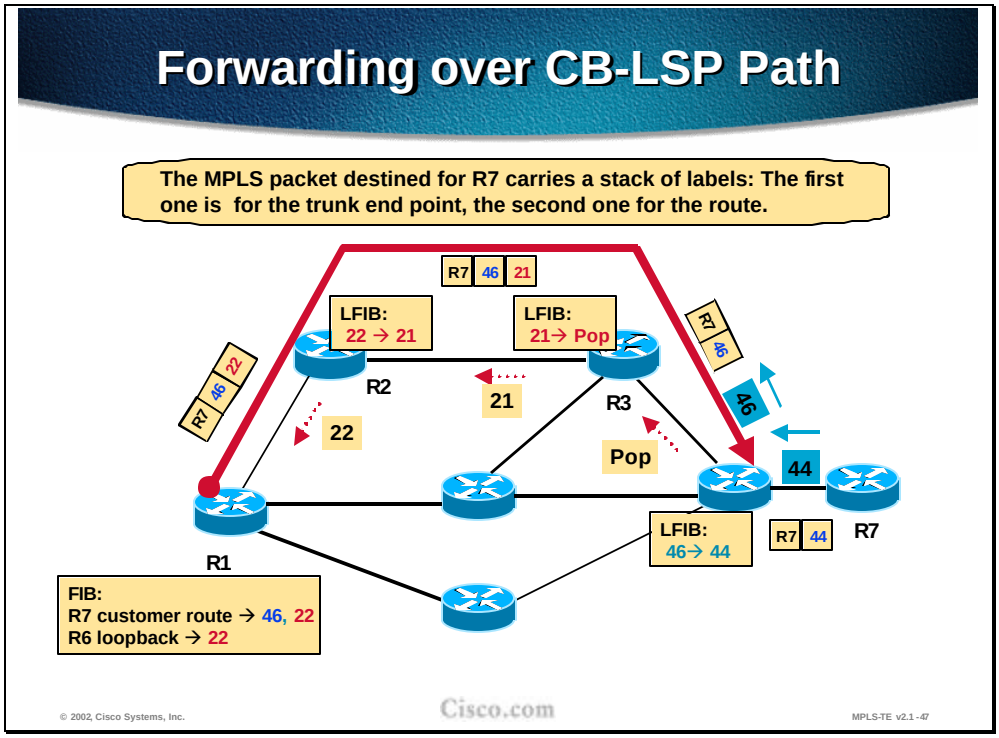
Therefore, when the RSVP reservation is completed (the RESV message arrives at the source router), the MPLS LSP is also completed.



A new item in this network diagram are routes coming from R7. This router could be a customer router attached to the ISP network (R1 to R6).

It would be possible to route packets through MPLS up to R6, 'pop' them back into the IP layer and then route normally into the customer network. However, it makes sense to try to keep the packet inside the level-2-switched MPLS layer.

In order to achieve this, further LDP/TDP hello messages are sent explicitly along the path out of the end of the Traffic Trunk (R6) and into the customer network (R7). These hellos create extra MPLS labels for the last part of the path defining the route into the customer network inside MPLS. A label stack (of two labels) will be required to reach the customer network from the R1 router.



To route into the customer network inside MPLS, a stack of labels is created:

- The first, top-most label, of the label stack (label 22 at R1) defines the path inside the ISP network (the RSVP LSP identified in the previous diagrams).
- When this top-most label is ‘pop’ped off the label stack (at R3), another label comes to the top of the label stack. . This second label identifies the label into the customer router (label 46 at R1).
- As the ‘pop’ happens at R3, the MPLS label (the second label) for the customer route must be defined between R6 and R3. On R6 it may refer to another MPLS label in the customer network (as in this case) or alternatively be ‘pop’ped to arrive natively at R7 itself.

Traffic destined for R6 itself would have only the top-most label in the label stack (label 22 at R1). ‘Pop’ping this label off the stack at R3 leaves an empty MPLS label stack at R6 and therefore the packet reverts to the IP layer on the link ? R3-R6, and arrives at R6 as an IP packet ready to be routed.

Practice

Q1) Explain the role of RSVP in MPLS-TE.

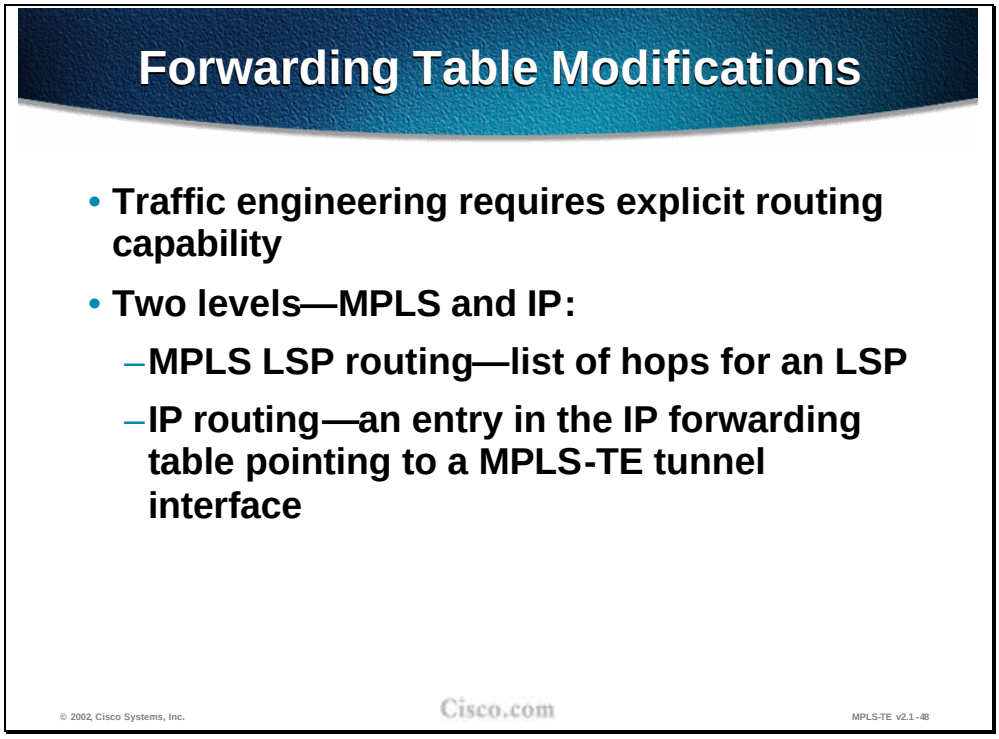
A) RSVP interacts with IGP to allocate labels and reserve resources for the Traffic Trunk.

B) RSVP is used in LSP path signaling to ensure the label allocation and bandwidth reservation.

C) RSVP is used in LSP path signaling only to ensure bandwidth reservation.

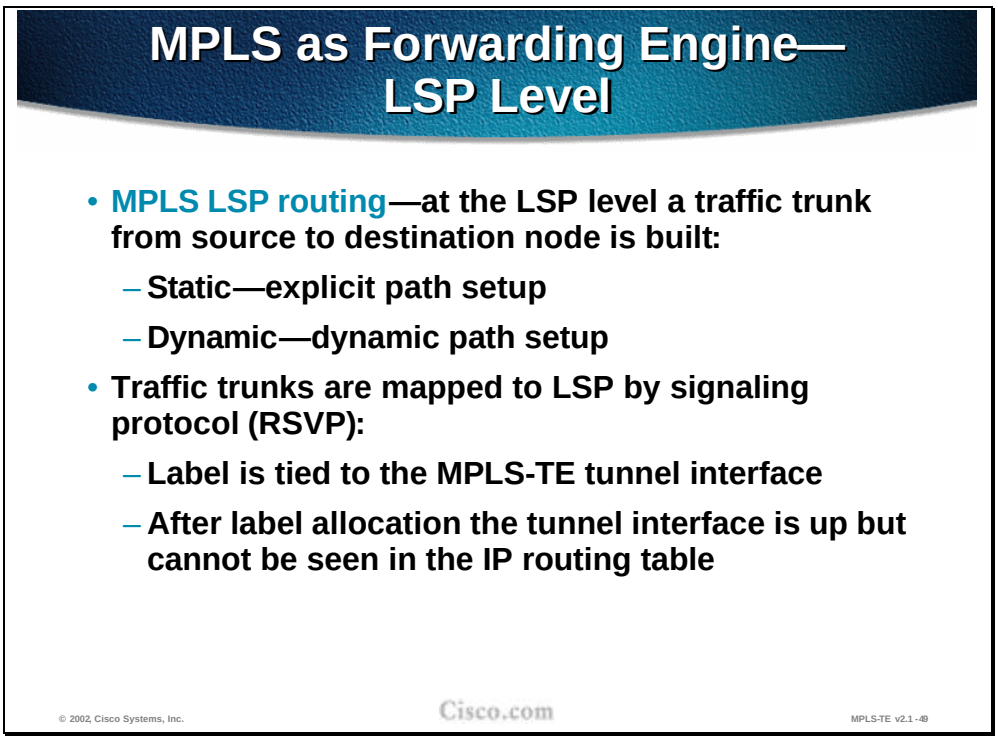
D) RSVP is used to compute a list of IP next-hop address between the tunnel endpoints.

Forwarding Table Modifications



In order to use the traffic engineered tunnels some modifications must be made to the forwarding tables and to the mechanisms they are built with. Explicit routing capability is required at the MPLS level and at the IP forwarding level as well:

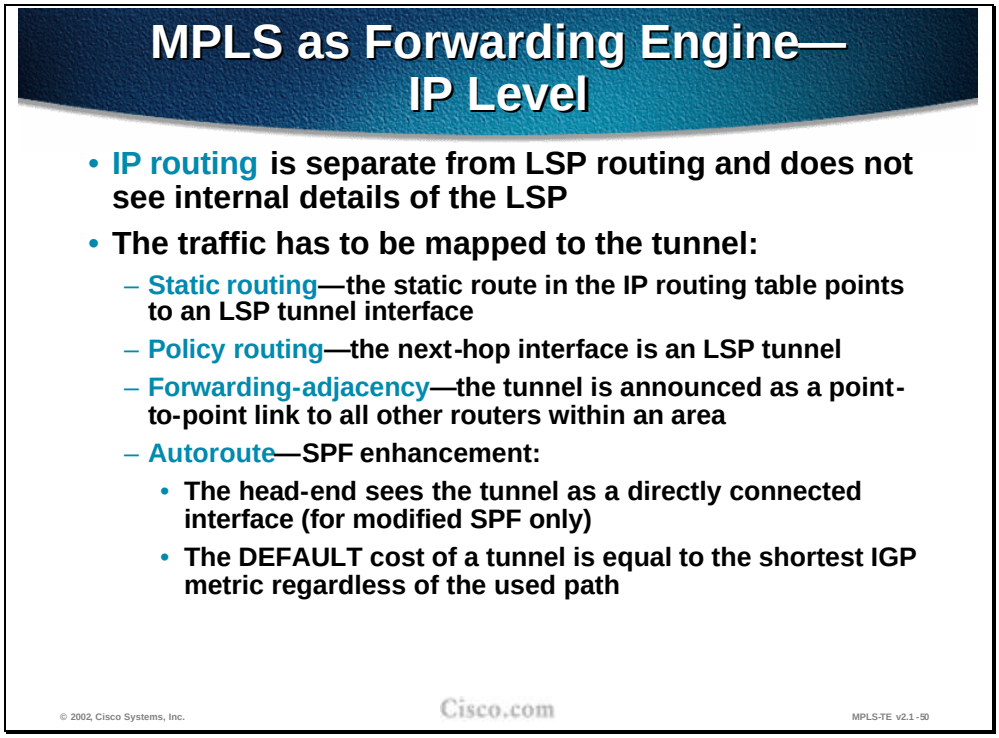
- The MPLS LSP routing requires the list of hops for an LSP (explicit path).
- For IP routing, an entry in the IP forwarding table has to point to the MPLS-TE tunnel interface. This tunnel follows the established MPLS LSP.



For a traffic trunk, an LSP path must be built from the source to the destination (from the traffic trunk head-end to its endpoint, tail-end). The LSP path can be:

- Statically defined (manually defining a list of hops towards the destination)
- Dynamically built (by using constraint-based path computations)

The traffic trunks are mapped to the LSP using the signaling protocol (RSVP). With label allocation to the MPLS-TE tunnel interface at the head-end of the trunk, the tunnel comes up but does not appear in the IP routing table. The traffic engineered tunnel itself does not appear in SPF calculations for the destinations behind the trunk tail-end.



The tunnel normally does not appear in IP routing table.The IP routing process does not see the tunnel so the tunnel is normally not included in any SPF calculations. The IP traffic can be mapped onto a tunnel in three different ways:

- Using static routes that point to the tunnel interfaces.
- Using policy based routing and set the next hop for the destination to the tunnel interface.
- Using forwarding-adjacency the tunnel will be announced via OSPF or ISIS like any other Unidirectional Link (UDL). In order to be used for data forwarding such a tunnel has to be set up bidirectionally.
- Using the autoroute feature, which is an SPF enhancement that includes the tunnel interface into the route calculation as well. The result of the autoroute feature is that the tunnel is seen at the head-end (and only there) as a directly connected interface. The metric (cost) of the tunnel is set to the normal IGP metric from the tunnel head-end to the tunnel end-point (over the least cost path, regardless if the tunnel is actually using the least cost path or not).

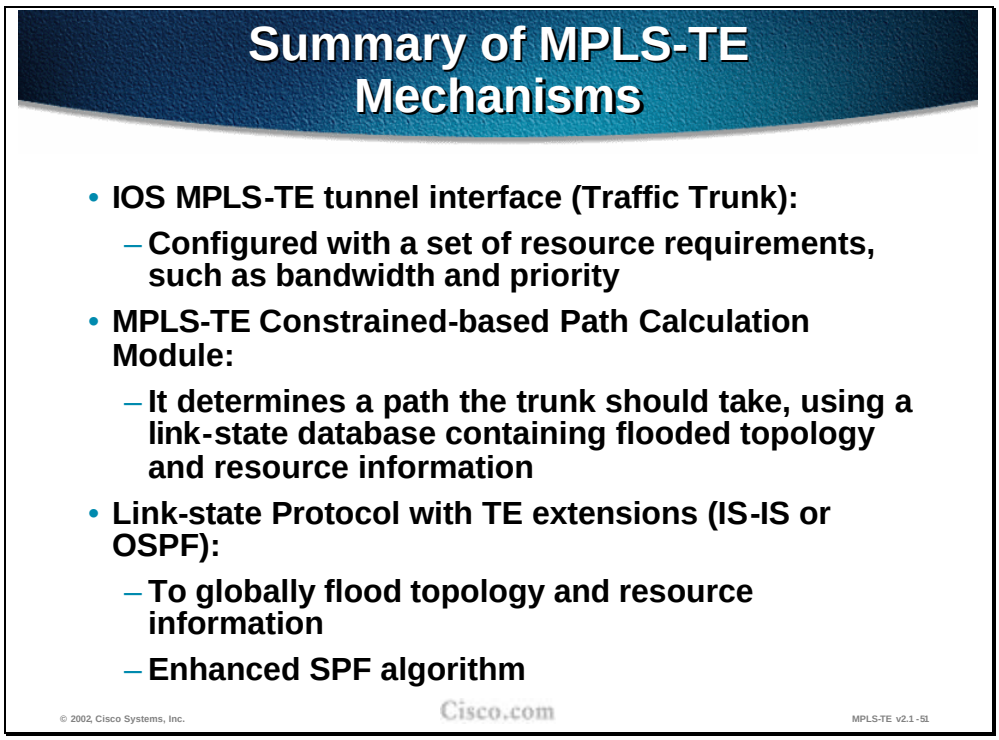
Note With the autoroute feature, the traffic engineered tunnel appears in the IP routing table as well but this appearance is restricted to the tunnel head-end only.

The first two options are not very flexible or scalable. The traffic for each destination that needs to use the tunnel must be manually mapped to the tunnel.

For example, when using static routes, the tunnel is used only for the explicit static routes. Any other traffics not covered by the explicit static routes, including traffic

for the tail-end router (even though the tunnel terminates on it) will not be able to use the tunnel, instead, it will follow the normal IGP path.

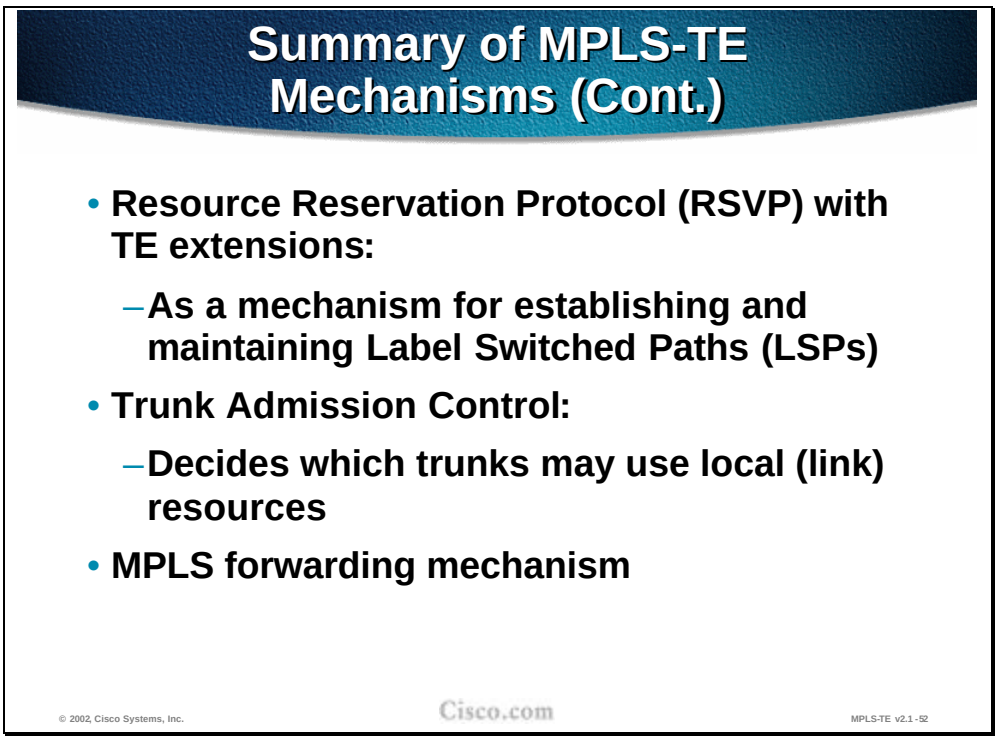
Note The autoroute and forwarding-adjacency features are explained in details in Assigning Traffic to Traffic Trunks lesson.



Overall, the MPLS-TE mechanisms include several components that interact in a complex yet effective way to provide the engineered tunnels across the MPLS enabled networks.

The main component of MPLS-TE is the **MPLS-TE tunnel interface** itself which is the Traffic Trunk (TT), and which is configured with a set of resource requirements including the required bandwidth and priority.

The **Constraint-based Path Calculation** determines the path (Label Switched Path, LSP) the trunk should take using the link-state database that contains the resource information. The resource information is flooded throughout the network with modified link-state Interior Gateway Protocols (IGP) that include resource information in their link-state updates. There are two routing protocols with TE extensions: Integrated IS-IS and OSPF. The SPF algorithm is modified as well to take into account the resource information when calculating the LSP path.



The computed LSP path must be established using a signaling protocol. The Resource reSerVation Protocol (RSVP) with TE extensions is used:

- To reserve the required bandwidth.
- To establish and maintain the MPLS labels for the LSP.

The bandwidth reservation is done via the Trunk Admission Control that decides which trunks may use link resources if available.

After the LSP path is established and MPLS labels allocated, the MPLS forwarding mechanism ensures that the traffic mapped onto the tunnel is forwarded along the LSP path.

Practice

Q1) How is traffic mapped to the MPLS-TE tunnel?

A) By manually turning on CSPF under the IS-IS or OSPF configuration.

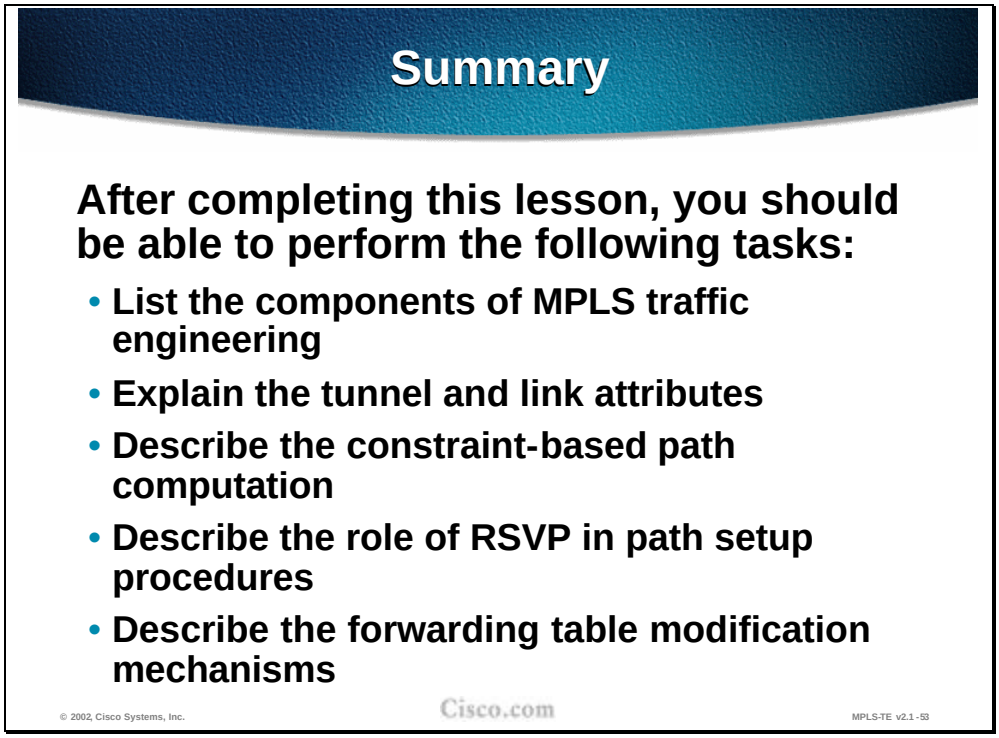
B) The MPLS-TE tunnel is by definition represented as a routing interface.

C) Statically using static routes, with policy routing, or dynamically using the autoroute feature.

D) By enabling any IGP protocol over the tunnel.

Summary

This section summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Constraint-based Path Computation

Lesson Review

Instructions

Answer the following questions:

1. What are the characteristics of a traffic trunk?
2. What modifications are needed to the IGP to support MPLS-TE?
3. What is a result of Constraint-based path calculation?
4. Explain the role of RSVP in MPLS-TE.
5. How is traffic mapped to the MPLS-TE tunnel?

Constraint-Based Path Computation

Overview

This lesson describes the details of link attribute propagation with an IGP protocol and constraint-based path computation.

Importance

This lesson is a mandatory for the students planning to improve the usage of their network resources with MPLS traffic engineering.

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- Describe the detailed structure of MPLS-TE link attributes
- Explain the role and usability of guaranteed bandwidth sub-pool
- Describe the usability of affinity bits
- Implement MPLS TE constraints with affinity bits
- Avoid usage of links or nodes for MPLS TE tunnels using IP address exclusion
- Describe the propagation of link attributes through an Interior Routing Protocol (OSPF or IS-IS)
- Describe the constraint-based path computation algorithm

- Describe the interaction between link attributes and trunk attributes during the constraint-based path computation

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Cisco Certified Internetwork Professional (CCIP) level of knowledge or equivalent level of IP routing and Cisco IOS knowledge as well as solid understanding of MPLS and link state protocols (OSPF or Integrated IS-IS).

Mandatory Prerequisites:

- AMVS course

Optional prerequisites:

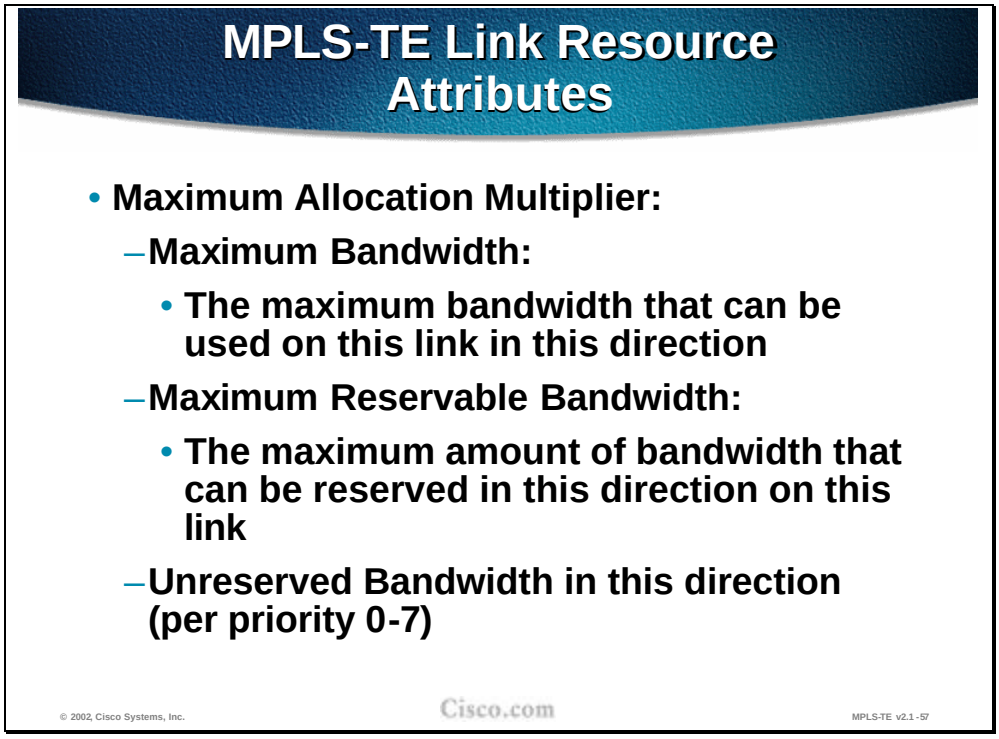
- CISIS course for students deploying MPLS TE in IS-IS environments

Outline

This lesson includes these sections:

- Overview
- MPLS TE Link Attributes
- MPLS TE Trunk Attributes
- Implementing TE Policies with Affinity Bits
- Avoid Usage of Links or Nodes for MPLS TE Tunnels Using IP Address Exclusion
- Propagating MPLS TE Link Attributes with Link-State Routing Protocol
- Constraint-Based Path Computation
- Guaranteed-Bandwidth Sub-Pool
- Summary
- Lesson Review

MPLS TE Link Attributes



The Constraint-based path computation that takes place at the head-end of the traffic engineered tunnel must be provided with several resource attributes before the LSP path is actually determined. These attributes include:

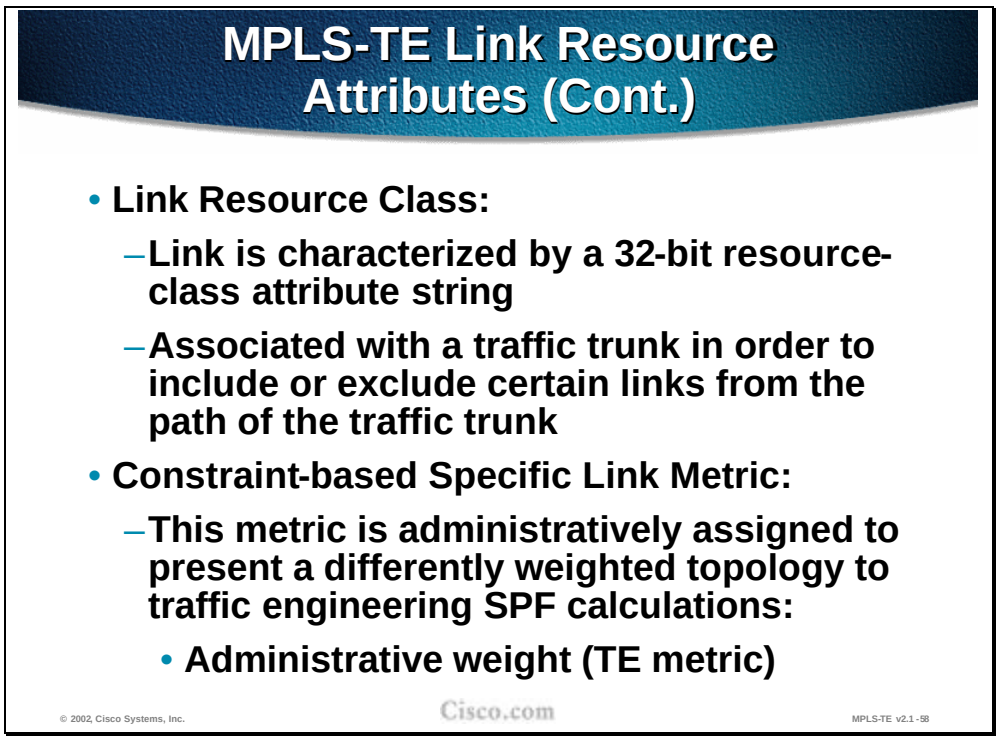
- Link Resource Attributes that provide information on each link's resources.
- Traffic Trunk Attributes that characterize the Traffic Trunk.

Among Link Resource Attributes, the most important is the **Maximum Allocation Multiplier**. This attribute deals with the amount of bandwidth available on the specified link. 'Available' means 'not already allocated' rather than 'presently in use' and is a measure of allocation not utilization. Furthermore, because there are priority levels for Traffic Trunks, this availability information needs to be configured for each priority level on the link. Normally, the bandwidth at the upper priority level is always higher than at lower levels (0-7 levels). Due to over-subscriptions the total amount of bandwidth can exceed the actual bandwidth of the link. There are three components of this attribute:

- Max. Bandwidth provides information on the maximum bandwidth that can be used on the link, per direction, since the traffic trunks are unidirectional. This parameter is usually set to the configured bandwidth of the link.
- Max. Reservable Bandwidth provides information on the maximum bandwidth that can be reserved on the link per direction. By default it is set to 75% of the Max. bandwidth.

- Unreserved Bandwidth provides information on the remaining bandwidth that has not yet been reserved.

Note	Higher priority can preempt lower priority but lower priority can't preempt higher priority.



For each link, another Link Resource Attribute, the Link **Resource Class**, is provided as well. The link is characterized by a 32-bit link resource-class attribute string, which is matched with traffic trunk Resource Class Affinity attribute and allows inclusion or exclusion of the link into or from the path for the trunk.

Each link has a cost or metric for calculating routes in the normal operation of the IGP. It may be that, when calculating paths for Traffic Trunks, the link should use a different metric than the IGP metric. Hence a '**Constraint-Based Specific Metric**', the administrative weight, may be administratively assigned as well.

Practice

Q1) Choose the three major MPLS-TE link attributes that influence the LSP path computation.

A) Maximum Allocation Multiplier (Max. Bandwidth, Max. Reservable Bandwidth for each of the seven priority levels, Unreserved Bandwidth)

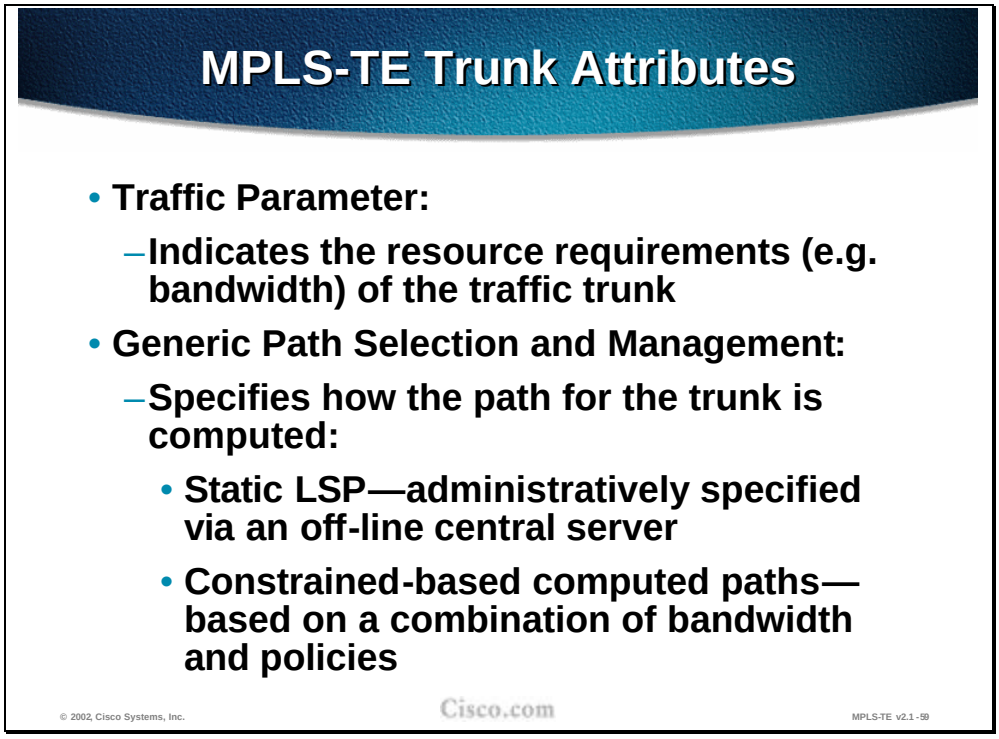
B) Link Resource Class

C) Constraint-based Specific Bandwidth

D) Maximum Allocation Multiplier (Max. Bandwidth, Max. Reservable Bandwidth, Unreserved Bandwidth per priority)

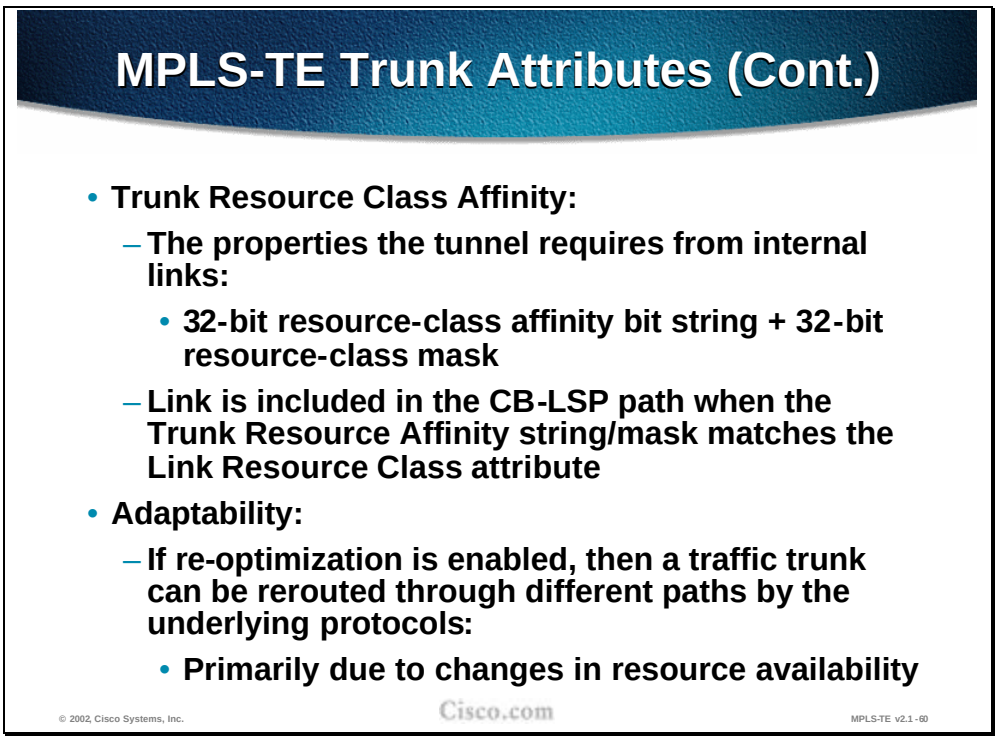
E) Constraint-based Specific Link Metric

MPLS TE Trunk Attributes



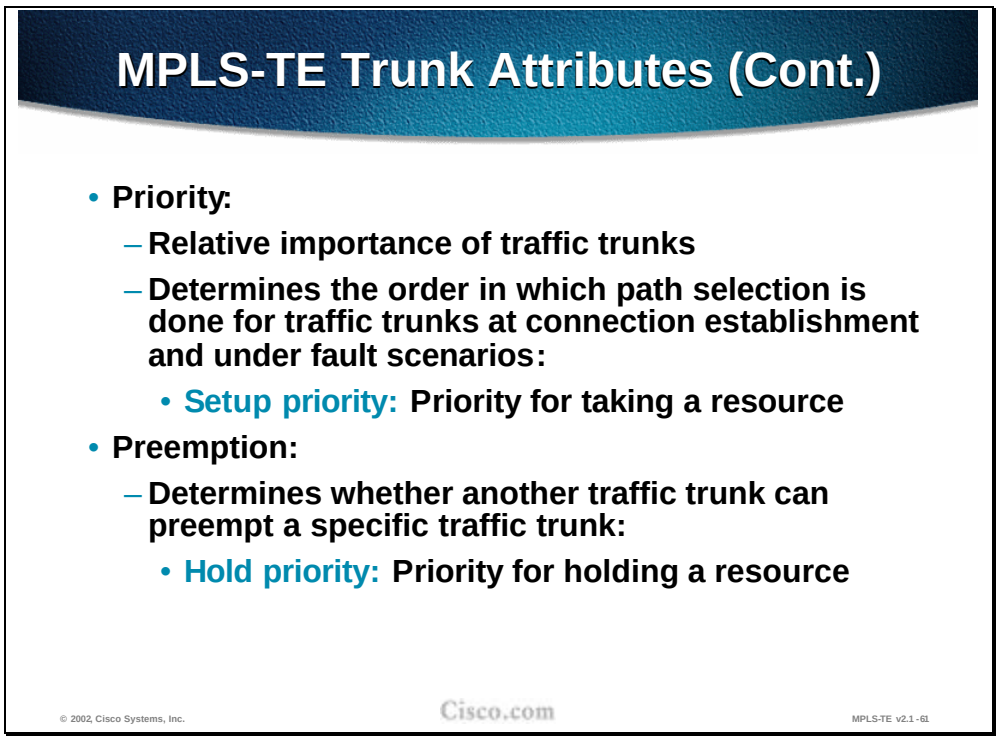
The Traffic Trunk (TT) is characterized by several attributes that affect the path setup and maintenance:

- Traffic Parameter (**Bandwidth**) attributes specify (among other traffic characteristics) the amount of bandwidth required by the Traffic Trunk. The traffic characteristics may include peak rates, average rates, permissible burst size, etc. From a traffic engineering perspective, the traffic parameters are significant because they indicate the resource requirements of the traffic trunk. These characteristics are useful for resource allocation.
- Path Selection and Management attributes (**Path Selection Policy**) specifies the way in which the head-end routers should select explicit paths for traffic trunks. The path can be configured manually or computed dynamically using the Constraint-based path computation, both taking the resource information and policies into account.



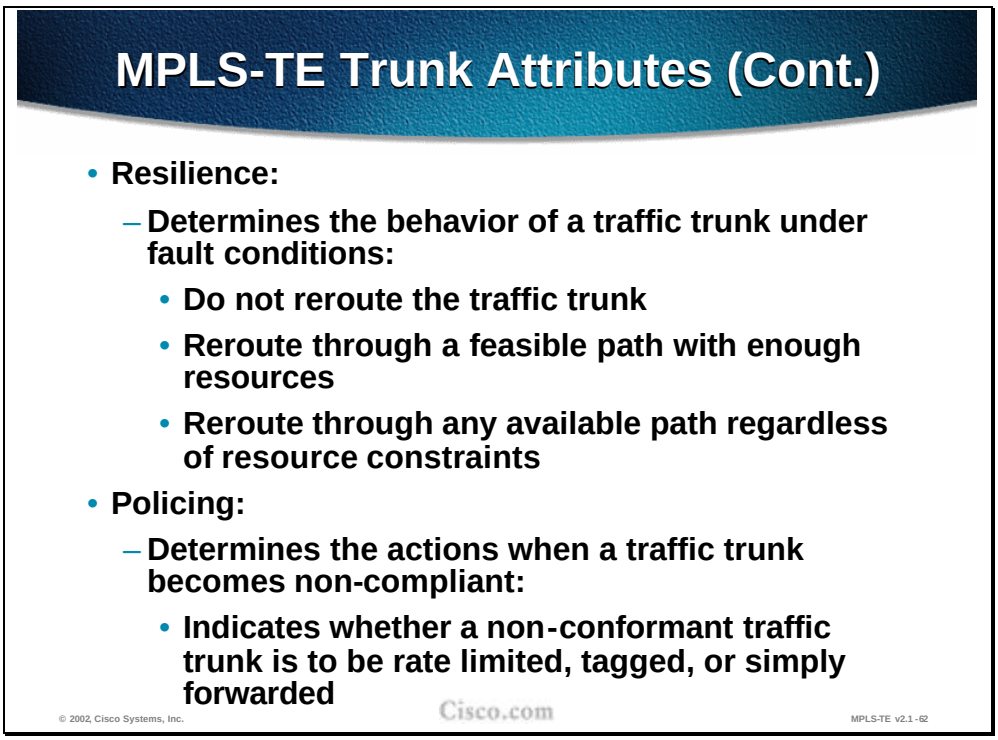
Additional trunk attributes that play a role in LSP path computation and maintenance are:

- The **Resource Class Affinity** attribute allows the network operator to apply path selection policies by administratively including or excluding network links. Each link may be assigned a Resource Class attribute. Resource Class Affinity specifies whether to explicitly include or exclude links with resource classes in the path selection process. The Resource Class Affinity is a 32-bit string accompanied by a 32-bit resource-class mask. The mask indicates which bits in the resource class need to be inspected. The link is included in the Constraint-based LSP when the Resource Class Affinity string or mask matches the Resource Class attributes.
- The **Adaptability** attribute indicates whether the traffic trunk should be re-optimized and consequently rerouted to another path primarily due to the changes in resource availability.



Continuing with the Trunk Attributes the following two are closely associated and play an important role in competitive situations where more traffic trunks compete for the link resources. Two types of priorities are assigned to each traffic trunk:

- Setup priority (**Priority**) defines the relative importance of traffic trunks and determines the order in which path selection is done for traffic trunks at connection establishment and during rerouting due to faulty conditions. Priorities are also important at implementation, permitting pre-emption because they can be used to impose a partial order on the set of traffic trunks according to which pre-emptive policies can be actualized.
- Holding priority (**Pre-emption**) defines the pre-emptive rights of competing trunks and specifies the priority for holding a resource. The attribute determines whether a traffic trunk can preempt another traffic trunk from a given path, and whether another traffic trunk can preempt a specific traffic trunk. Pre-emption can be used to assure that high priority traffic trunks can always be routed through relatively favorable paths within a differentiated services environment. Pre-emption can also be used to implement various prioritized restoration policies following fault events.



Two additional Trunk attributes define the behavior of the tunnel in faulty conditions or in cases when the trunk becomes non-compliant with trunk attributes (e.g. required bandwidth):

- The resilience attribute determines the behavior of the trunk under faulty conditions and can specify:
  - Not to reroute the traffic trunk at all.
  - To reroute the trunk trough a path that can provide the required resources.
  - To reroute the trunk though any available path irrespective of available link resources.
- The policing attribute determines the action in situations where the trunk becomes non-compliant. Non-compliance is understood to be when the amount of traffic on the trunk exceeds the required (reserved) bandwidth. Three actions can be performed:
  - The traffic on the tunnel is rate limited (excessive traffic is dropped).
  - The excessive traffic is tagged but still forwarded.
  - The traffic is unconditionally forwarded.

Practice

Q1) Choose the three major traffic trunk attributes that influence the LSP path computation.

A) Setup and Hold priority

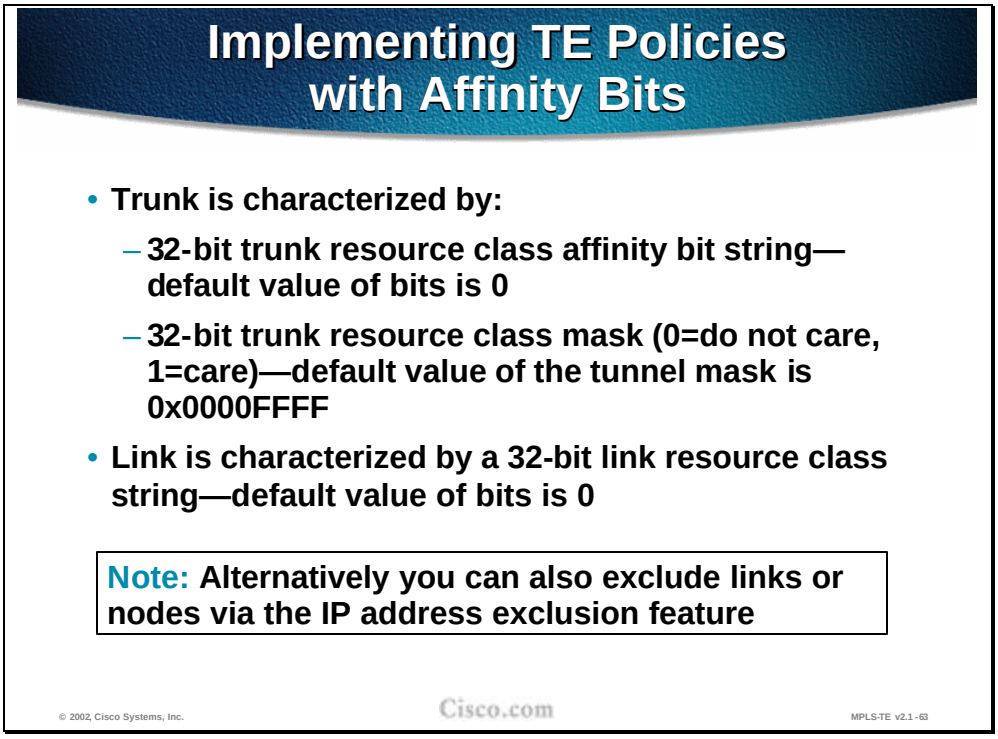
B) Trunk Resource Class Affinity

C) Link Resource Class

D) Traffic Parameter (required bandwidth)

E) Maximum Allocation Multiplier

Implementing TE Policies with Affinity Bits

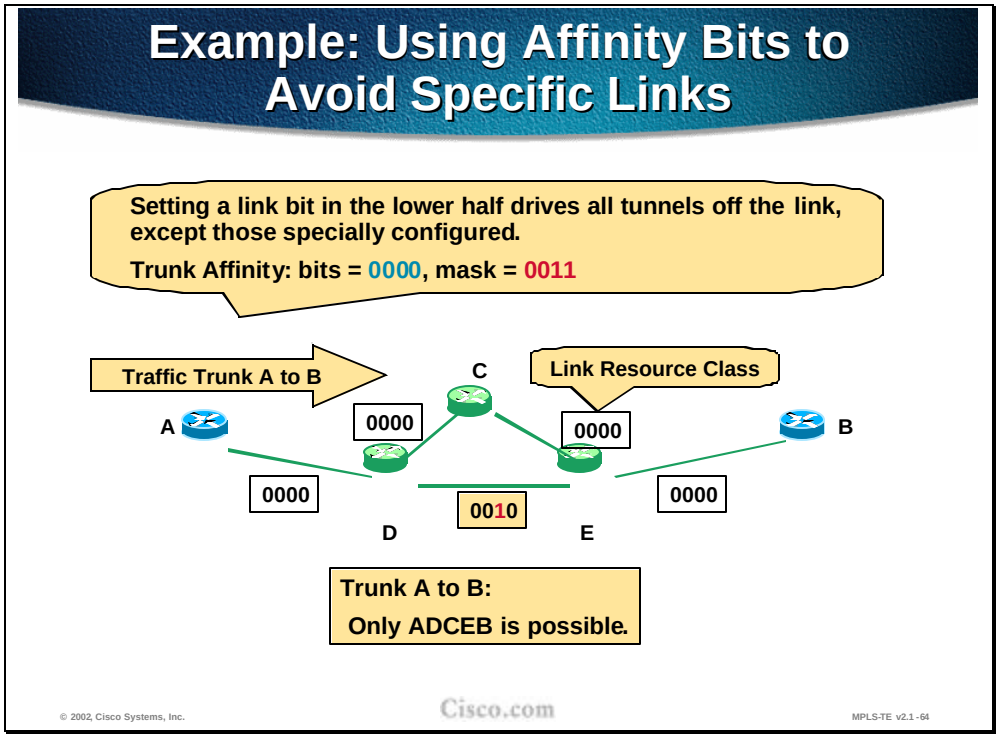


The policies during the LSP path computation can be implemented using the resource class affinity bits of the traffic trunk and the resource class bits of the links over which the trunk should pass (following the computed LSP path).

Each traffic trunk is characterized by a 32-bit resource class affinity string accompanied by a respective resource class mask. The zero bits in the mask exclude the respective link resource class bits from being checked.

Each link is characterized by its resource class 32-bit string, which is set to 0 by default. The matching of the tunnel trunk resource class affinity string with the resource class string of the link is performed during the LSP path computation.

Note There is also the possibility to exclude links or nodes using the **IP address exclusion** feature when configuring tunnels via the explicit-path command.

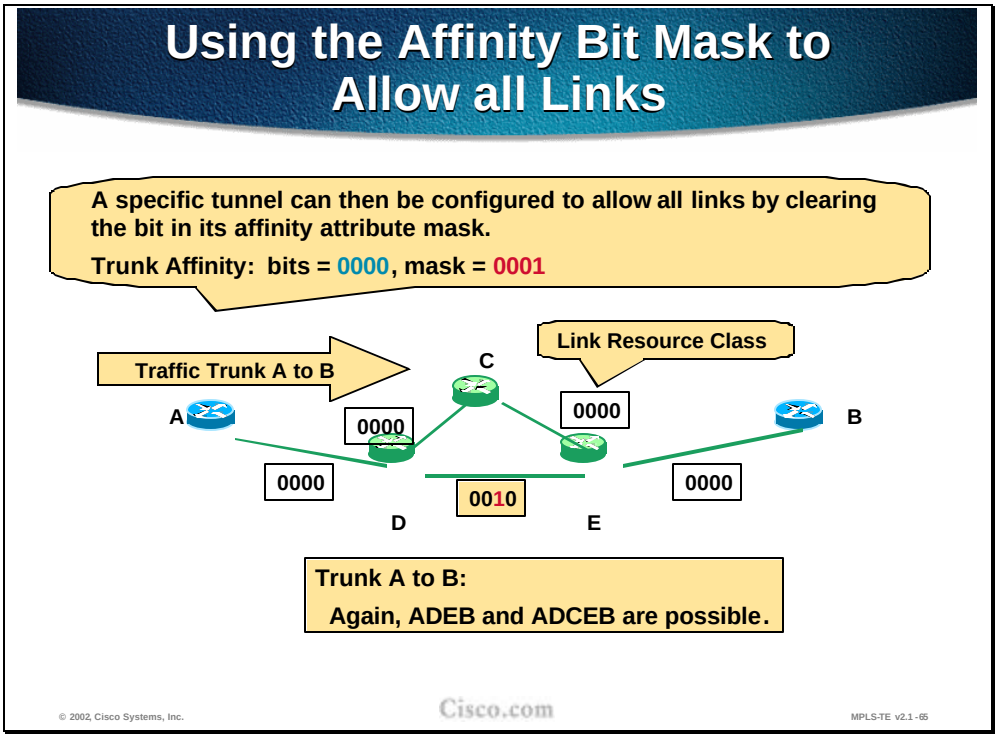


The example in the figure shows a sample network with the trunk resource class affinity bits and link resource bits. For simplicity only the four bits (of the 32-bit string) are shown. The trunk should be established between routers A (head-end) and B (tail-end).

With the trunk resource class affinity bits and the link resource class bits at their default values of 0, the Constraint-based path computation would have two possible paths: A-D-E-B or A-D-C-E-B.

Because it is desirable to move all dynamically computed paths away from the link D-E, the link resource class bits were set to a value 0010 and the trunk mask was set to 0011.

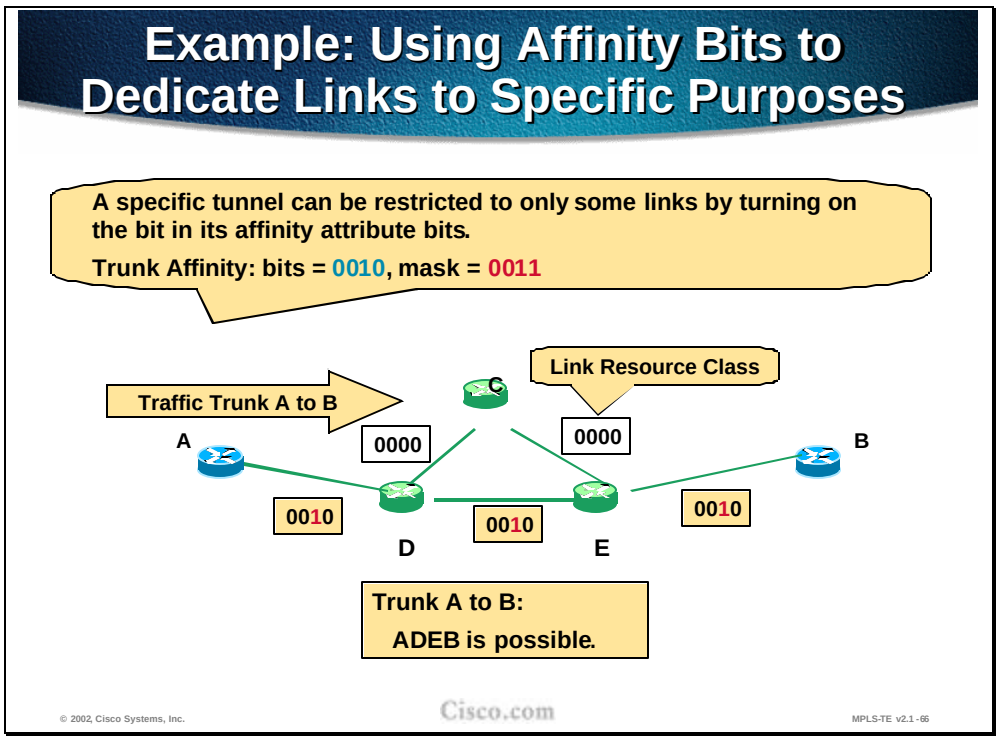
In the example, the trunk mask requires that only the lower two bits require matching. The 00 of the traffic affinity does not match the 10 of the link D-E resource class and results in the exclusion of this link as a possible path for the trunk. The only remaining alternative path is D-C-E, on which the default values of the resource class string (all zeros) match the trunk affinity bits.



Continuing with the sample network, only the lower bit was set in the trunk mask. The trunk affinity bits remain unchanged as well as the resource class bits on the D-E link.

The matching between the trunk resource class affinity bits and the link resource class bits is done on the lowest bit only (due to the mask setting of 0001). The 0 of the trunk affinity bit (the lowest bit) matches with the 0 of the link resource class bit (the lowest bit) and therefore the link D-E remains in the possible path computation (along with the D-C-E link).

Which path will actually be used depends on other trunk and link attributes, including the required and available bandwidth.



The last example with the sample network deals with setting the trunk resource class affinity bits and the link resource class bits to force the tunnel to follow a specific path. Links A-D-E-B are all configured with the resource class value 0010. The trunk resource class affinity bits are set to a value 0010 and the mask to 0011. Only the lower two bits will be compared in the Constraint-based path computation. The 10 of the trunk resource class affinity matches the 10 of the link resource class on all links configured with that value. The 10 does not match the 00 set on the path D-C-E and thus the only possible LSP path remains (A-D-E-B).

Practice

Q1) How can a certain link be excluded from the LSP path computation?

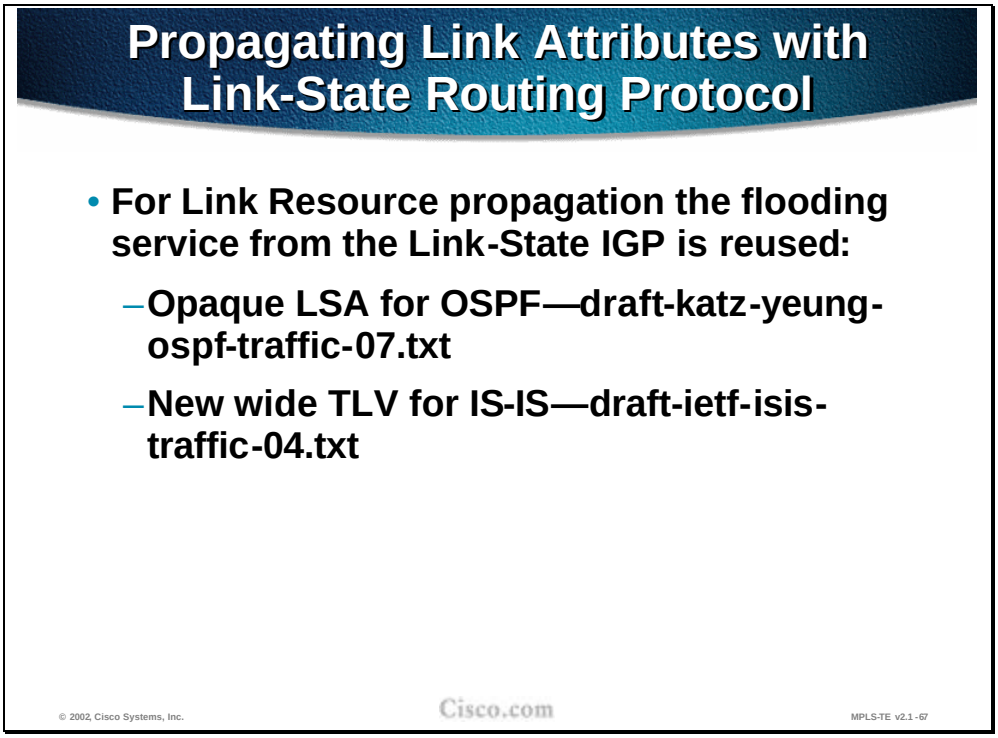
A) Using the proper setting of resource class affinity bits settings and link resource class bit along with resource class mask.

B) By setting the MPLS-TE cost of the link to zero.

C) Using the proper setting of link resource class bit settings and trunk resource class affinity bits along with resource class mask.

D) By manually specifying a path that bypasses a given link.

Propagating MPLS TE Link Attributes with Link-State Routing Protocol



The link resource attributes must be propagated throughout the network to be available at the head-end of the traffic trunk when the LSP path computation takes place.

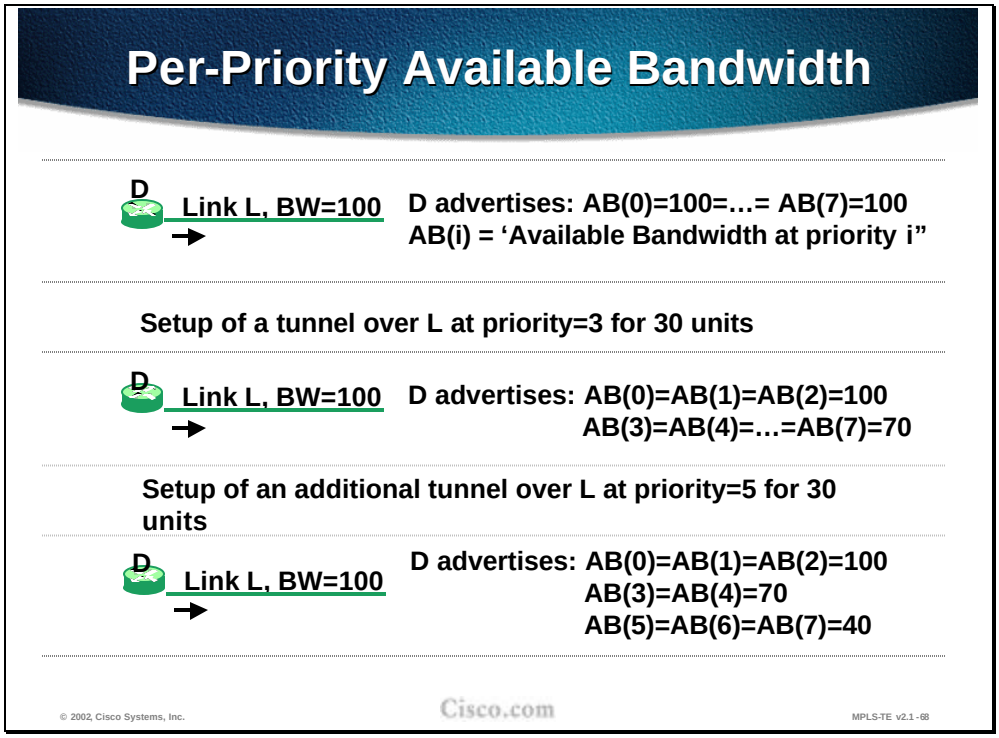
Since the propagation (flooding) of the attributes can only be achieved by link-state routing protocols (Interior Gateway Protocols), OSPF and IS-IS were extended to support the MPLS-TE features.

The OSPF uses new Link State Advertisements (Opaque LSA) and IS-IS uses new Type-Length-Value attributes in its Link State Packets (LSP).

The details on OSPF and ISIS extensions for MPLS-TE can be found in the following documents:

- Opaque LSA (type 10) for OSPF, draft-katz-yeung-ospf-traffic-07.txt, IETF
- New wide TLV (type 22) for IS-IS, draft-ietf-isis-traffic-04.txt, IETF

Note As drafts have a limited lifetime and are replaced by new ones, or converted to RFC documents, have a check at www.ietf.org for the latest versions.

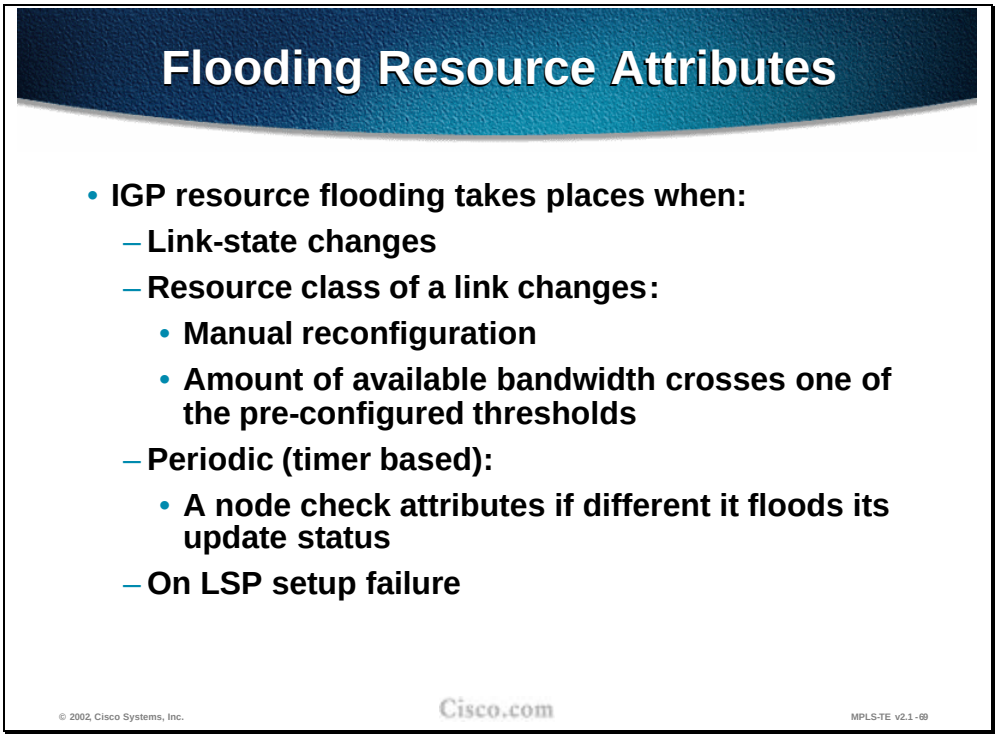


Another important factor in the LSP path computation is the available bandwidth on the link that the traffic trunk will pass. These bandwidths are configured per priority level (8 levels, 0 being the highest, 7 the lowest) and communicated in respective IGP link-state updates, again per priority.

When a certain amount of the bandwidth is reserved at a certain priority level, this amount is subtracted from the available bandwidth at that level and at all levels below. The bandwidth at upper levels remains unchanged.

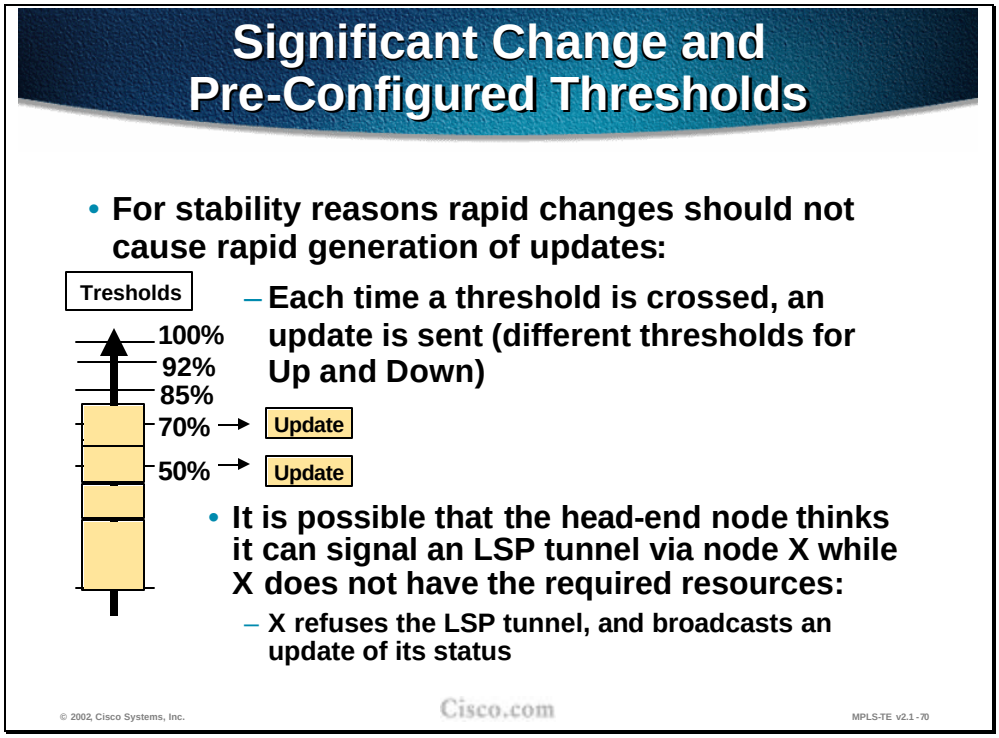
In the example the max. bandwidth is set to the bandwidth of the link, which in is 100 (assuming a Fast Ethernet link). The system allows to set the AB to a higher value than the interface bandwidth, but when making a reservation, any bandwidth above the interface bandwidth will be rejected. The available bandwidth (AB) is advertised in the link-state packets of the router D and the value is 100 at all priority levels before any tunnel is setup. After that a tunnel at priority level 3 requiring 30 units of bandwidth is set up across the link L. The available bandwidth at all priority levels above (0, 1 and 2) remains unchanged at 100. On all other levels, 30 was subtracted from 100 which resulted in available bandwidth of 70 on priority level 3 and below (4-7).

Another tunnel is set up at priority level 5 requiring 30 units of bandwidth across the link L. The available bandwidth at all priority levels above remains unchanged with 100 on 0 to 2 and 70 on 3 and 4. On all other levels 30 was subtracted from 70, which resulted in an available bandwidth of 40 on priority level 5 and below (6-7).



The flooding of resource attributes by the IGP takes place at certain conditions and events:

- When the link changes its state (up, down).
- When the resource class of the link changes due to a manual reconfiguration or in case some pre-configured thresholds are crossed by the available bandwidth.
- Periodically (based on a timer), a node checks resource attributes and if the resource attributes were changed, the update is flooded.
- When the LSP path setup fails.



For stability purposes the significant rapid changes in available link resources should not trigger the updates immediately. The drawback of not propagating the change immediately is that in some cases the head-end sees the link as available for the LSP path and includes the link in its path computation even though the link may be down or does not have the required resource available. When the LSP path is actually being established, a node with the link lacking the required resources cannot establish the path and floods an immediate update to the network.

The thresholds for resources are set both for an up direction (resources exceeding the threshold) and a down direction (resources dropping below the threshold). When the threshold is crossed (in either direction) the node generates an update carrying the new resource information.

The graphic shows the threshold values for up direction (100%, 92%, 85%, 70% and 50%) and two updates being sent out. Each one immediately when the margin is crossed.

Practice

Q1) How are link attributes known to the head-end of the traffic trunk?

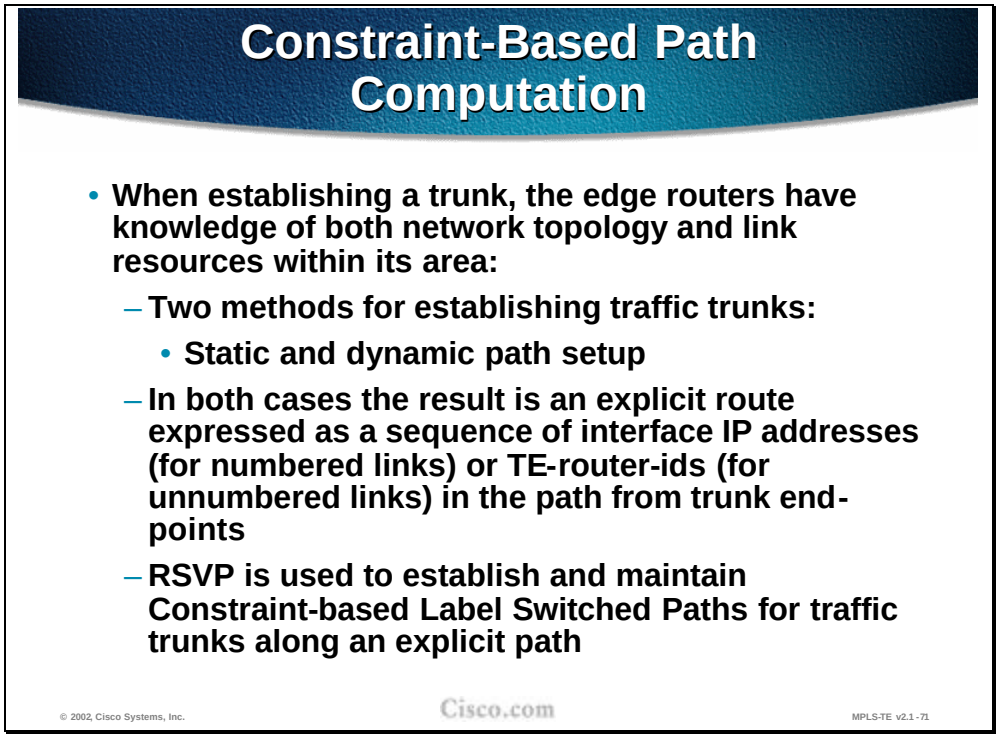
A) Using the modified distance-vector IGP that floods link resources.

B) Using the modified link-state IGP that floods link resources.

C) Through the LDP Request and Response messages.

D) Using the standardized link-state IGP that floods link resources.

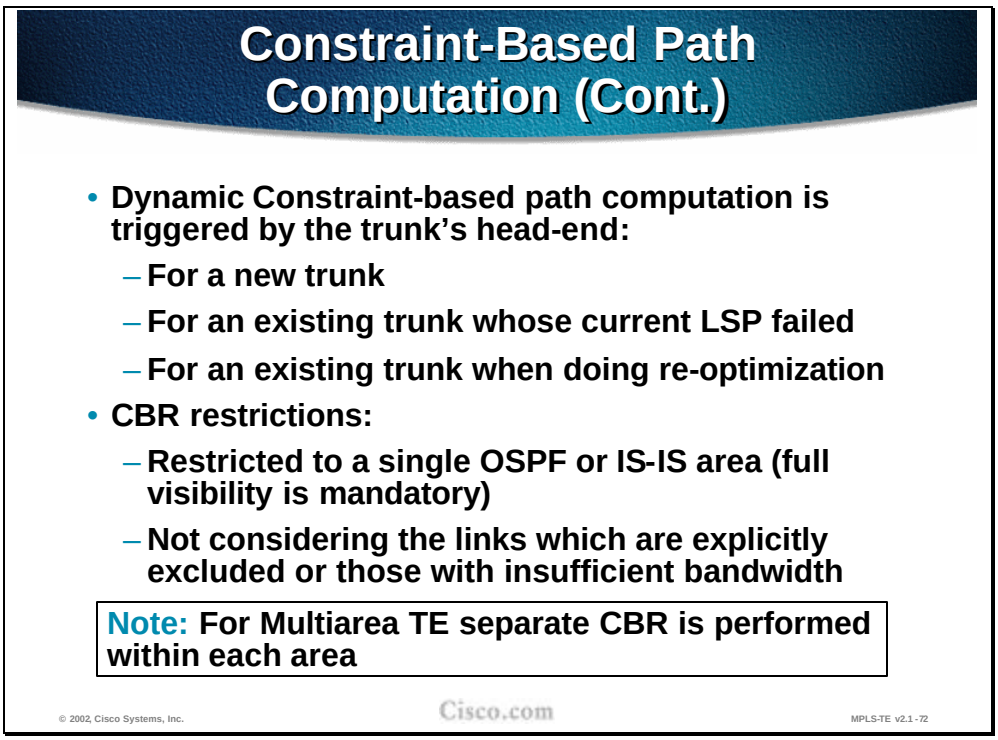
Constraint-Based Path Computation



The head-end of the traffic trunk has the visibility both of the network topology and network resources. This information is flooded throughout the network via a link-state IGP.

The LSP path for the traffic trunk can be statically defined or computed dynamically. The computation takes the available resources and other trunk and link attributes into account (thus constraint-based path computation). The result of the constraint-based path computation is a series of IP-addresses representing the hops on the LSP path between the head-end and tail-end of the traffic trunk.

For LSP signaling and the final establishment of the path, the RSVP is used.



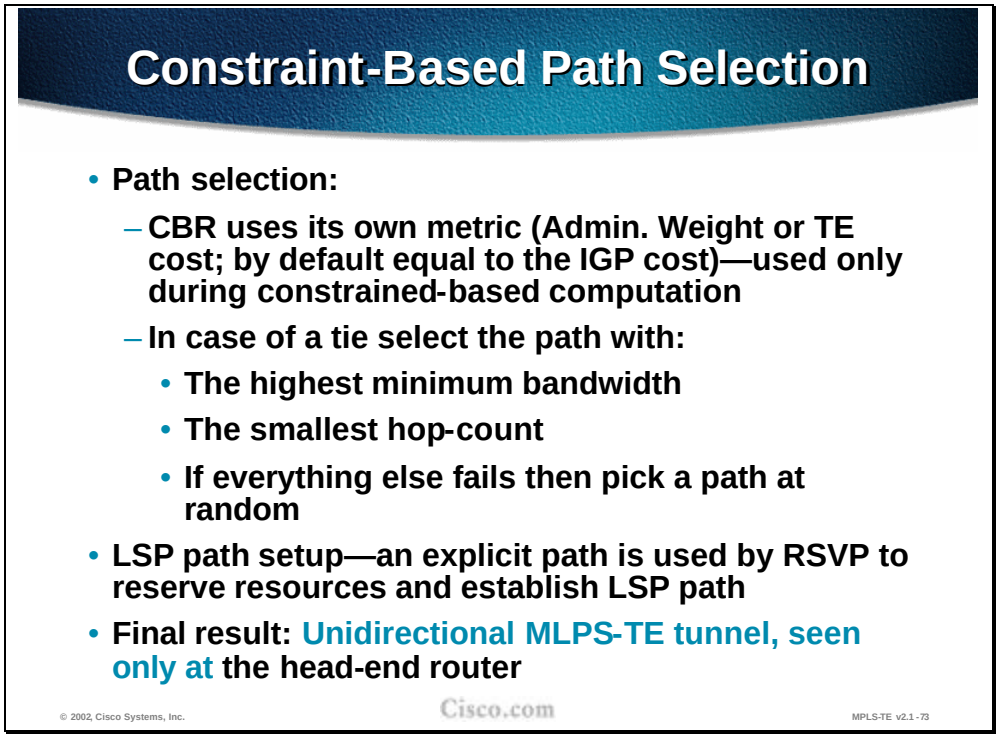
The Constraint-based path computation is always performed at the traffic trunk head-end. The computation is triggered for:

- A new trunk
- An existing trunk whose Label Switched Path setup has failed
- The re-optimization of an existing traffic trunk

The LSP path computation is restricted by several factors (constraint-based). The LSP path can only be computed if:

- The endpoints of the trunk are in the same OSPF or IS-IS area (due to link-state flooding of resources).
- The links that are explicitly excluded via the link resource class bit string, or that cannot provide the required bandwidth, are pruned from the computation.

As the linkstate database only contains the relevant information within a single area, for **multiarea Traffic Engineering** separate CBRs have to be performed within each area.



The Constrained-based path computation selects the path that the traffic trunk will take based on the administrative weight (TE cost) of each individual link. This administrative weight is by default equal to the IGP link metric. The value is used only during the constraint-based path computation.

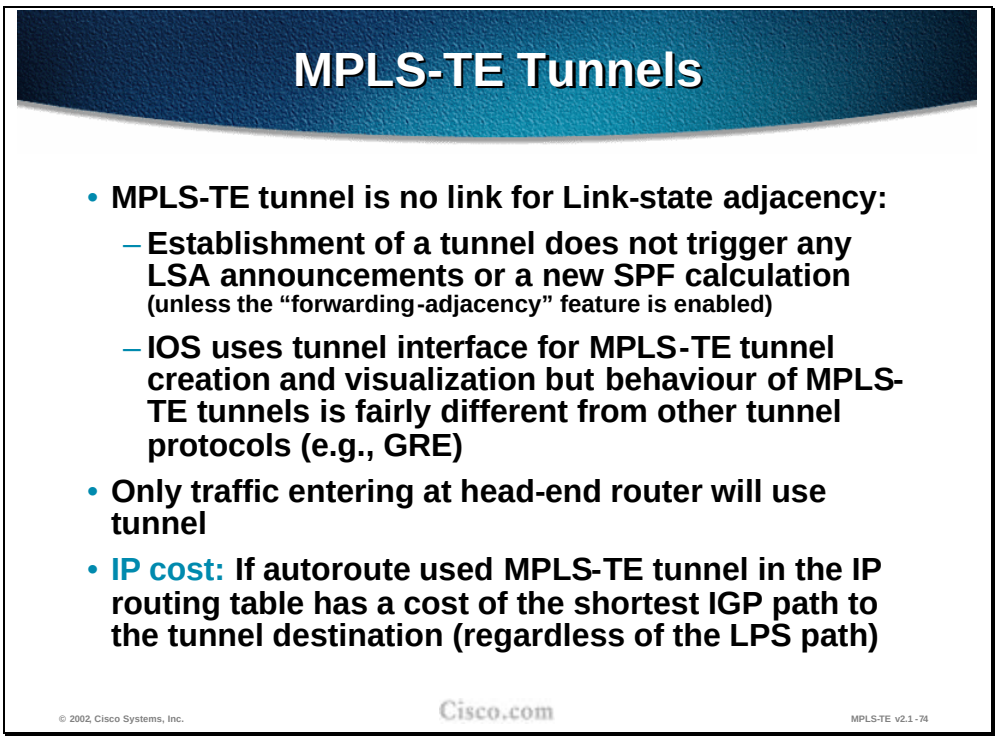
If there are more candidates for the LSP path (several paths with the same metric) then the selection criteria is (in sequential order):

- The highest minimum bandwidth on the path takes precedence.
- The smallest hop count takes precedence.

If after applying all the criteria still more than one path exists the path is randomly chosen.

When the LSP path is computed, the RSVP is used to actually reserve the bandwidth, to allocate labels for the path, and finally to establish the LSP path.

The result of a constraint-based path computation is a unidirectional MPLS-TE tunnel (traffic trunk) that is seen only at the tunnel endpoints (head-end and tail-end).



From the perspective of an IGP routing, the traffic trunk (tunnel) is not seen as an interface at all and is not included in any IGP route calculation (apart from other IP tunnels like Generic Route Encapsulate tunnels). The traffic engineered tunnel, when established, does not trigger any link-state update or any SPF calculation.

This behavior can be changed by using the **mpls traffic-eng forwarding-adjacency** command and by defining two tunnels in a bidirectional way.

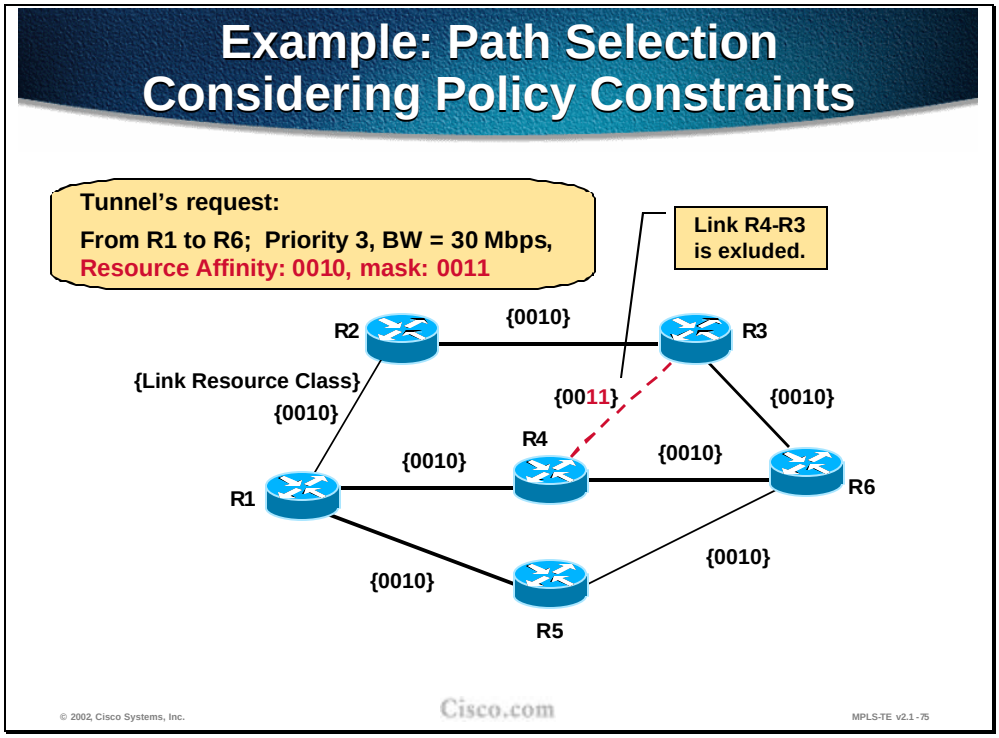
The Cisco IOS software uses the tunnel mainly for visualization. The rest of the actions associated with the tunnel are done by the MPLS forwarding and other MPLS-TE related mechanisms.

The IP traffic that will actually use the traffic engineered tunnel is forwarded to the tunnel only by the head-end of the tunnel. In the rest of the network, the tunnel is not seen at all (no link-state flooding).

With the autoroute feature, the traffic trunk (tunnel):

- Appears in the routing table
- Has an associated IP metric (cost equal to the best IGP metric to the tunnel endpoint)
- Is also used to forward the traffic for destinations behind the tunnel endpoint

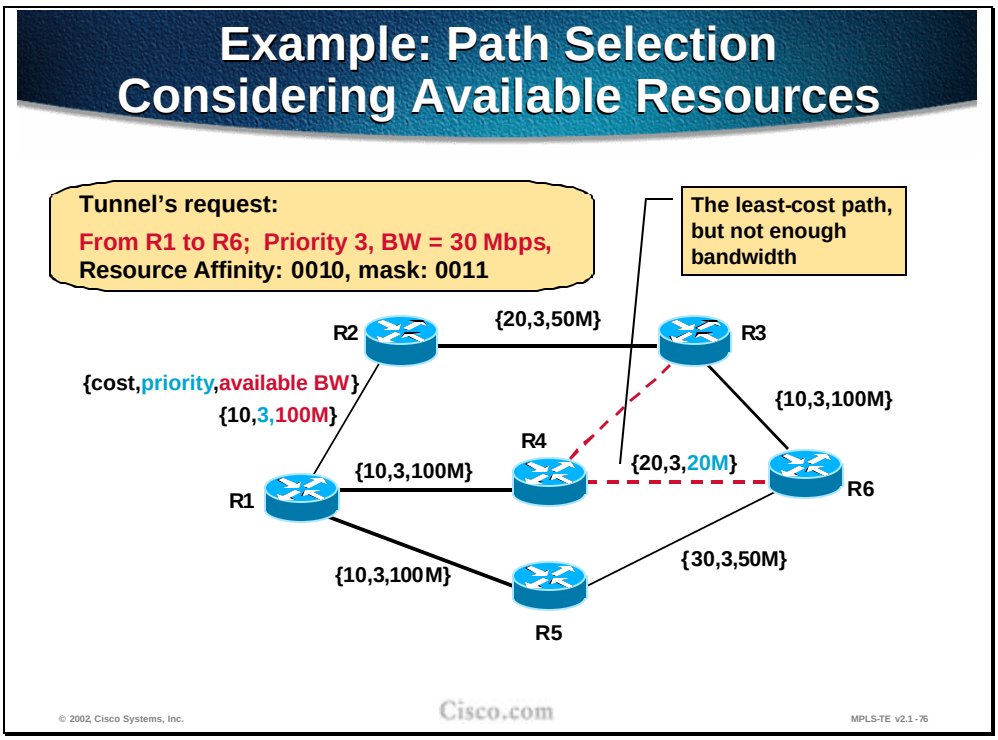
Even with the autoroute feature, the tunnel itself is not used in link-state updates and the rest of the networks still does not have any knowledge of it.



The example of the constraint-based path computation and LSP path selection requires that the traffic trunk (tunnel) be established between R1 (head-end) and R6 (tail-end). The traffic trunk requirements are as follows:

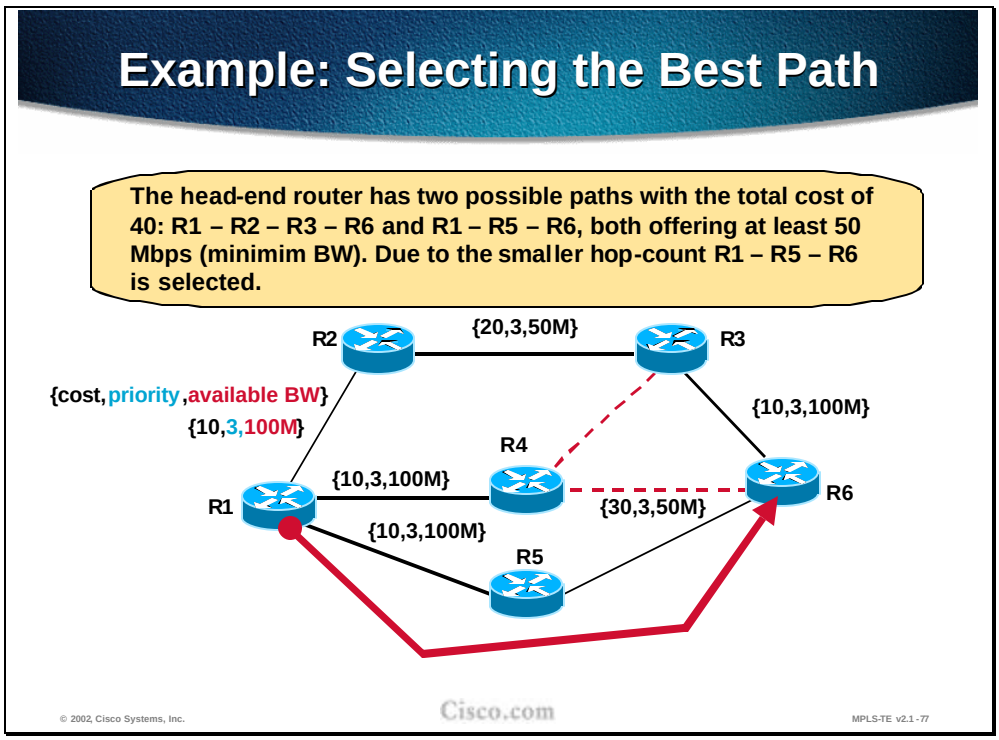
- The required bandwidth at priority level 3 is 30 Mbps
- The resource class affinity bits are set to 0010 and the trunk's mask is 0011. The checking will be done only on the lower two bits.

The link R4-R3 should be excluded from the LSP path and its resource class bit string is set to 0011 accordingly. When the traffic trunk resource class affinity bits are compared to the link R4-R3 resource class bits, there is no match, and the link is effectively excluded from the LSP path computation.



The next parameter checked during the constraint-based path computation is the TE cost (administrative weight) of each link through which the tunnel will possibly pass. The lowest cost is calculated across the path R1-R4-R6 and the overall cost is 30. All other possible paths have a higher overall cost.

When resources are taken into account, the constraint-based path computation finds that on the lowest-cost path there is not enough bandwidth to satisfy the traffic trunk requirements (30 Mbps required, 20 Mbps available). As a result, the link R4-R6 is effectively excluded from the LSP path computation.



The resulting LSP paths (after exclusion of the links that do not satisfy the traffic trunk requirements) in the example are: R1-R2-R3-R6 and R1-R5-R6. Both paths have a total cost of 40 and the tie has to be resolved using the tie-break rules.

First the highest minimum bandwidth on the path is compared. After the comparison still both paths are candidates since both can provide at least 50 Mbps of the bandwidth.

The next rule, the minimum number of the hops on the LSP path, is applied. Since the lower path (R1-R5-R6) has a lower hop-count, this path is finally selected and the constraint-based computation is concluded.

The next step toward final establishment of the LSP path for the traffic engineered tunnel is the signalization of the path via RSVP.

Practice

Q1) Which path is selected when there are several equal-cost LSP path candidates?

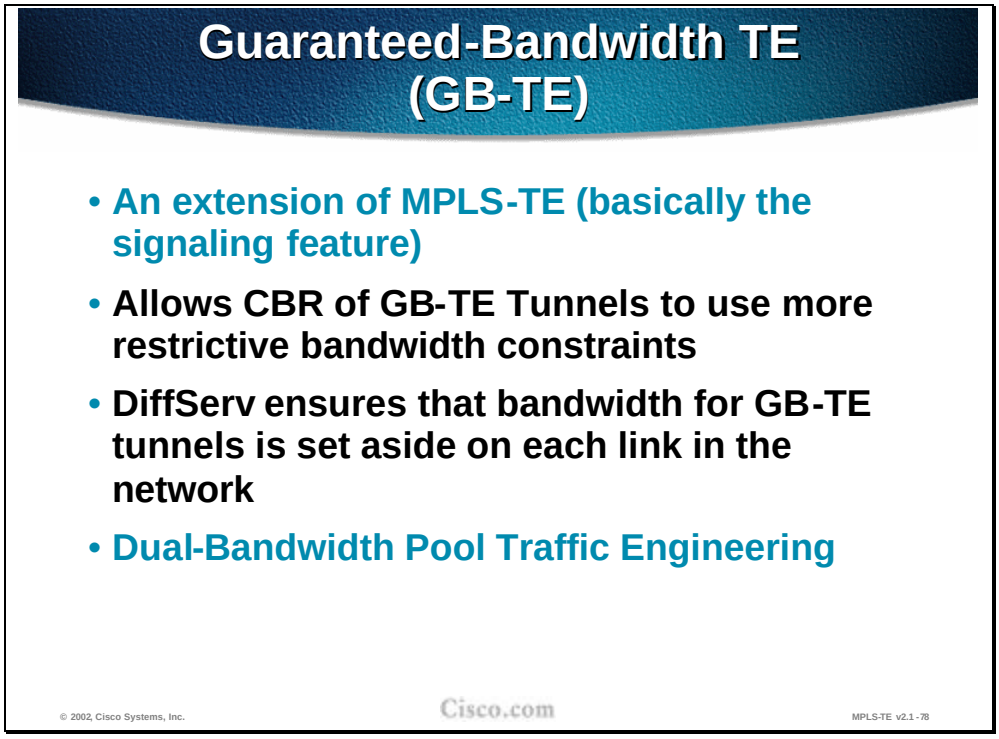
A) The path with the highest minimum bandwidth, then the paths with lower hop count, then random selection.

B) The path with the highest average bandwidth, then the paths with lower hop count, then random selection.

C) The very same one as IGP would select.

D) The path with the lower hop count, then highest minimum bandwidth, then random selection.

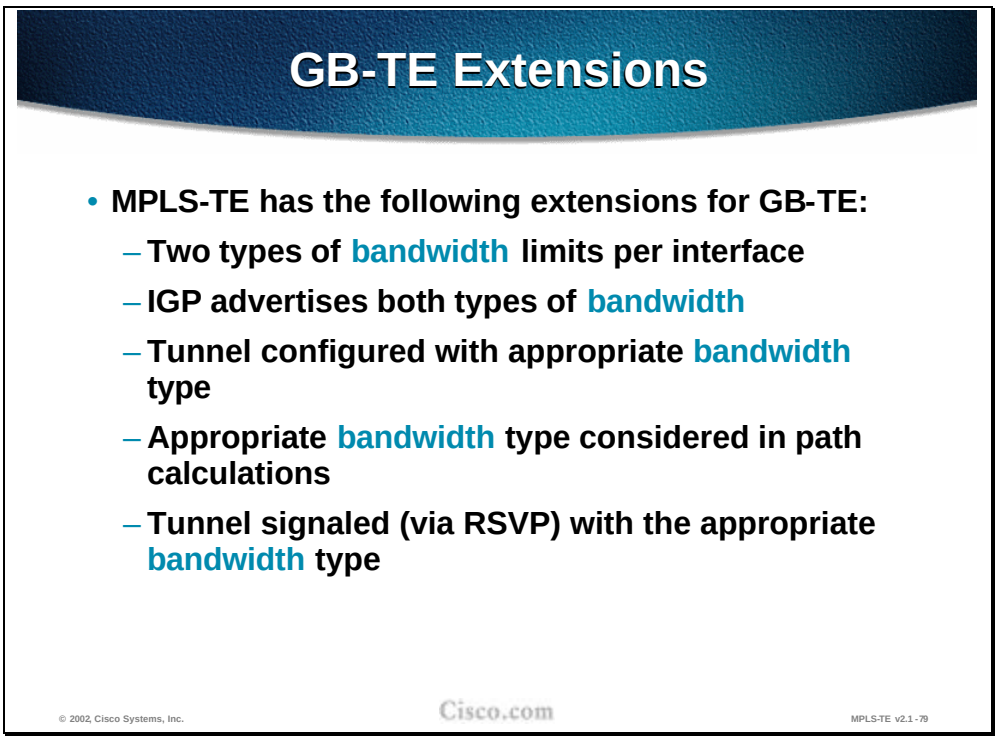
Guaranteed-Bandwidth Sub-Pool



Guaranteed Bandwidth Traffic Engineering (GB-TE) extends the current MPLS Traffic Engineering capabilities to introduce the awareness of a particular Class of Traffic, which is the Guaranteed Bandwidth traffic. GB-TE enables the Service provider to perform a separate admission control and route computation of the Guaranteed Bandwidth traffic. The GB-TE is another signaling feature of IGP and RSVP.

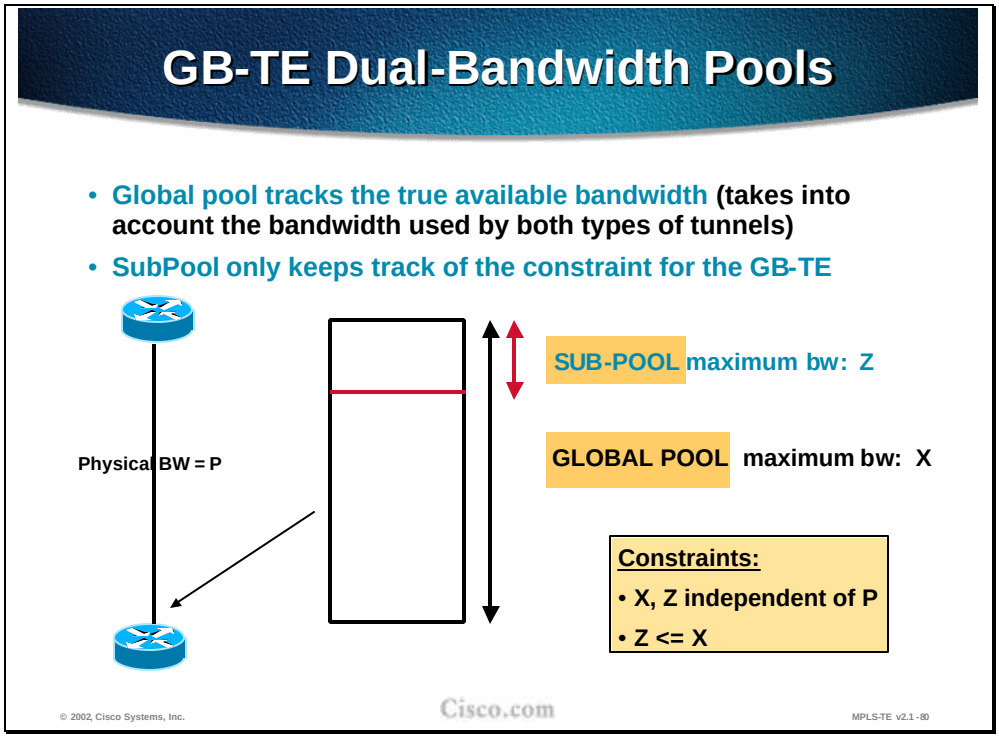
With only a single bandwidth pool on the link in traditional MPLS-TE, when the bandwidth is reserved for the tunnel, the traffic within a tunnel is considered as a single class. For example, when voice and data are inter-mixed within the same tunnel, the QoS mechanisms cannot ensure better service for the voice. Normally, class-based weighted-fair queueing (CB-WFQ) can be performed for the tunnel.

The idea of GB-TE is to **guarantee** the bandwidth for GB-TE tunnels across the network. For critical applications (e.g. voice), a separate GB-TE tunnel is created. Thus two bandwith pools are used, one for traditional MPLS-TE tunnels and one for GB-TE tunnels. The DiffServ Quality of Service mechanisms (low-latency queueing (LLQ)) ensure that bandwidth for GB-TE tunnels is dedicated for these tunnels.In the initial phase, the GB-TE supports a single Class of Guaranteed Bandwidth. It is expected that subsequent phases of GB-TE will extend capabilities such as the support of multiple Classes of Guaranteed Bandwidth and the dynamic re-programming of queuing or scheduling mechanisms.



GB-TE tunnels are similar to regular TE tunnels. To support GB-TE, some modifications to regular MPLS-TE mechanisms were done:

- There are two types of bandwith per each link in the network (two bandwidth pools – the global pool and the sub-pool).
- These bandwidths are both announced in the link-state updates carrying resource information.
- The tunnel (traffic trunk) parameters include the bandwidth type the tunnel will use.
- The Constraint-based path calculation is done with respect to the type of the bandwidth the tunnel requires.In RSVP messages, it is always indicated whether the LSP to be set-up is a regular MPLS-TE tunnel or GB-TE tunnel..
- Intermediate nodes perform admission control and bandwidth allocation (“locking” for the GB-TE) on the appropriate bandwidth pool.



On each link in the network two bandwidth pools are established:

- Global (main) pool that keeps track of the true available bandwidth. The pool takes into account the bandwidth used by both of the tunnels.
- Sub-pool (GB-TE) which tracks only the bandwidth for the GB-TE tunnels.

The bandwidths specified for both pools are independent of the actual physical bandwidth of the link (providing for over-subscription). The same applies also to the traditional MPLS-TE with one bandwidth pool.

The only constraint for the two pools is that the bandwidth of the sub-pool (dedicated to GB-TE tunnels) must not exceed the bandwidth in the global pool.

Practice

Q1) What is the purpose of Guaranteed-Bandwidth? (Choose two.)

A) The GB-TE is another signaling feature of IGP and LDP.

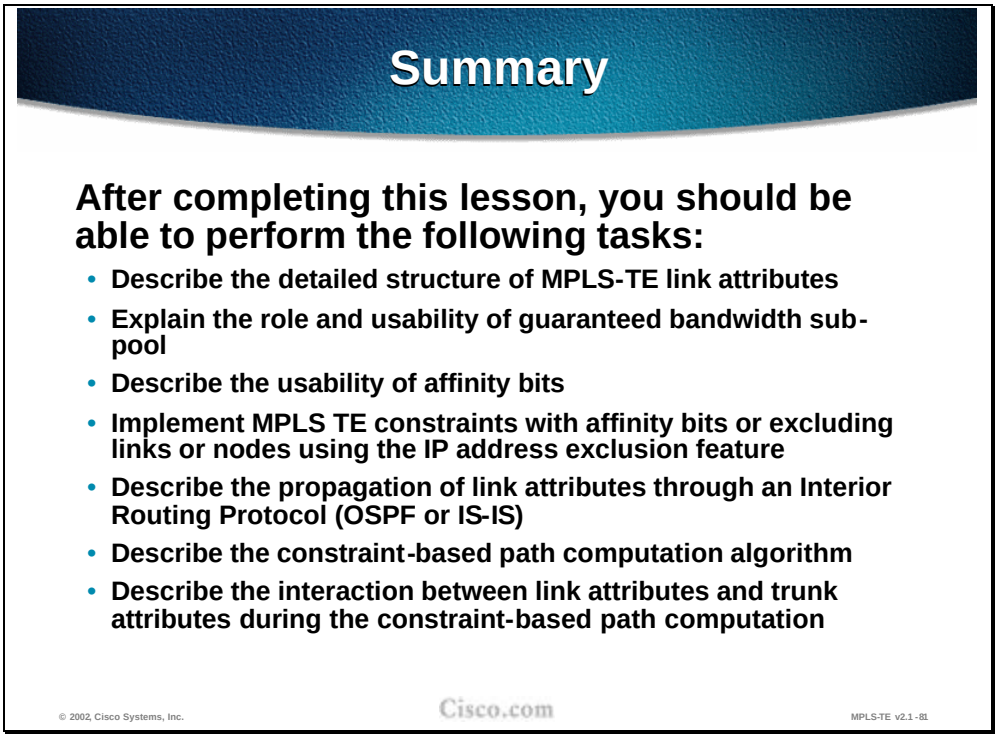
B) The GB-TE is another signaling feature of IGP and RSVP.

C) GB-TE supports multiple Classes of Guaranteed Bandwidth and the dynamic re-programming of queuing or scheduling mechanisms.

D) Guaranteed Bandwidth enables the Service provider to perform a separate admission control and route computation for the differentiated trunks.

Summary

This section summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Path Setup and Maintenance

Lesson Review

Instructions

Answer the following questions:

1. List the major MPLS-TE link attributes that influence the LSP path computation.
2. List the major traffic trunk attributes that influence the LSP path computation.
3. How are link attributes known to the head-end of the traffic trunk?
4. How can a certain link be excluded from LSP path computation?
5. Which path is selected when there are several equal-cost LSP path candidates?

Path Setup and Maintenance

Overview

This lesson describes the details of MPLS traffic engineering tunnels including path setup procedures and path maintenance.

Importance

This lesson is a mandatory for the students planning to improve the usage of their network resources with MPLS traffic engineering.

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- Describe the MPLS-TE path setup procedures
- Explain the details of RSVP assistance in MPLS TE path setup
- Describe the functions of trunk and link admission control
- Explain path monitoring and rerouting
- List the methods for path and link protection
- Explain the traffic trunk reoptimization and bandwidth requirement adjustments

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Cisco Certified Internetwork Professional (CCIP) level of knowledge or equivalent level of IP routing and Cisco IOS knowledge as well as solid understanding of MPLS and link state protocols (OSPF or Integrated IS-IS).

Mandatory Prerequisites:

- AMVS course

Optional prerequisites:

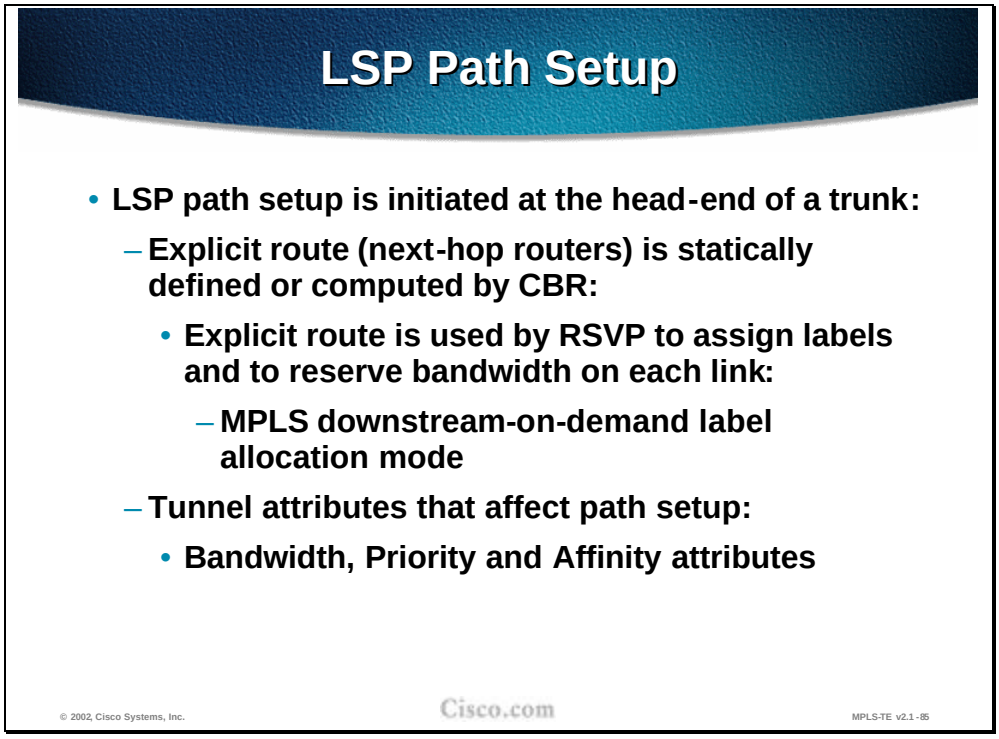
- CISIS course for students deploying MPLS TE in IS-IS environments

Outline

This lesson includes these sections:

- Overview
- LSP Path Setup
- RSVP Usage in Path Setup
- Example—Hop-by-Hop Path Setup with RSVP
- Trunk and Link Admission Control
- Path Monitoring
- Path Re-Routing
- Path Re-Optimization
- Path and Link Protection
- Path Adjustment with Autobandwidth
- Summary
- Lesson Review

LSP Path Setup

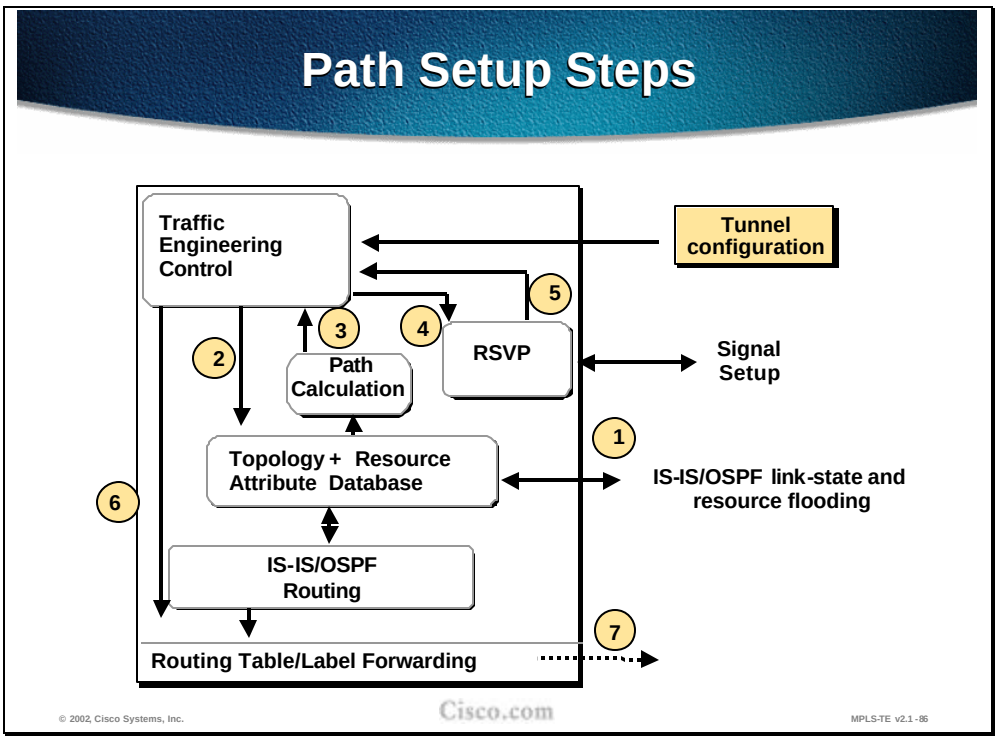


The Label Switched Path (LSP) setup is always initiated at the traffic trunk head-end. The explicit route for the traffic trunk is composed of the list of next-hop routers towards the trunk endpoint. The LSP tunnels can be statically defined or computed with constraint-based routing and thus routed away from network failures, congestion, and bottlenecks.

The explicit route is used by the Resource reSerVation Protocol (RSVP) with traffic engineering extensions to assign labels and to reserve the bandwidth on each link. Labels are assigned using the downstream-on-demand allocation mode.

The path setup is affected by the following tunnel attributes:

- Bandwidth
- Priority
- Affinity attributes



The figure represents a conceptual block diagram of the pieces that form the constraint-based routing and path computation. In the upper left corner there is a Traffic Engineering control module where the control algorithms run. The module looks at the tunnels that have been configured for constraint-based routing.

The Traffic Engineering control module will periodically check the constraint-based routing topology database (shown in the middle of the block diagram) to calculate the best current path from the current device to the tunnel destination. Once the path is calculated, the module will pass the path off to the RSVP module to signal the circuit setup across the network. If the signalization succeeds, the signaling message will eventually return to the device, and RSVP will announce back to Traffic Engineering control module that the tunnel has been established. Consequently the Traffic Engineering control module will tell the IGP routing module that the tunnel is available for use. The IGP routing module will include the tunnel information into its routing table calculation and use it to affect what routes are put into the routing table.

Practice

Q1) How is an LSP path setup initiated?

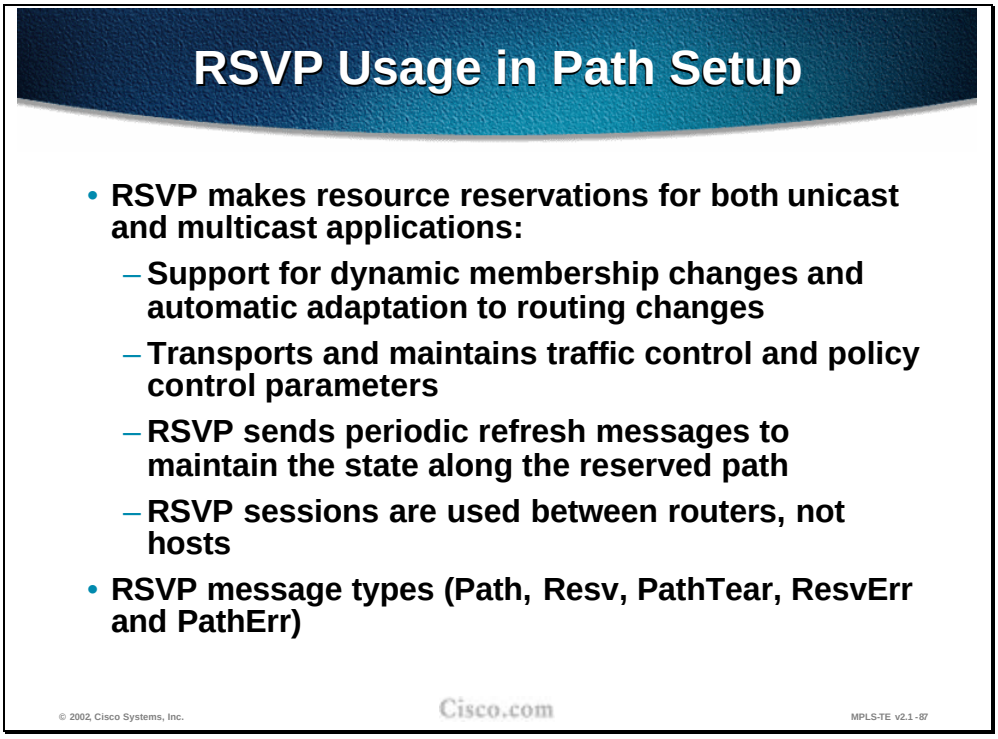
A) By the head-end using the RSVP signalization.

B) By the tail-end using the RSVP signalization.

C) By the head-end using independent allocation mode.

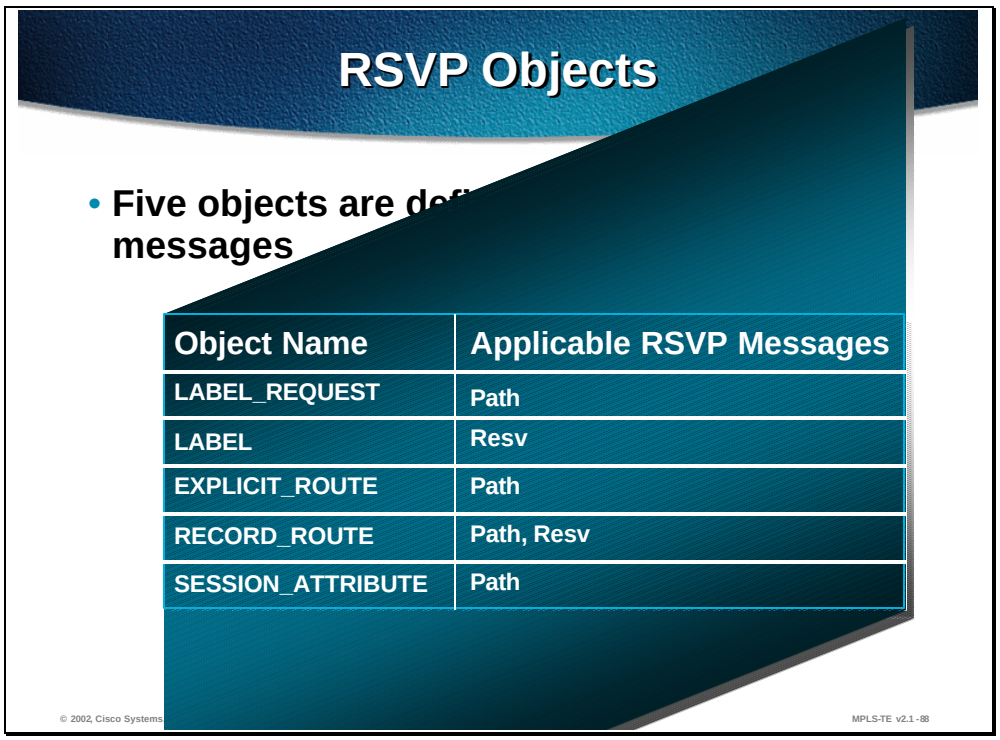
D) By the head-end using the LDP signalization.

RSVP Usage in Path Setup



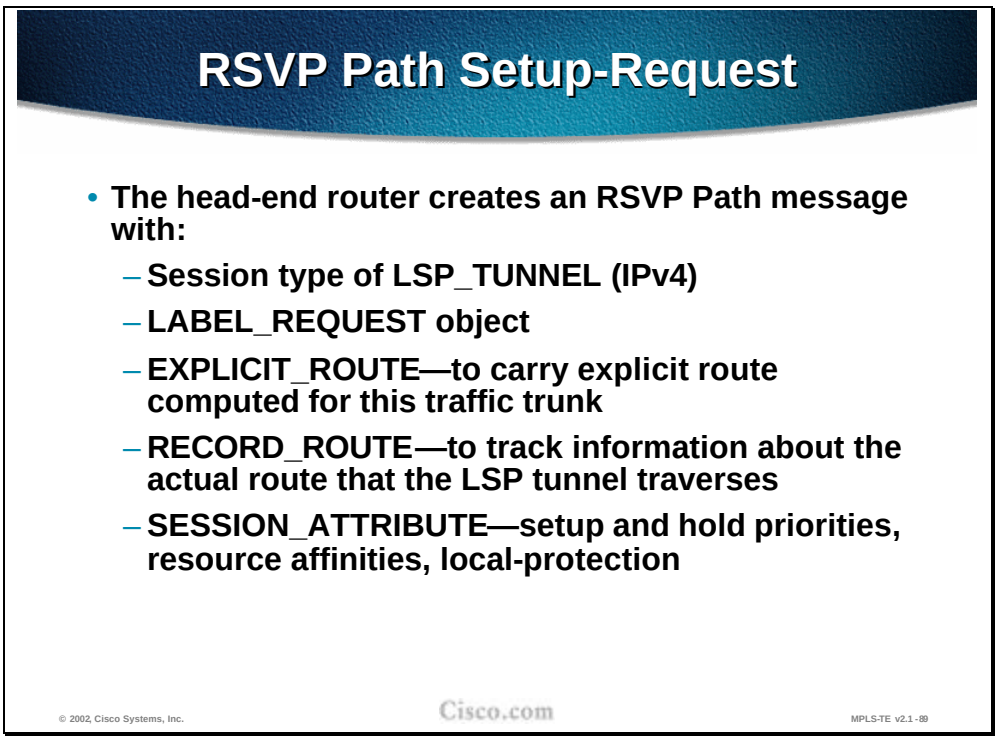
The RSVP plays a significant role in the path setup for LSP tunnels and supports both unicast and multicast applications. The RSVP dynamically adapts to changes either in membership (e.g. multicast groups) or in routing itself (changes in routing tables). Additionally the RSVP transports the traffic parameters and maintains the control and policy over the path. The maintenance is done by periodic refresh messages sent along the path to maintain the state. In the normal usage of RSVP, the sessions are run between hosts. In traffic-engineering, the RSVP sessions are run between the routers on the tunnel endpoints. The following RSVP message types are used in path setup:

- Path
- Resv
- PathTear
- ResvErr
- PathErr



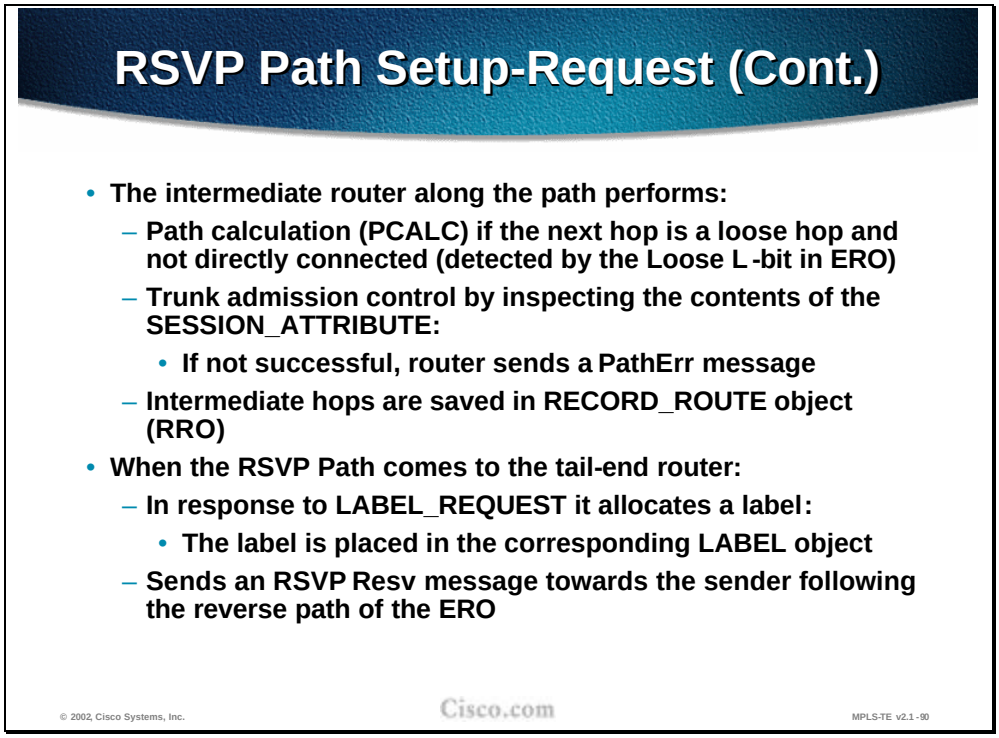
In the Path and Resv messages of the RSVP there are five objects that are traffic engineering related:

- A Label_Request object is carried in the Path message and requests the label assignment. A request to bind labels to a specific LSP tunnel is initiated by an ingress node through the RSVP Path message.
- A Label object is returned with the Resv message. Labels are allocated downstream and distributed (propagated upstream – from tail-end to the head-end) by means of the RSVP Resv message.
- An Explicit_Route object (ERO) is carried in the Path message to request or suggest a specific route for the traffic tunnel (in the form of a concatenation of hops which constitutes the explicitly routed path). The object is used if the sender node has knowledge of a route that has a high likelihood of meeting the tunnel's QoS requirements, or that makes efficient use of network resources.
- A Record_Route object (RRO) is added to the Path and Resv message to enable the sender node to receive information about the actual route that the LSP tunnel traverses.
- A Session Attribute object can be added to Path messages to aid in session identification and diagnostics. Additional control information, such as setup and hold priorities, resource affinities, and local-protection, are also included in this object.



The path setup is initiated at the head-end router with a RSVP path message carrying the following information:

- Session Type of the LSP tunnel that denotes that the destination address is an IPv4 (or v6?) address.
- Label Request object requests intermediate routers to provide a label binding for the session. If a node is incapable of providing a label binding, it sends a PathErr message with an “unknown object class” error. If the Label Request object is not supported end to end, the sender node will be notified by the first node, which does not provide this support.
- Explicit Route object is sent by the head-end if it knows of a route that has a high likelihood of meeting the tunnel’s requirements (either statically configured or computed. The object in the Path message requests the intermediate nodes to forward the Path message towards its destination along a path specified by the object itself.
- Record Route object in the Path message is used by the sender to receive information about the actual route that the LSP tunnel traverses. Since the Record Route object is analogous to a path vector, it can be used for loop detection as well.
- Session Attribute object is populated by the sender with path control information, such as setup and hold priorities, resource affinities, and local-protection.

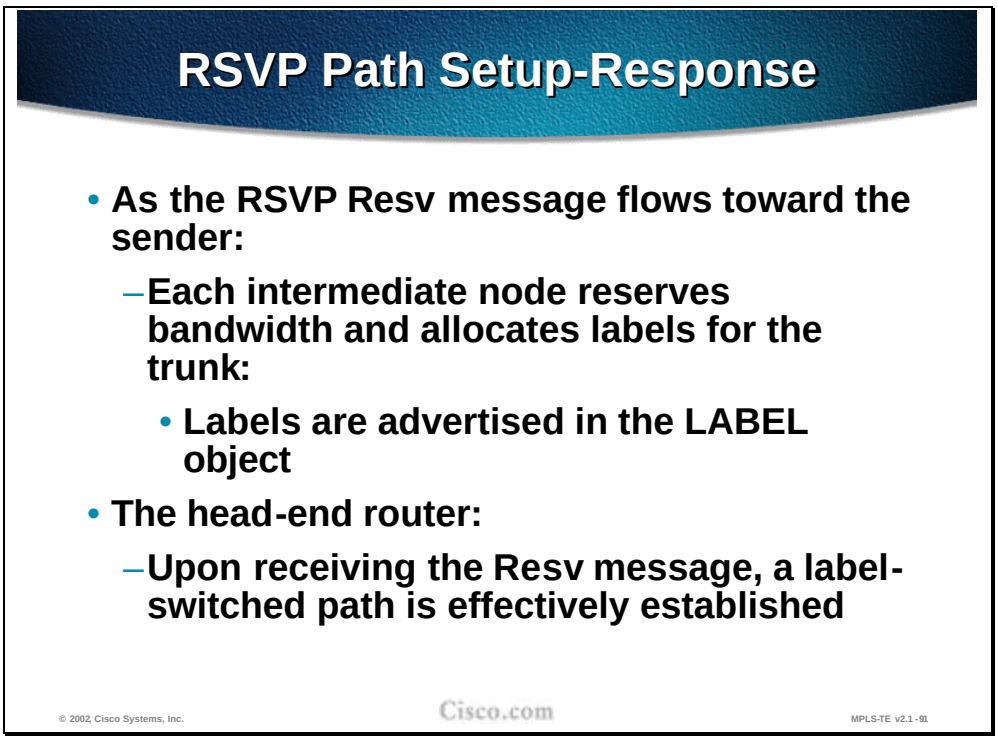


First an intermediate router checks the ERO and looks into the L-bit (loose) regarding the next hop information. If this bit is set and the next hop is not on a directly connected network the node performs a CBR calculation (PCALC) using its traffic engineering database specifying this loose next-hop as destination.

In that way the ERO is augmented by the new results, now forming a hop by hop path up to the next loose node specification.

Then the intermediate routers along the path (indicated in the Explicit Route) perform the traffic trunk admission control by inspecting the contents of the Session Attribute object. If the node cannot meet the requirements it generates the PathErr message. If the requirements are met, the node is saved in the Record Route objects.

When the RSVP Path message arrives at the tail-end router (the end-point of the trunk), the Label Request message triggers the path label allocation. The label is placed to the corresponding Label object of the RSVP Resv message that is generated. The RSVP message is sent back to the head-end following the reversed path recorded in the Explicit Route object (ERO) and stored at each hop in its path state block.



When the RSVP Resv message flows back towards the sender, the intermediate nodes reserve the bandwidth and allocate the label for the trunk. The labels are placed into the Label object of the Resv message.

When the RSVP Resv message arrives to the head-end router the required LSP path setup has been effectively established.

Practice

Q1) List the four main components (objects) of RSVP messages that help establish the MPLS-TE tunnel.

A) Session_Attribute

B) Tunnel source and destination address

C) Explicit_Label

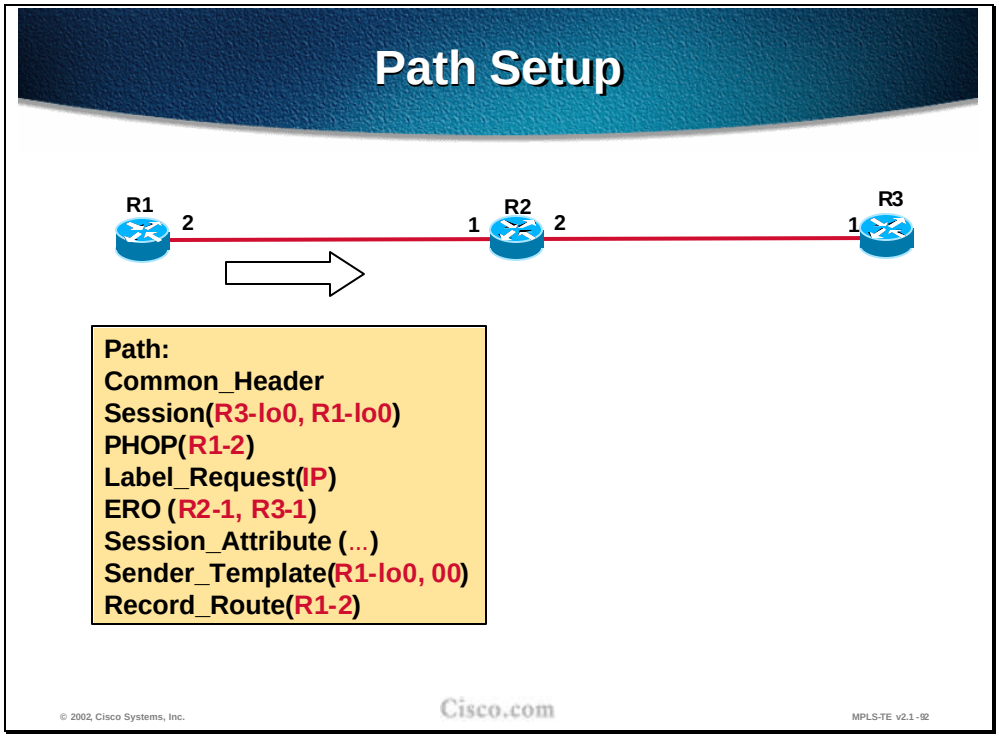
D) Implicit_Route

E) Label_Request

F) Explicit_Route

G) Record_Route

Example—Hop-by-Hop Path Setup with RSVP

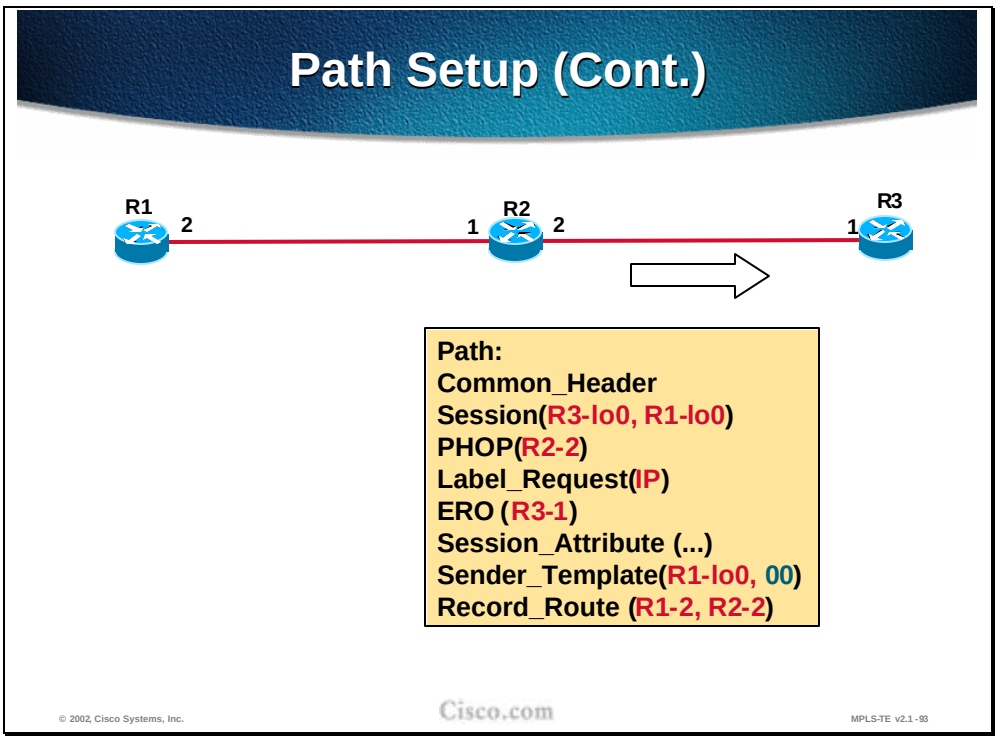


The LSP tunnel path setup is initiated by the RSVP Path message initiated by the tunnel head-end (Router R1 in this example). Some of the most important contents are explained and monitored in the following example.

The RSVP Path message contains several objects including the Session identification (R3-lo0, 0, R1-lo0 in the example), which uniquely identifies the tunnel. The traffic requirements for the tunnel are carried in the Session_Attribute. The Label request that is present in the message is handled by the tail-end router, which allocates the respective label for the LSP path.

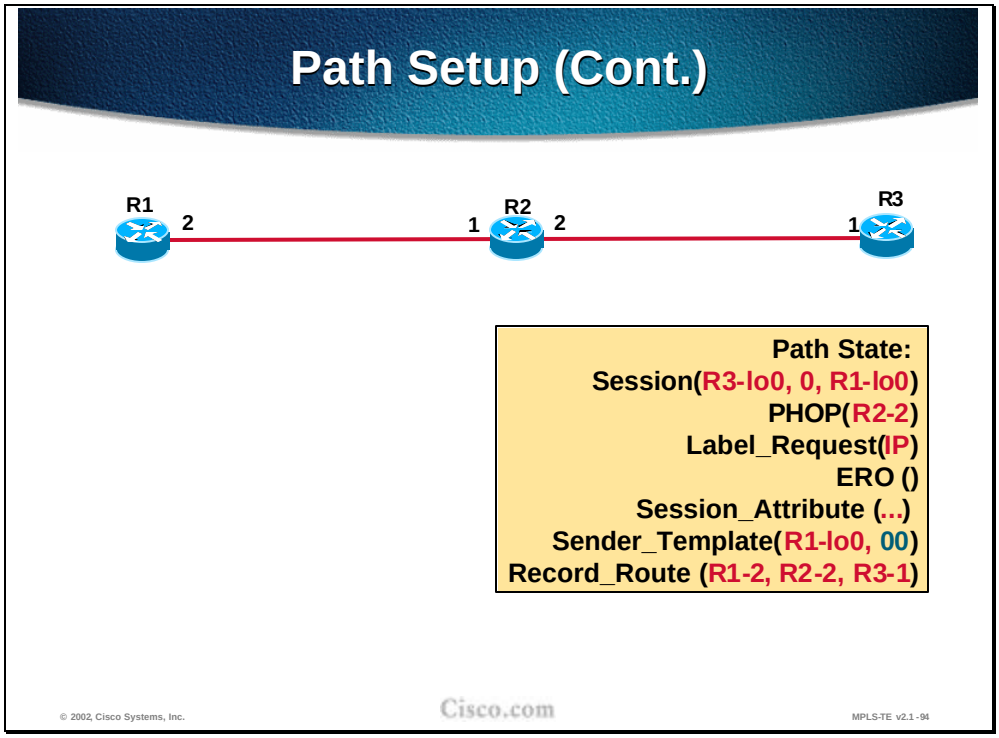
The Explicit Route object (ERO) is populated by the list of next-hops that are either manually specified or calculated by the CBR (where R2-1 is used to represent the interface labeled “1” at the R2 router in the slide) The PHOP (Previous hop) is set to the router`s outgoing interface address. The Record_Route object (RRO) is populated with the same address as well.

Note The Sender_Template is used in assigning unique LSP path identifiers (R1-lo0 - loopback interface 0 which identifies the tunnel head-end, 00 – stays for LSP_ID), It can happen that the same tunnel takes two possible LSP paths (one primary and another secondary). In such a case the head-end must take care of assigning unique IDs to these paths.



As the next hop router (R2) receives the RSVP Path message, it places the contents of ERO into its path state block and removes itself from the ERO (R2 removed the R2-1 entry from the ERO). Router R2 adjustthe PHOP to the address of its own interface (the “2” interface at R2, R2-2) and adds the address (R2-2) to the RRO. The Path message is then forwarded to the next-hop in the ERO.

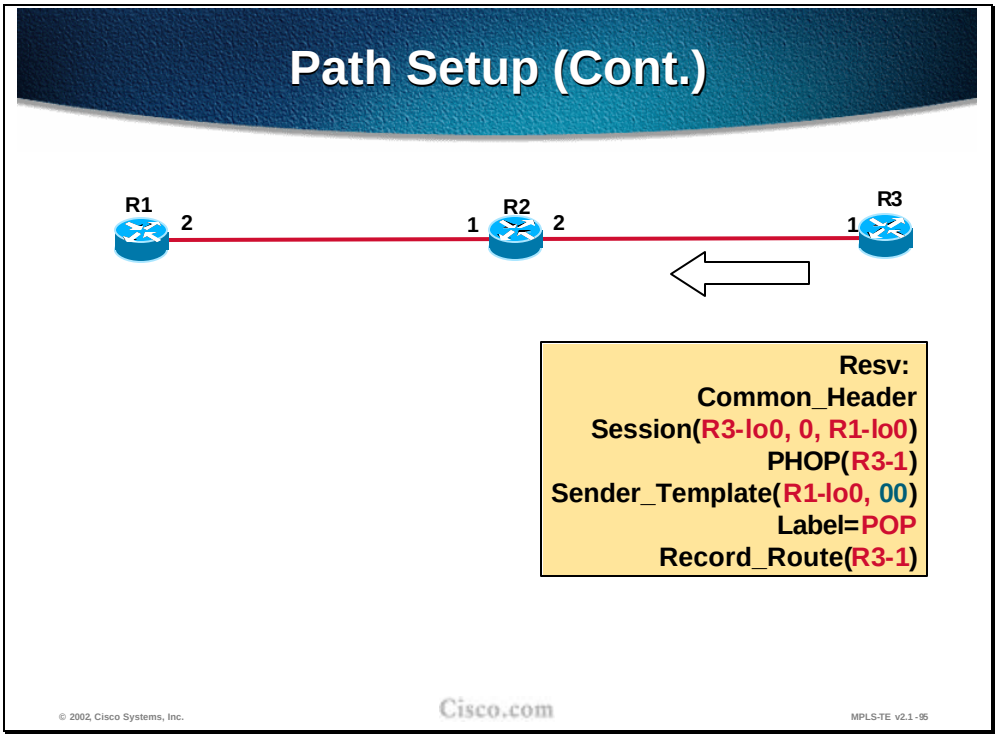
Note	Several other functions are performed at each hop as well, including the traffic admission control.



When the RSVP Path message arrives to the tail-end router (R3), the path state block is created and the ERO becomes empty (after removing the router`s own address from the list) indicating it has reached the tail-end of the tunnel. The RRO at this moment contains the entire path from the head-end router.

The RSPV Resv message must be generated.

The Label Request object in the RSVP Path message requires the tail-end router to allocate a label for the specified LSP tunnel (session).

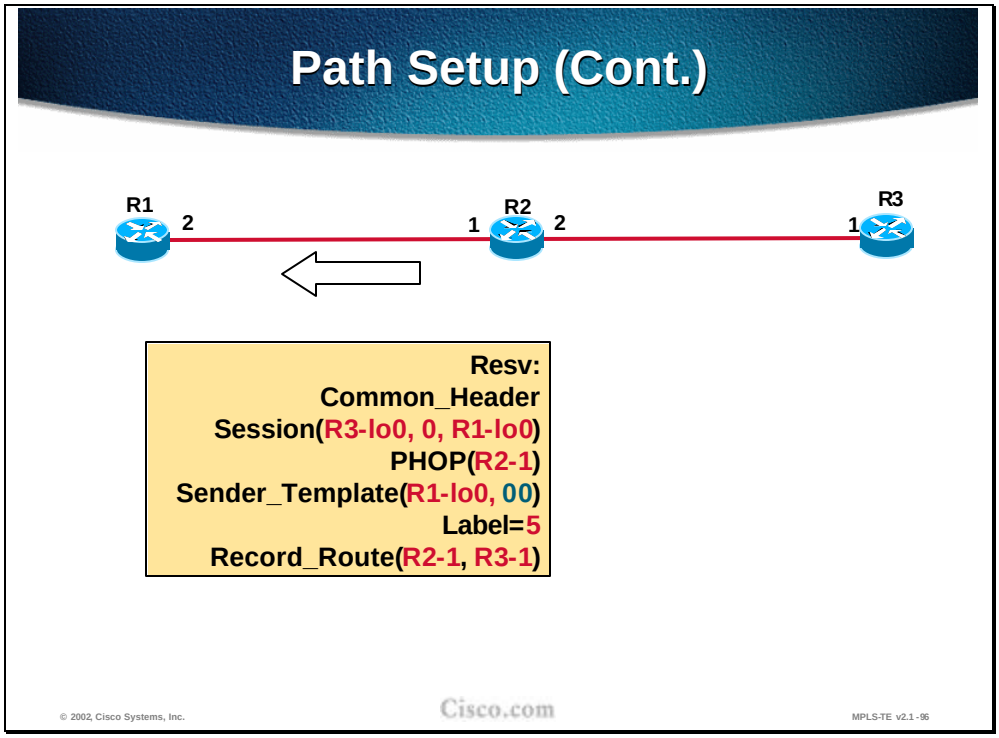


Since R3 is the tail-end router, it does not allocate a specific label for the LSP tunnel. The implicit-null label is used instead (the value “POP” in the Label object).

The PHOP in the RSVP Resv message is populated by the tail-end router’s interface address and this address is copied to the RRO as well.

Note The RRO is re-initiated in the RSVP Resv message.

The Resv message is forwarded to the next-hop address in the path state block of the tail-end router. The next hop information in the path state block was established when the Path message was traveling in the opposite direction (head-end to tail-end).



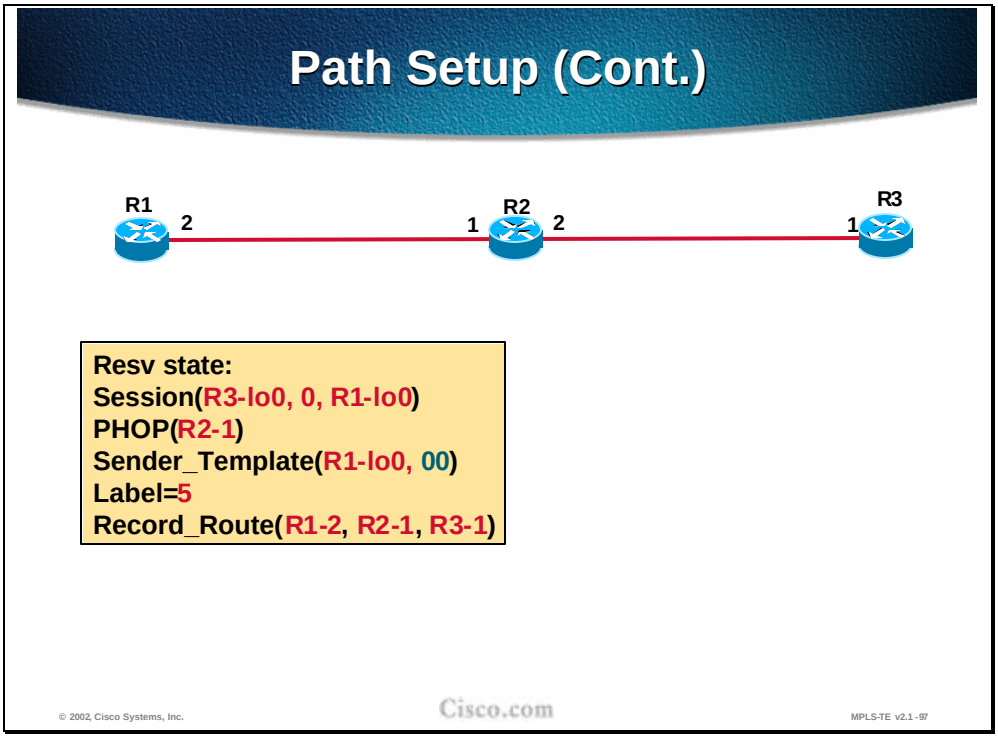
The RSVP Resv message travels back to the head-end router. On each hop (in addition to the admission control itself) label handling is performed. From the RSVP Resv message shown in the figure it is seen that the following actions were performed at the intermediate hop (R2):

The R2's interface address was put into the PHOP field and added to the beginning of the RRO list.

The incoming label (5) was allocated for the specified LSP path.

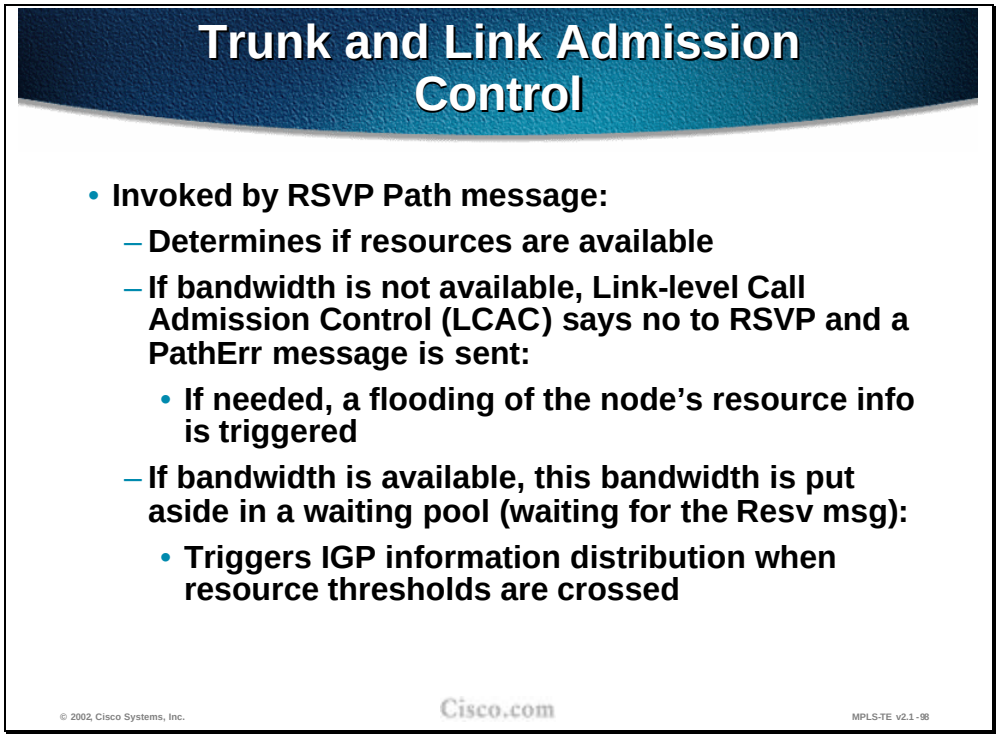
Note The label switch table is not shown but contains the information for label switching (in this particular case the label "5" is replaced with an implicit-null label ("POP").

The Resv message is forwarded towards the next hop listed in the path state block of the router. The next hop information in the path state block was established when the Path message was traveling in the opposite direction (head-end to tail-end).



When the RSVP Resv message arrives at the head-end router (R1) the LSP path setup is concluded. The label (5) allocated by the next-hop router towards the tunnel endpoint (PHOP = R2-1) is stored and the explicit route taken by the tunnel is present in RRO. The LSP tunnel is established.

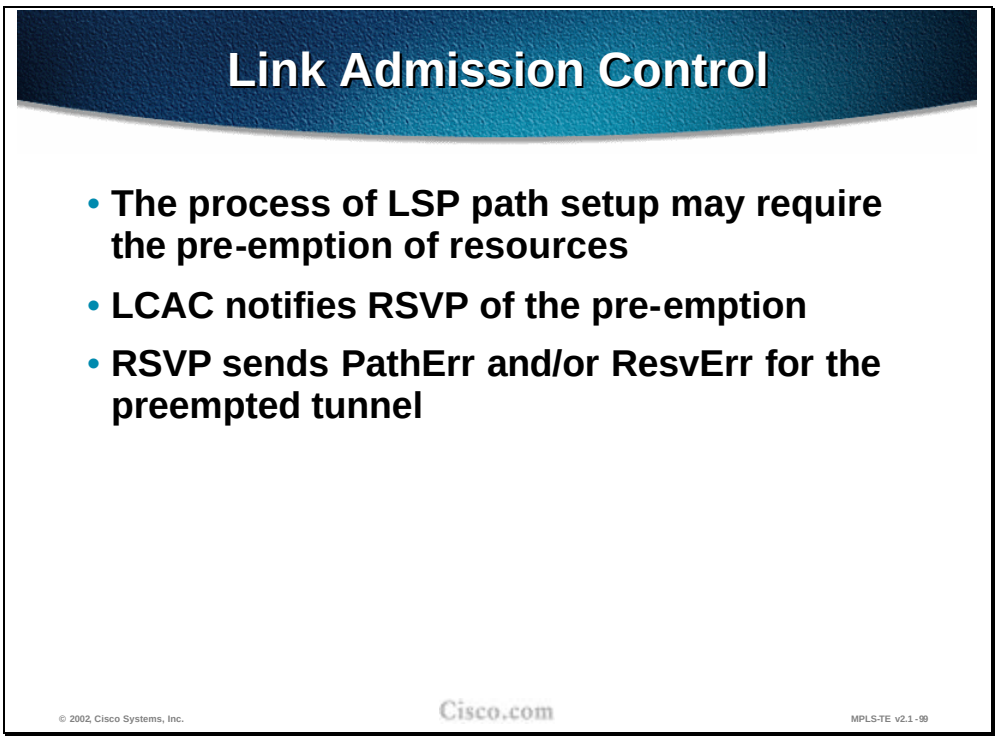
Trunk and Link Admission Control



One of the essential steps performed at each hop of the route to the LSP tunnel endpoint (the trunk) is admission control, which is invoked by the RSVP Path message traveling from the head-end to the tail-end router.

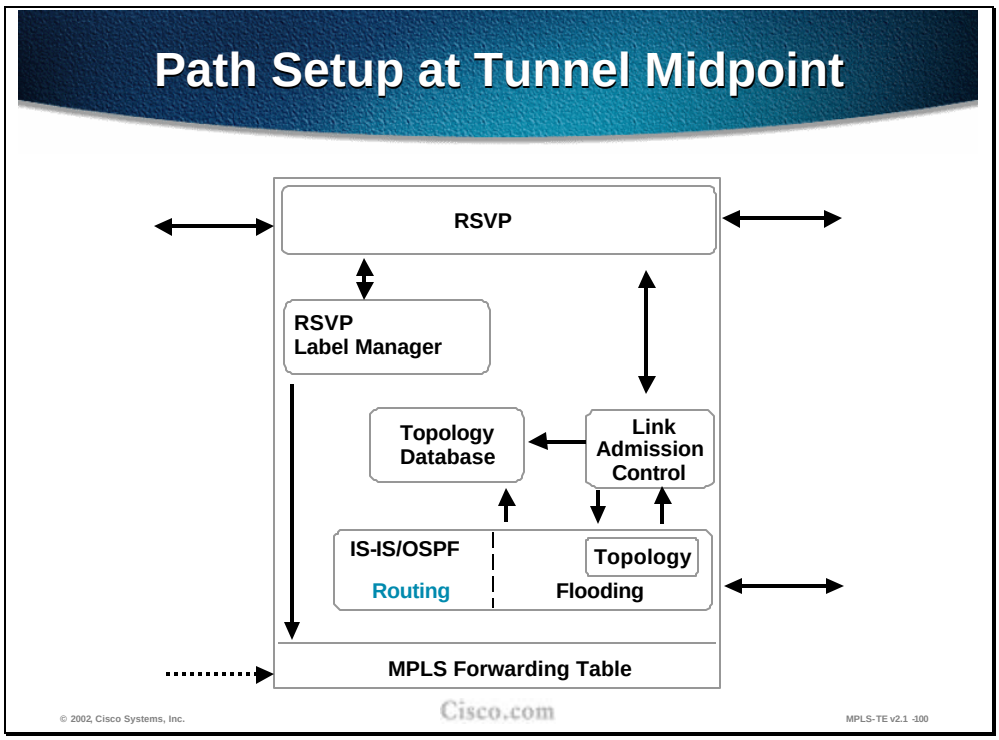
Each hop on the way determines if the available resources specified in the Session Attribute object are available. Two situations can appear:

- There is not enough bandwidth on a specified link through which the traffic trunk (tunnel) should be established. The Link-level Call Admission Control (LCAC) module informs RSVP about the lack of resources and RSVP respectively generates the RSVP Patherr message with the code "Requested bandwidth unavailable." Additionally, the flooding of the node`s resource information (by the respective link-state IGP) can be triggered as well.
- If the requested bandwidth is available, the bandwidth is reserved and is put into a waiting pool waiting for the Resv message to confirm the reservation. Additionally, if the resource threshold is exceeded, the respective IGP triggers the flooding of the resource info.



During the admission control, the priorities are checked as well. If the requested bandwidth is available, but is in use by lower priority sessions, then lower priority sessions (beginning with the lowest priority) **may** be pre-empted to free the necessary bandwidth. There are 8 levels of priority, 0 being the highest, 7 being the lowest.

When pre-emption is supported, each pre-empted reservation triggers a ResvErr and/or PathErr message with the code “Policy Control failure”.



The figure shows a conceptual model of the actions taken at the tunnel midpoint. The RSVP module is receiving signaling messages (Path) from upstream routers, passing them onto the destination, receiving the reverse path messages (Resv) from the destination, and passing them back towards the head-end of the tunnel.

The first action the RSVP performs (in addition to a regular RSVP setup) is to invoke the MPLS-TE link admission control module. The module determines if resources are available to admit the session (tunnel) or if existing sessions need to be pre-empted. The information is signaled to the RSVP module.

Depending on the resource allocation associated with the session, the RSVP module may invoke the IGP flooding module to cause the flooding of the new reservation. .

If the session was admitted by link admission control, RSVP needs to take the label received from the downstream router and establish proper entry in the MPLS forwarding table via the RSVP label manager. The label is consequently communicated to the upstream neighbor.

Practice

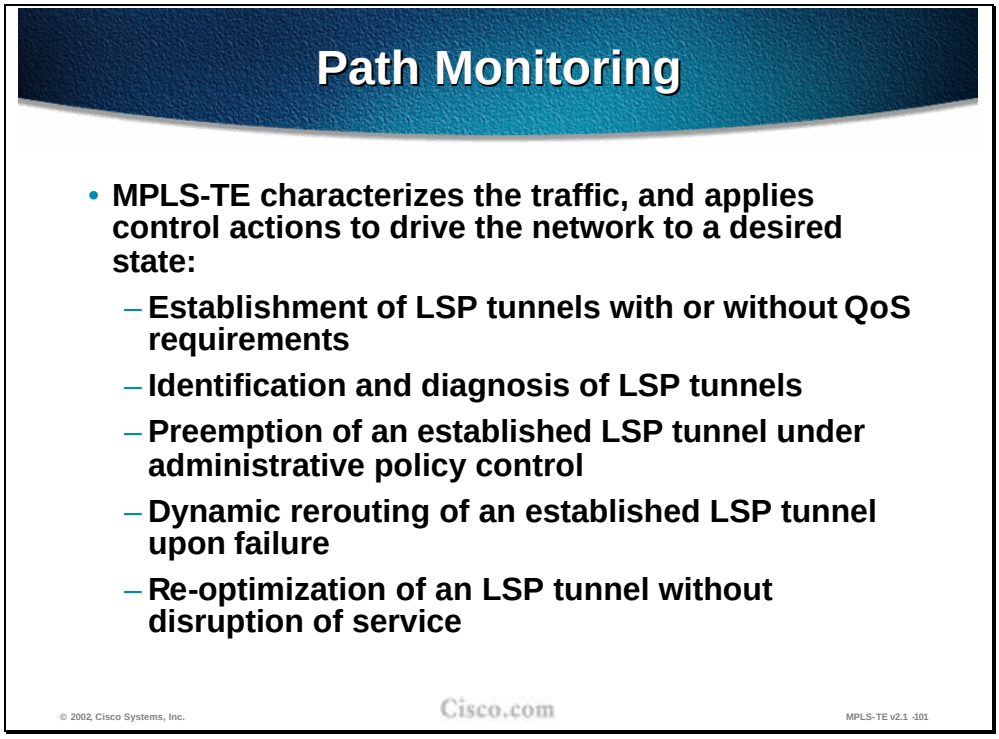
Q1) Put the steps for Link-level Call Admission Control (LCAC) signaling the inability to reserve the required bandwidth in correct order.

A) If needed, the resources information are flooded by the IGP

B) LCAC notifies the RSVP

C) RSVP in turn sends a PathErr message

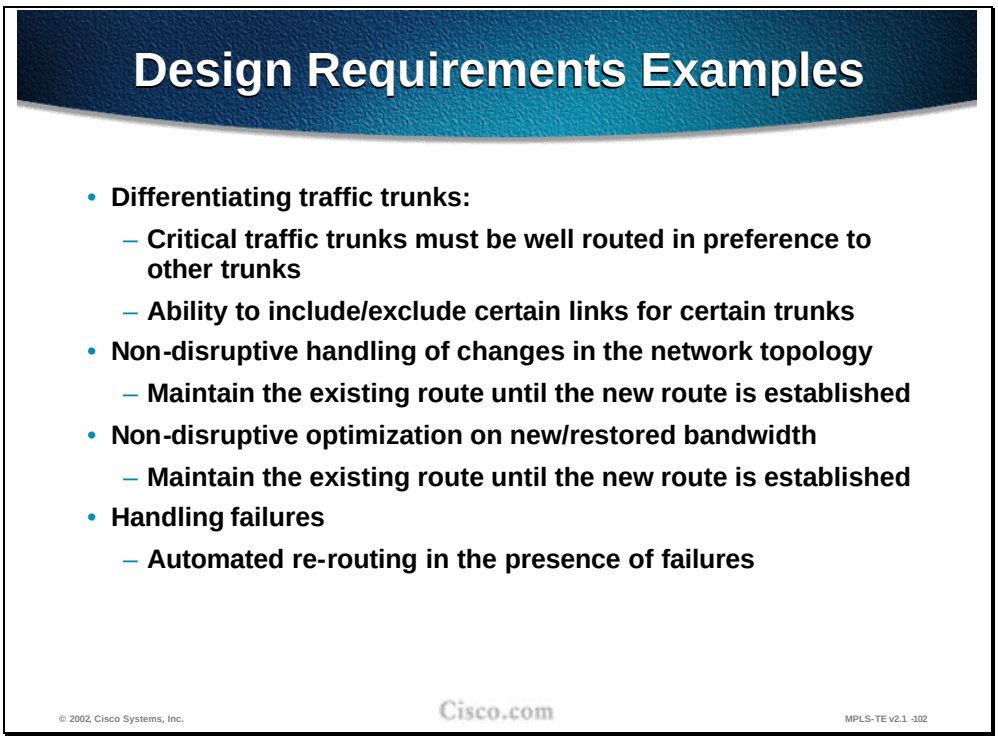
Path Monitoring



After the LSP path establishment, the path is constantly monitored to maintain the network traffic trunk in a desired state. The Quality of Service (QoS) attributes (like min. bandwidth and etc.) may be taken into account during the LSP path setup and monitored in order to provide for the re-optimization of the tunnel as well.

Various diagnostics are performed on the identified LSP paths and if necessary, the tunnels are pre-empted following the administrative policy control or dynamically rerouted in a case of network topology changes. The tunnels are also monitored for re-optimization in a case of changes in available resources.

Note It is highly desirable not to disrupt traffic while the tunnel rerouting is in progress. This smooth rerouting requirement requires establishing a new LSP tunnel and transferring traffic from the old LSP tunnel onto the new one before tearing down the old LSP tunnel. The concept is called “make-before-break”.



To constantly provide optimal paths for the tunnels and to continuously meet the traffic requirements when designing traffic trunks, several design requirements and guidelines must be followed:

- Traffic trunks must be differentiated. This differentiation includes both the traffic trunk attributes as well as the physical link attributes.. Some critical traffic trunks must have higher priority in preference to other trunks (that might have to be pre-empted). Some physical links must be included/excluded for certain trunks.
- The changes in the network topology must be non-disruptive. Before the new tunnel is fully established the existing path must be taken. It is better to lose some traffic than forwarding traffic to a “black hole.”
- The same policy as with topology changes applies to traffic trunks with changes in link attributes. If changes in available bandwidth require re-optimizing the tunnels, the former path must be used until the newly established tunnel is established.
- Traffic trunks must be automatically re-routed in a case of network failure.

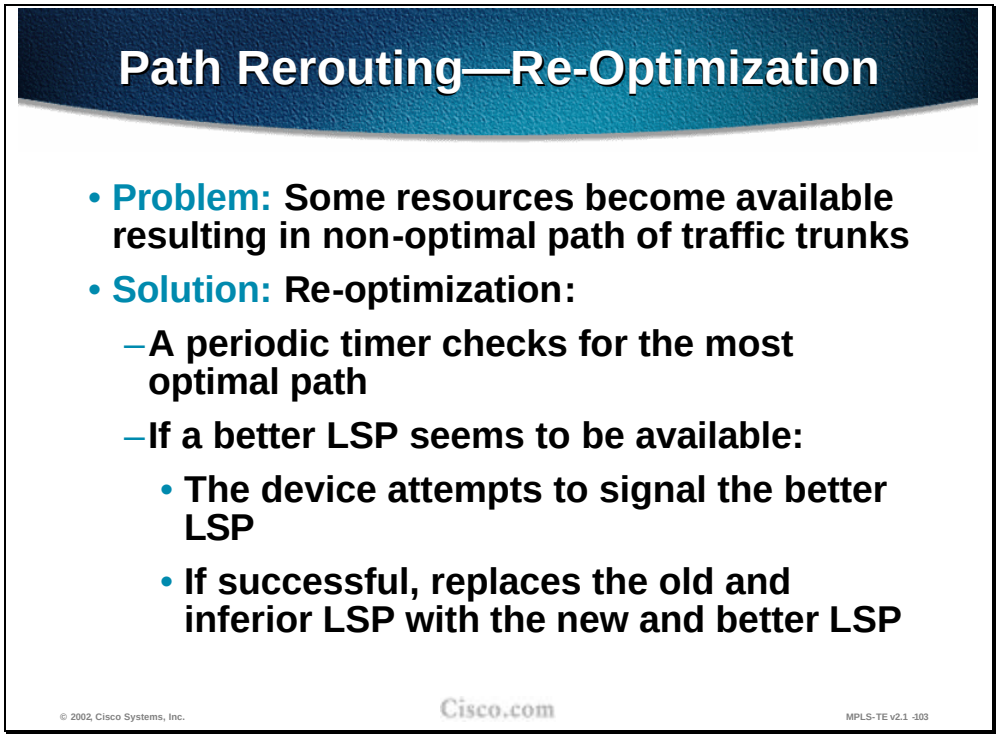
Practice

Q1) Re-optimization occurs when a device examines tunnels with established LSPs, to see if better LSPs are available.

A) True

B) False

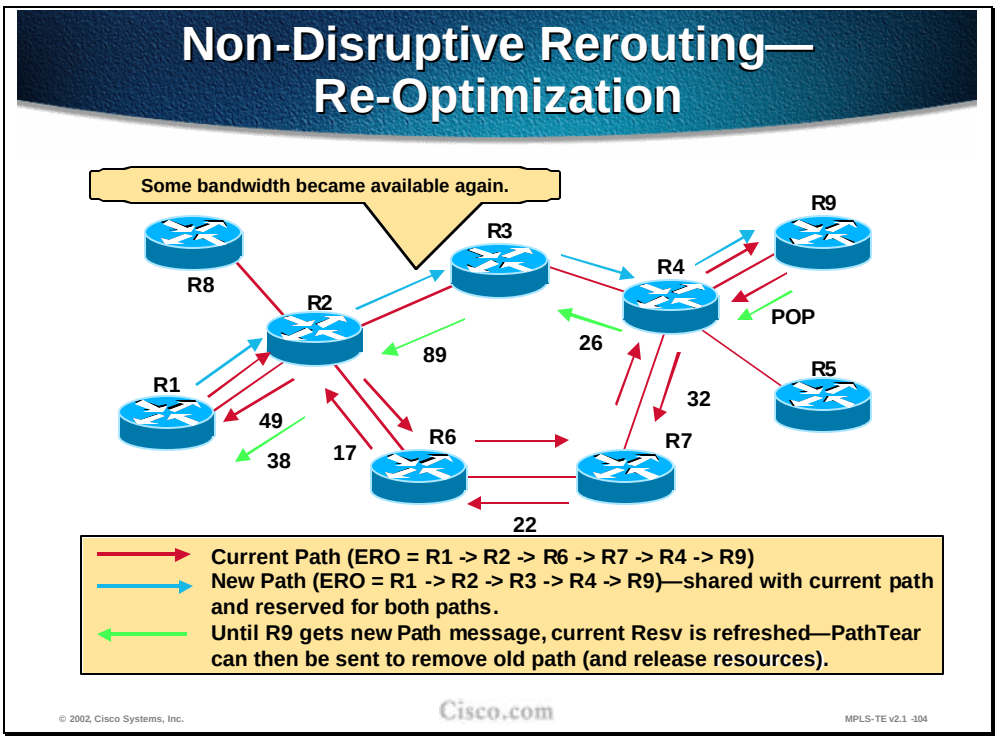
Path Re-Routing



The LSP path must be re-routed when there are physical (topology) failures or when certain changes in resource usage require it. As resources in another part of the network become available, the traffic trunks may have to be re-optimized.

The re-optimization is done on a periodic basis. At certain intervals, the checks for the most optimal paths for LSP tunnels are done and if the current path is not the most optimal, trunk re-routing is initiated.

The device (head-end router) first attempts to signal a better LSP and only after the new LSP path setup has been established successfully, will the traffic be re-routed from the former trunk to the new one.



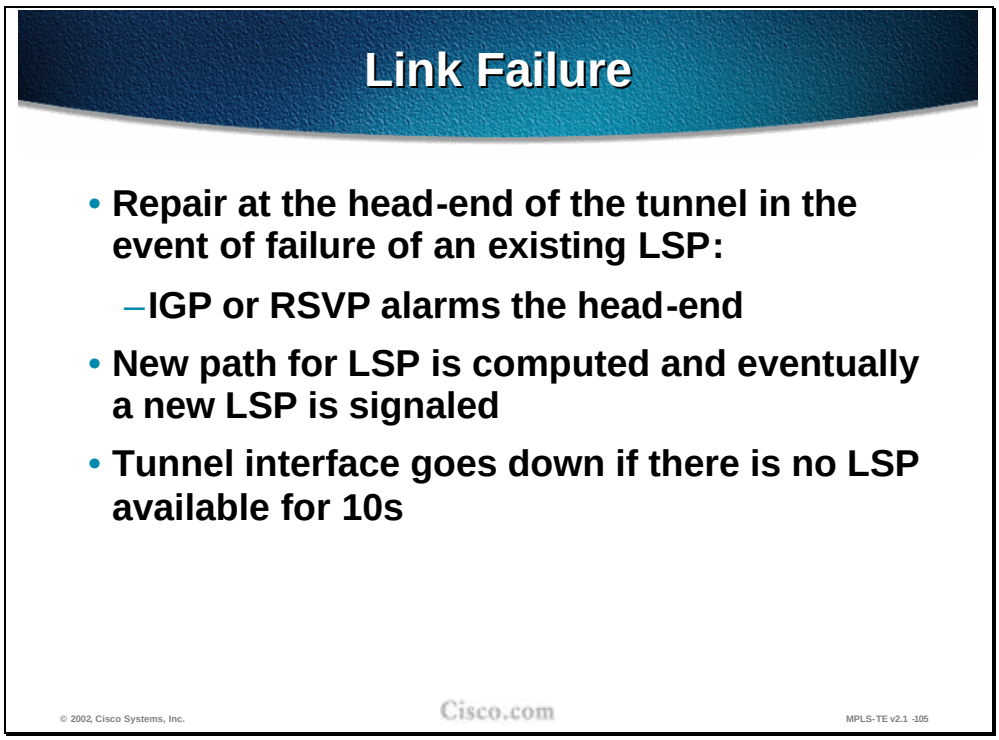
The example in the figure shows how the non-disruptive re-routing of the traffic trunk is performed. Initially the Explicit Route Object (ERO) lists the LSP path R1-R2-R6-R7-R4-R9, with R1 as the head-end and R9 as a tail-end of the trunk.

The changes in available bandwidth on the link R2-R3 dictate that the LSP Path be re-optimized. The new path R1-R2-R3-R4-R9 is being signaled and parts of the path overlap with the existing path. Still the current LSP path is used.

Note On links that are common to the old and new LSPs, resources used by the old LSP tunnel should not be released before traffic is moved to the new LSP tunnel, and reservations should not be counted twice (this might cause the Admission Control to reject the new LSP tunnel).

After the new LSP path is successfully established, the traffic is rerouted to the new path and the reserved resources of the previous path are released.. The release is done by the tail-end initiating a RSVP PathTear message.

The labels that are allocated during the RSVP Path setup are shown as well. The tail-end router assigns the implicit-null (POP) label.

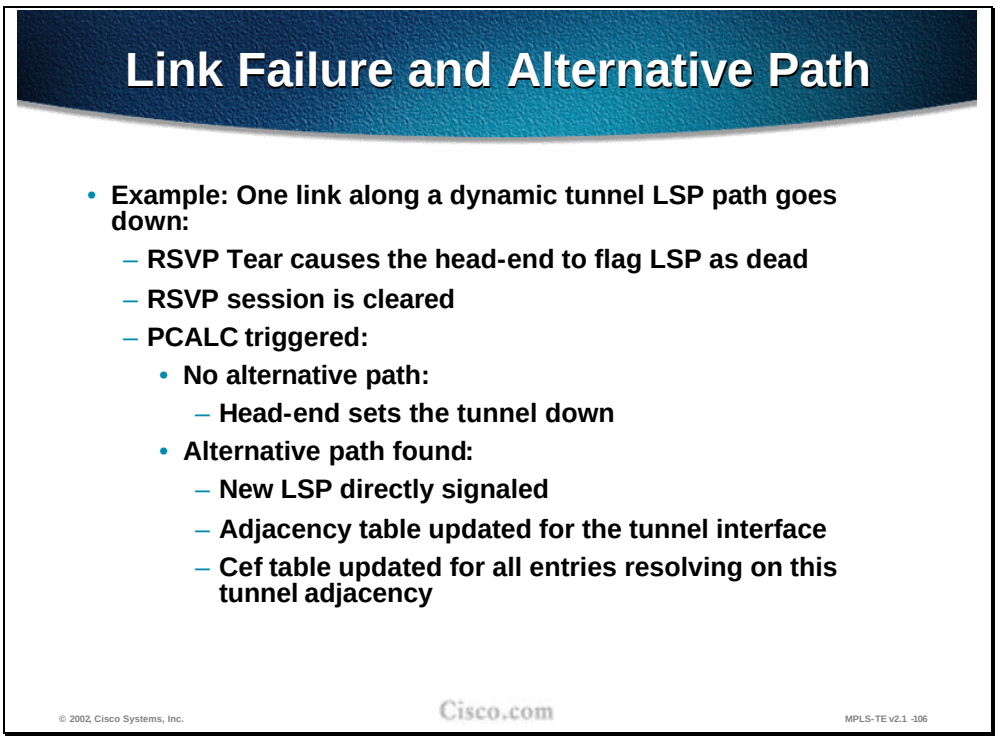


When a link passed by a certain traffic trunk fails, the head-end of the tunnel detects that failure by either:

- The IGP (OSPF or IS-IS) sends a new link-state packet with information about changes that have happened.
- RSVP alarms the failure by sending an RSVP PathTear message to the head-end.

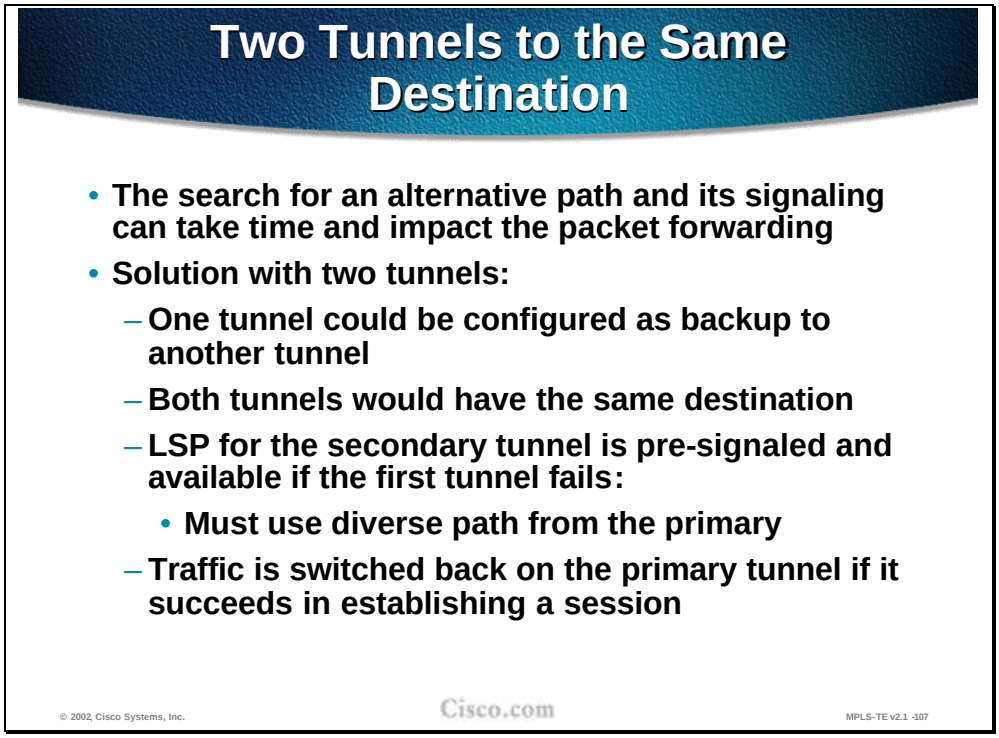
Link failure detection without any pre-configured or pre-computed path at the head-end results in a new path calculation (using a modified SPF algorithm) and consequently in a new LSP path setup.

Note	The tunnel interface used for the specified traffic trunk (LSP path) goes down if the specified LSP path is not available for 10 seconds. In the meantime the traffic intended for the tunnel continues using a broken LSP path resulting in black-hole routing.



When the router along the dynamic LSP path detects a link failure it sends the RSVP PathTear message to the head-end. This messages signals to the head-end that the tunnel is down. The head-end clears the RSVP session and a new Path calculation (PCALC) is triggered using a modified SPF algorithm. There are two possible outcomes of the calculation:

- No new path is found. The head-end sets the tunnel interface down.
- Alternative path is found. The new LSP path setup is triggered by RSVP signalization and adjacency tables for the tunnel interface are updated. Also the CEF table is updated for all the entries that resolve to this tunnel adjacency.



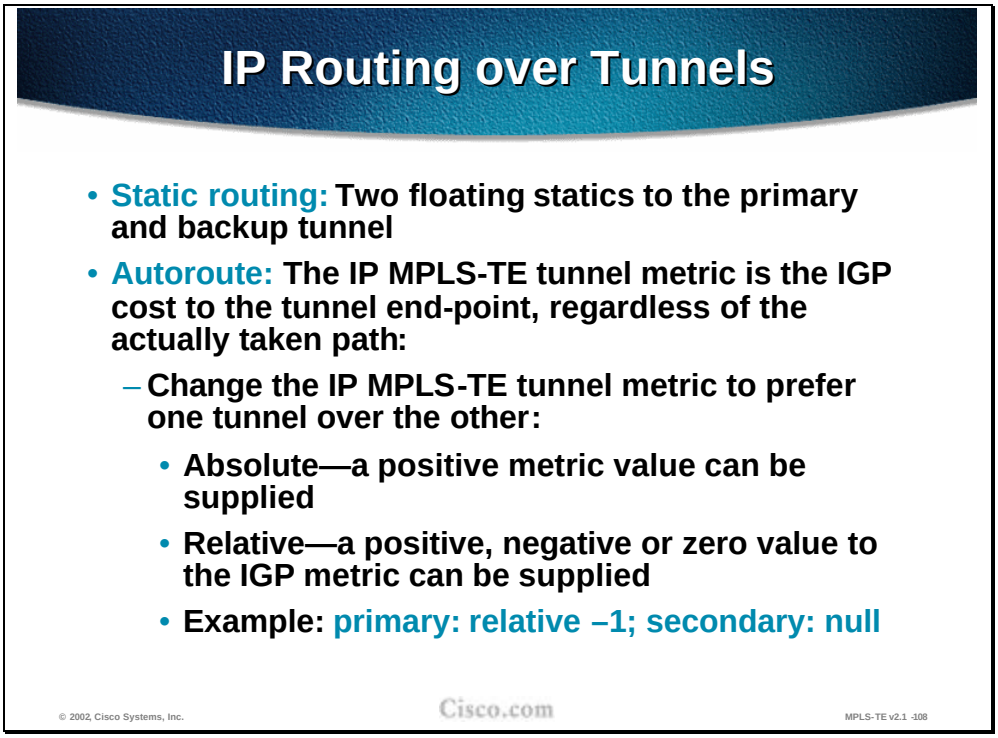
Since the time elapsed between the link failure detection and the new LSP path establishment can cause delays for critical traffic, there is a possibility of using alternative pre-established paths (backup). Therefore, there are two tunnels between the same endpoints at the same time.

Note The requirement is that pre-configured tunnels between the same endpoints must use diverse paths.

As soon as the primary tunnel fails the traffic is transitioned to the backup tunnel. The traffic is returned back to the primary tunnel if the conditions provide for the re-establishment.

Note Having two pre-established paths is the simplest form of MPLS-TE path protection. Another option is to use the pre-computed path only and establish the LSP path on-demand. In the latter case, there is no overhead in resource reservations.

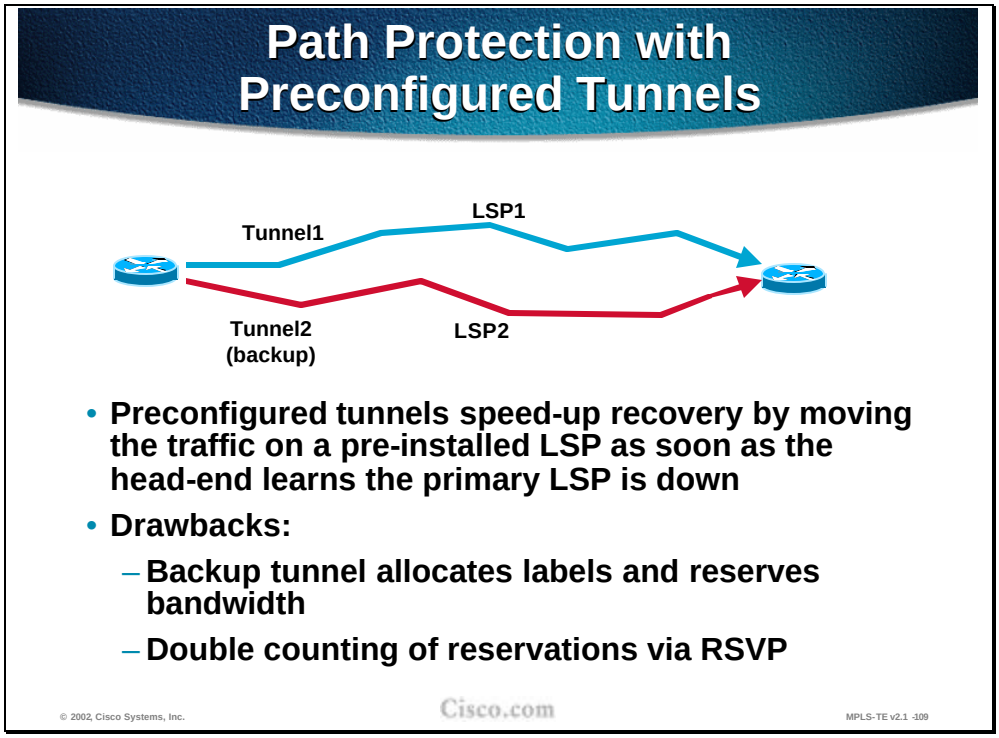
Several preparation steps must be taken in order for effective switching of the traffic between the tunnels.



In the presence of two tunnels, the primary (which is active) and the secondary, two routing options exist:

- Static routing with two floating static routes pointing to the tunnels.
- Autoroute feature" In this case the traffic trunk (tunnel) metric is the IGP cost to the tunnel end-point, regardless of the actual path taken. By adjusting this metric, the primary tunnel can be made preferential. The metric adjustments can be:
  - **Absolute** (a positive metric value is assigned to the tunnels)
  - **Relative** (the IGP metric is changed for a relative value, which can be either negative, 0 or positive). In the example, primary: relative –1; secondary: null, the secondary tunnel metric is not changed at all (0) and the primary tunnel metric is decreased by 1 (assuming that the lower metric is a better metric).

Note The autoroute feature is explained in detail in the **Assigning Traffic to Traffic Trunks** lesson, which follows this lesson.



The example in the figure shows two pre-configured tunnels: Tunnel1 (LSP1) is a primary tunnel and Tunnel2 (LSP2) is a backup tunnel. Their physical paths are diverse.

The switch-over to the backup tunnel is done at the head-end as soon as the primary tunnel failure is detected (via RSVP or via IGP). There is an obvious benefit to having a pre-configured backup tunnel. However, the solution presents some drawbacks as well:

- The backup tunnel requires all the mechanisms as the primary one. The labels must be allocated and bandwidth reserved for the backup tunnel as well.
- Looking only from the RSVP perspective, the resource reservations (bandwidth) are counted twice.

Practice

Q1) When autoroute is on, the metric of the Traffic Trunk metric is equal to:

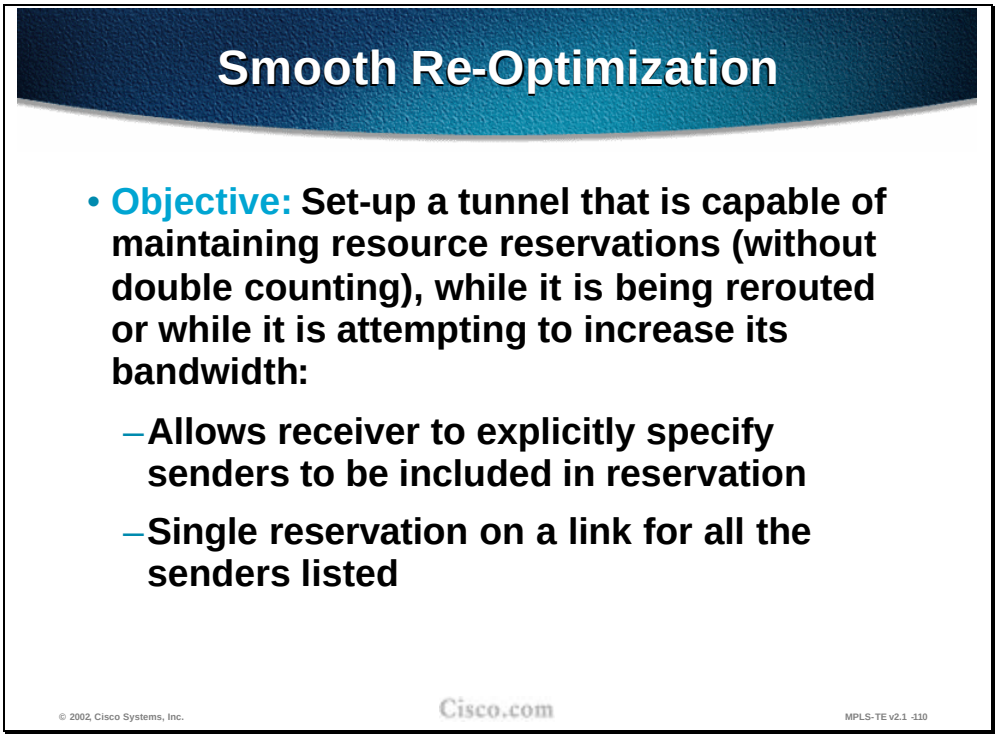
A) The IGP cost of the actual path taken.

B) The IGP cost to the tunnel end-point, regardless of the actual path taken.

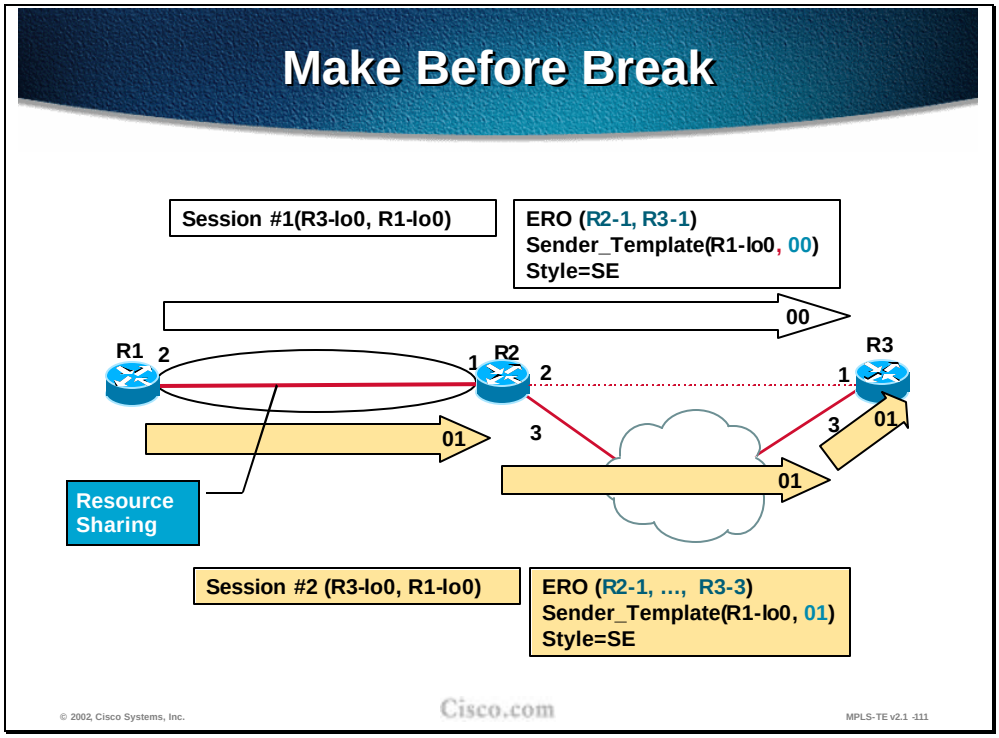
C) It is necessary to specify an absolute or relative metric.

D) The cost is always set to 0.

Path Re-Optimization



The drawbacks of configuring two tunnels and reserving resources twice can be overcome by single reservations on the same link (crossed by the same tunnel) that are not counted twice. These path re-optimization mechanisms are used during the path rerouting or while attempting to increase the tunnel bandwidth.



In the initial Path message, the head-end (ingress node) forms a Session object, and a Sender_Template (aLSP_ID is 00) with a “Shared Explicit” flag set. The tunnel set-up then proceeds according to the normal procedure.

On receipt of the Path message, the tail-end (egress node) sends a Resv message in which it indicates “Shared Explicit” path toward the ingress node. When an ingress node with an established path wants to change that path, it forms a new Path message. The existing Session object with a new LSP_ID (01) in the Sender_Template object is used. The ingress node creates an ERO for the new route. The new Path message is sent.

Since the receiver of the RSVP Path message realizes (based on the LSP_ID) that the second reservation belongs to the same session, it reserves resources only once.

Practice

Q1) How is the LSP path non-disruptively re-optimized?

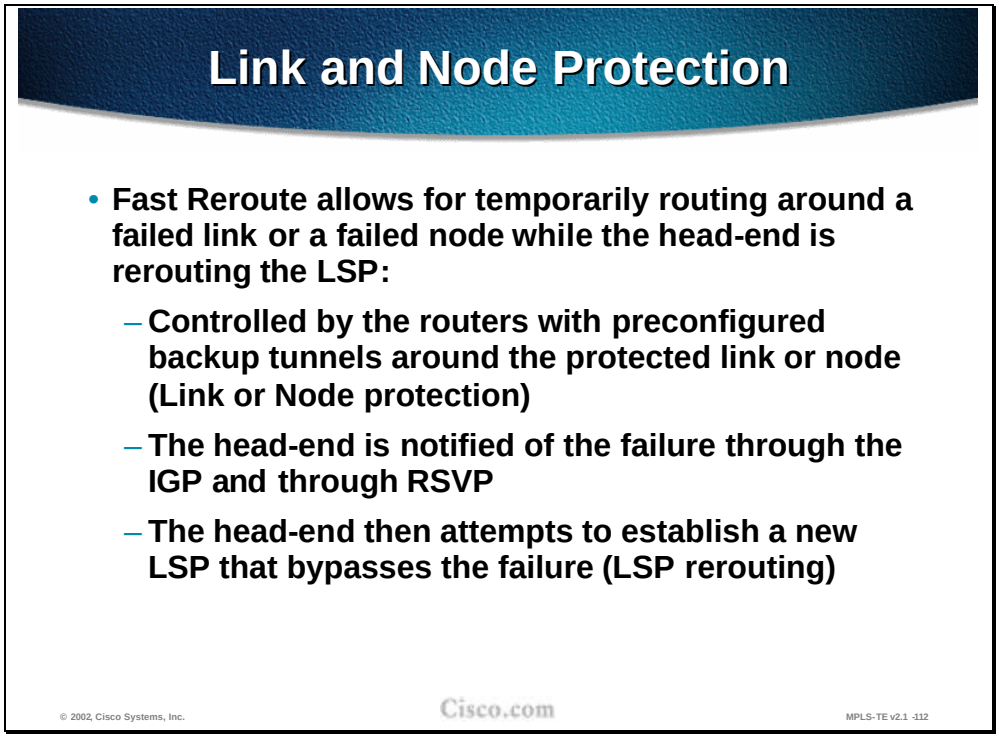
A) By establishing two LPSs in advance (make-before-break concept).

B) By establishing the new LSP first and then tearing down the old one (make-before-break concept).

C) By establishing the new LSP first and then tearing down the old one (fast reroute concept).

D) It is not possible. When the head-end is doing the re-optimization, the primary path must be torn down before stating considering an alternative option.

Link and Node Protection

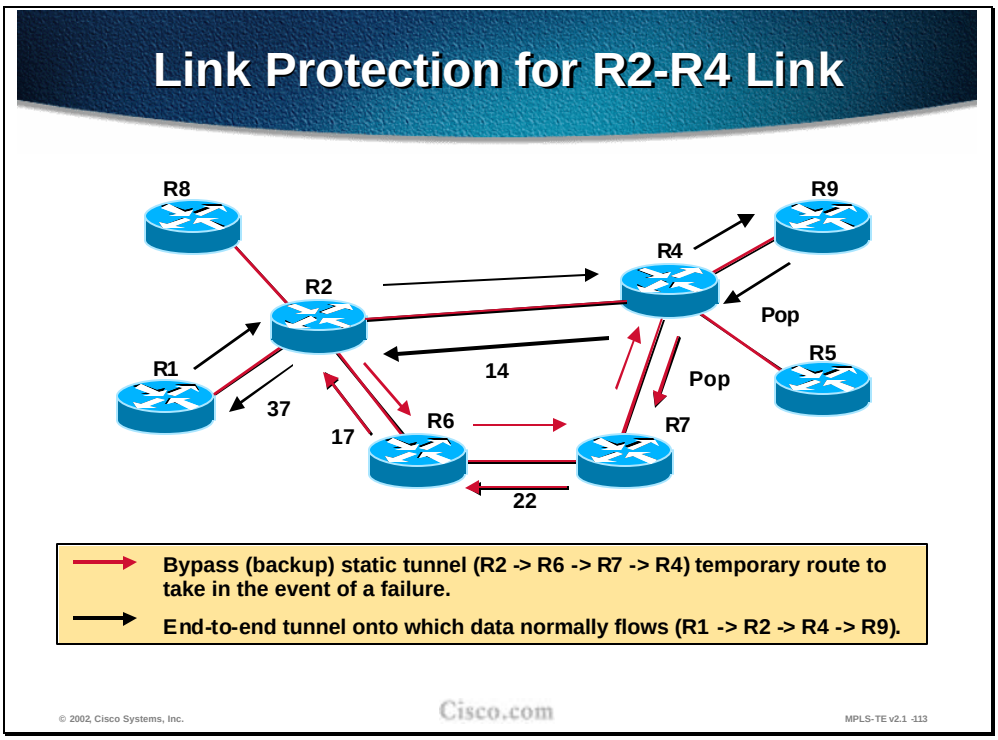


Paths for LSPs are calculated at the LSP head-end. Under failure conditions, the head-end determines a new route for the LSP. Recovery at the head-end provides for the optimal use of resources. However, due to messaging delays, the head-end cannot recover as fast as possible by making a repair at the point of failure.

To avoid packet flow disruptions while the head-end is performing new path calculation, the Fast ReRoute option of MPLS-TE is available to provide a protection from link or node failures (failure of a link or an entire router). The function is performed by routers directly connected to the failed link, as they reroute the original LSP to a pre-configured tunnel and therefore bypass the failed path.

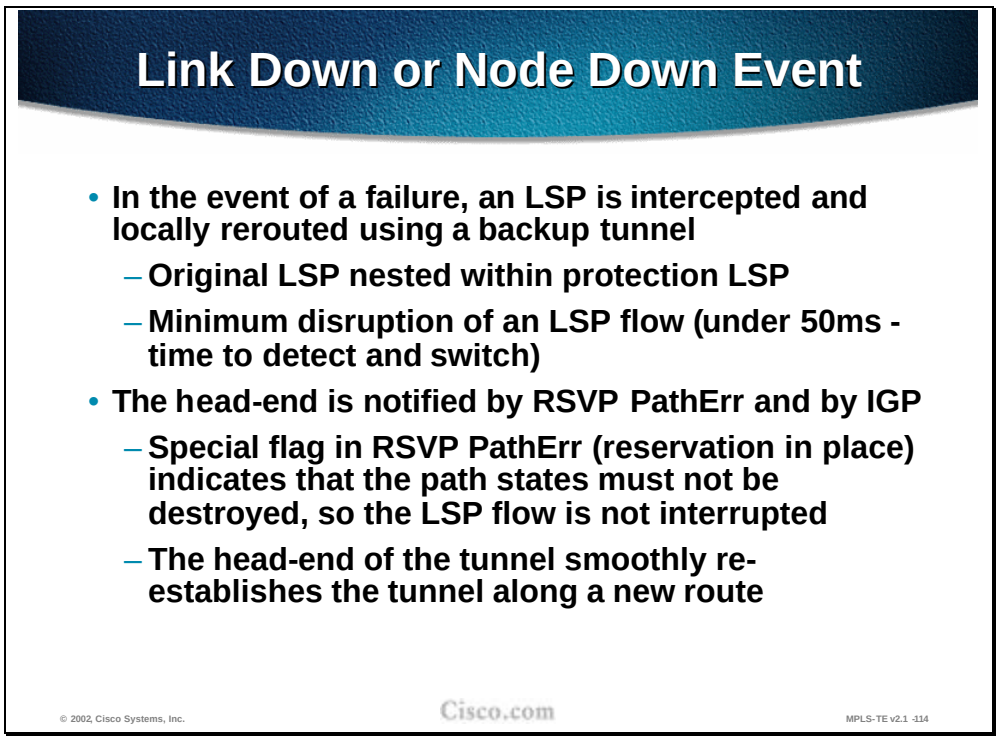
Note In terms of forwarding, it can be said that the original LSP is nested within the protection LSP.

The head-end of the tunnel is notified of the link failure through the IGP or through RSVP; the head-end then attempts to establish a new LSP.



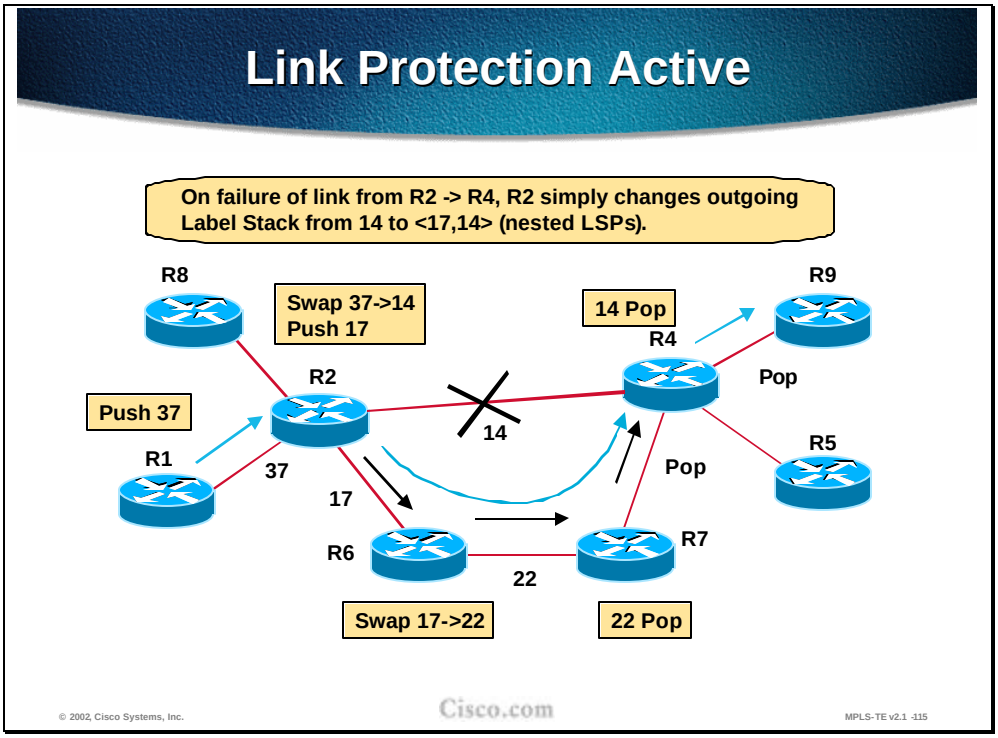
The example shows the link protection on the link between R2 and R4. The traffic trunk between R1 and R9 passes this link when the link is up and provides required resources.

The pre-configured tunnel between R2 and R4 takes the path R2-R6-R7-R4 and uses all the mechanisms of MPLS-TE (labels are allocated, resources reserved). This tunnel (link protection LSP) serves as a temporary backup in case the R2-R4 link fails.



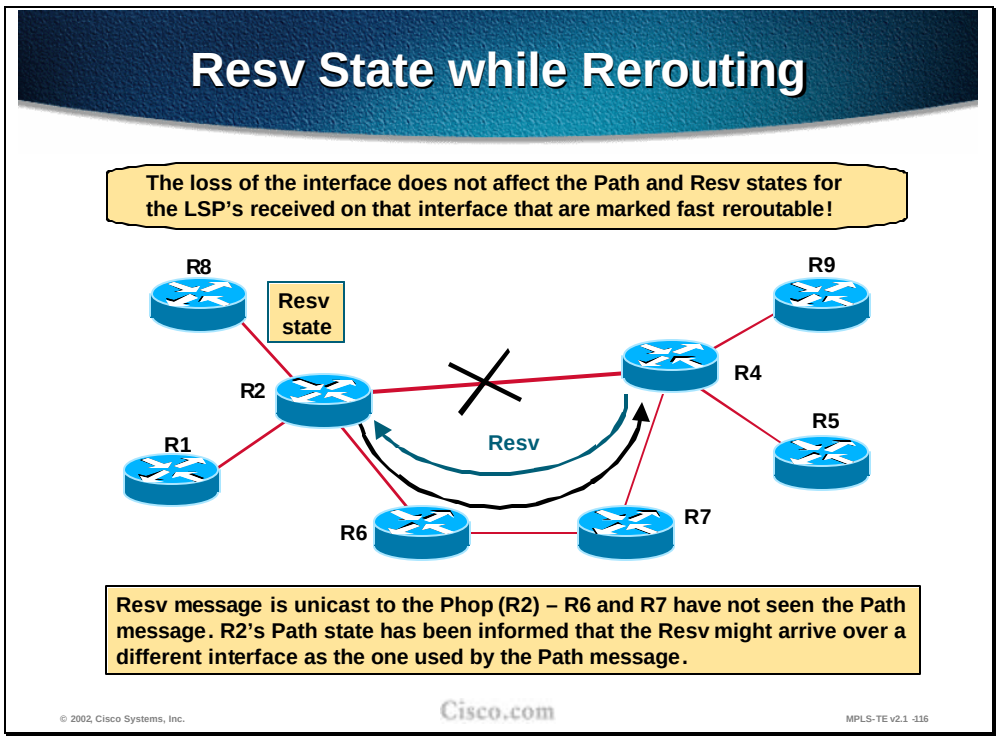
The reaction on a failure with such a pre-configured tunnel is almost instant. The local rerouting takes less than 50 milliseconds and the delay is caused only by the time it takes to detect the failed link and to switch the traffic to the link protection LSP.

When the protected link or node fails, the RSVP PathErr message and the normal IGP link-state mechanisms is used to notify the head-end. A special flag in the RSVP PathErr message indicates that the failed link already has a backup LSP, which allows for continuous forwarding of traffic while the router is re-establishing the failed path..



During the Fast reroute phase, the LSP labels must be handled as well. The router at the head-end of the link protection LSP changes the original outgoing label for the label of the pre-established LSP and nests the original label within the label stack.

In the example the original labels assigned to the LSP (listed from R1 to R9) are: 37-14-POP (implicit null). The labels on the link protection LSP are (listed from R2 to R4): 17-22-POP (implicit null). The only change in the link failure event happens on R2 where it swaps the incoming label 37 to 14. However since the link with assigned label 14 is not available, the path is moved to the link protection LSP. The original label 14 is put on the label stack of the link protection LSP to which an outgoing label 17 was assigned. Thus the original LSP path is effectively nested within the link protection path.



The RSVP states should not be affected by the Fast Reroute function. In fact, the Path and Resv messages still logically flow across the failed link. The PHOP for the Resv message traveling via R4 towards the head-end (R1) is unchanged and is still the R2 address. The R2 is aware that the response (Resv message) to the Path message might arrive via a different interface due to the link failure.

Since R6 and R7 have not seen the Path message (it passed the R2-R4 link) they could have problems in handling the Resv message flowing back (and thus maintaining RSVP states). To overcome the problem the Resv message is sent by the R4 directly to a unicast address of the R2.

Practice

Q1) What are the two major benefits of the Fast Reroute function?

A) The Fast ReRoute function allows for permanent rerouting around a failed link.

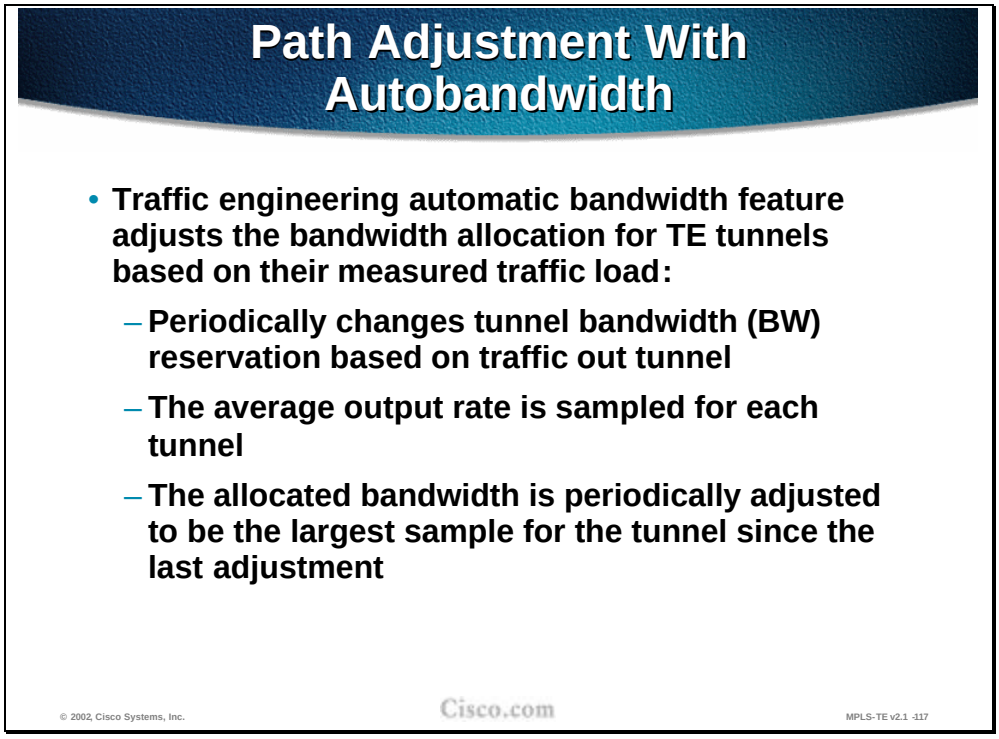
B) The Fast Reroute function is just another name for make-before-break concept.

C) The Fast ReRoute function allows for temporary rerouting around a failed link while the head-end is re-routing the path.

D) The Fast ReRoute function allows the head-end router to decide which backup route to use during the re-optimization.

E) In terms of forwarding, the original LSP path is effectively nested within the pre-established link protection path.

Path Adjustment with Autobandwidth



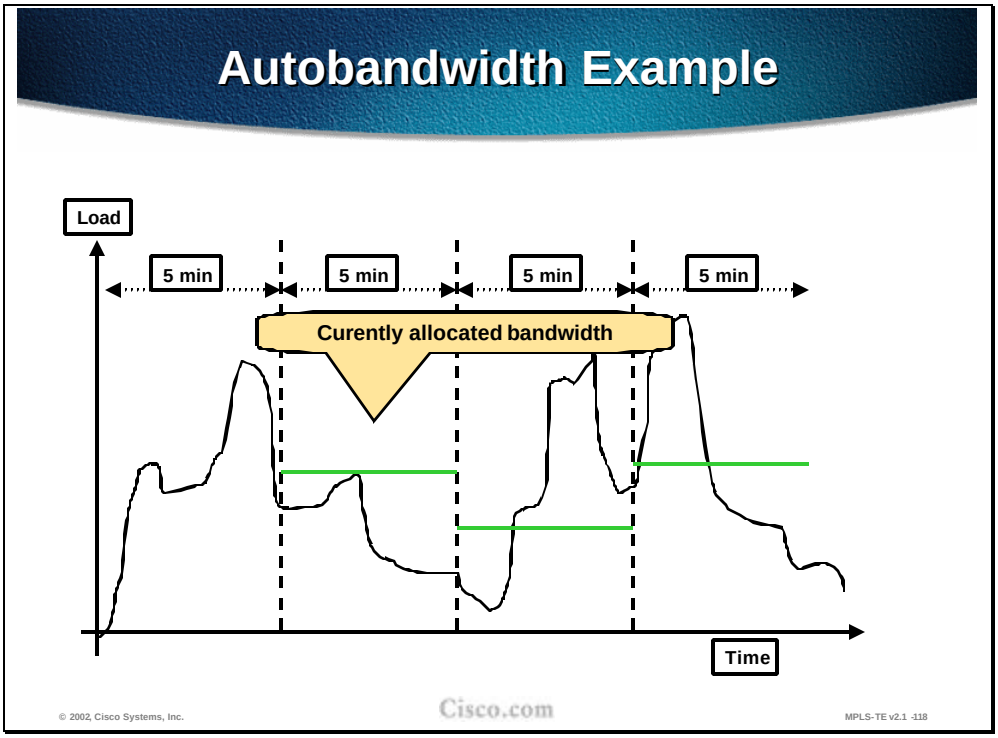
Traffic engineering automatic bandwidth adjustment provides the means to automatically adjust the bandwidth allocation for traffic engineering tunnels based on their measured traffic load.

Traffic engineering autobandwidth samples the average output rate for each tunnel marked for automatic bandwidth adjustment. For each marked tunnel, it periodically (for example, once per day) adjusts the tunnel's allocated bandwidth to be the largest sample for the tunnel since the last adjustment.

The frequency with which tunnel bandwidth is adjusted and the allowable range of adjustments is configurable on a per-tunnel basis. In addition, the sampling interval and the interval over which to average the tunnel traffic to obtain the average output rate, is user-configurable on a per-tunnel basis.

The benefit of the autobandwidth feature is that it makes it easy to configure and monitor the bandwidth for MPLS traffic engineering tunnels. If automatic bandwidth is configured for a tunnel, traffic engineering automatically adjusts the tunnel's bandwidth.

The automatic bandwidth adjustment feature treats each tunnel for which it has been enabled independently. That is, it adjusts the bandwidth for each such tunnel according to the adjustment frequency configured for the tunnel and the sampled output rate for the tunnel since the last adjustment, without regard for any adjustments previously made or pending for other tunnels.



The diagram shows the load on the tunnel and intervals of measurement. The input and output rates on the tunnel interfaces are averaged over a predefined interval (load-interval). In the example, the interval is the last 5 minutes.

The automatic bandwidth adjustments are done periodically, for example, once per day. For each tunnel for which automatic bandwidth adjustment is enabled, the platform maintains information about sampled output rates and the time remaining until the next bandwidth adjustment.

When the adjustments are done, the currently allocated bandwidth (shown as horizontal solid lines in the diagram) is reset to the maximum of:

- The largest average rate sampled during the time from the last bandwidth adjustment.
- The configured maximum value.

If the new bandwidth is not available, the previously allocated bandwidth is maintained.

Practice

Q1) What is the purpose of the autobandwidth feature of MPLS-TE?

A) To optimize the bandwidth usage by periodic adjustments of the allocated bandwidth with respect to the actual bandwidth usage by the tunnels.

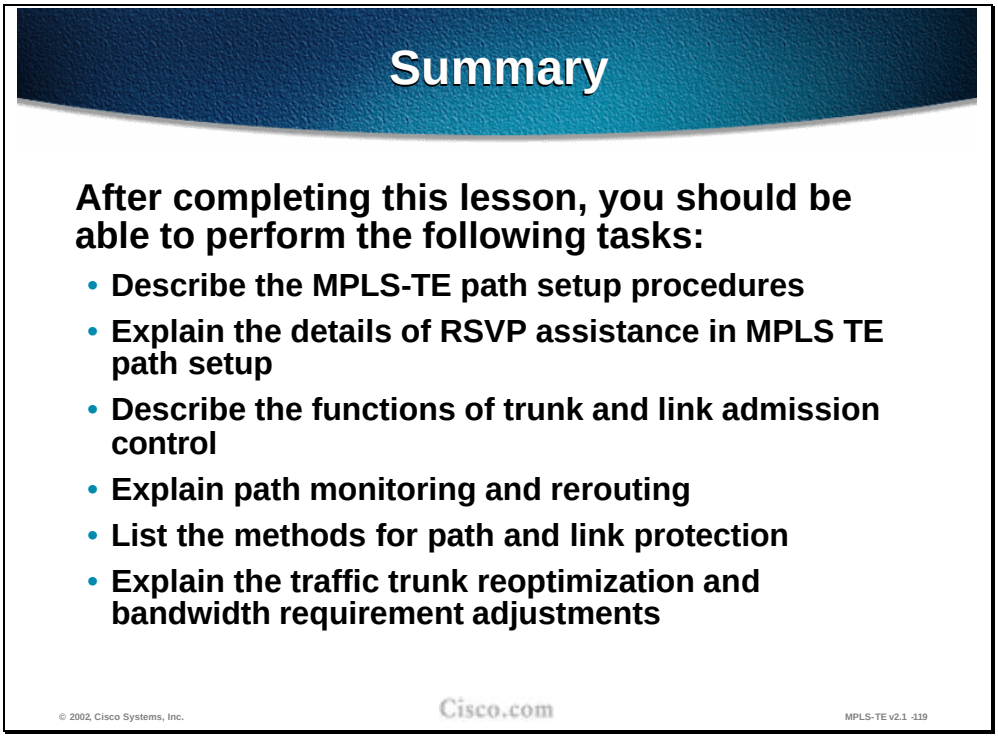
B) To optimize the bandwidth usage by periodic adjustments of the allocated bandwidth with respect to the available bandwidth in the network.

C) It periodically adjusts the tunnel's allocated bandwidth to be the average value for the tunnel since the last adjustment.

D) To measure the actual bandwidth usage and to adjust the bandwidth requested by the trunk with respect to the actual bandwidth usage by the tunnels.

Summary

This section summarizes the key points discussed in this lesson.



Next Steps

After completing this lesson, go to:

- Assigning Traffic to Traffic Trunks

Lesson Review

Instructions

Answer the following questions:

1. How is an LSP path setup initiated?
2. Explain the main components (objects) of RSVP messages that help establish the MPLS-TE tunnel.
3. How does Link-level Call Admission Control (LCAC) signal the inability to reserve the required bandwidth?
4. How is the LSP path non-disruptively rerouted?
5. List the LSP path protection methods.
6. What is a major benefit of a Fast reroute function?
7. What is the purpose of autobandwidth feature of MPLS-TE?

Assigning Traffic to Traffic Trunks

Overview

This lesson describes several well-known drawbacks of traditional IP routing: full-mesh requirements in IP-over-ATM overlay model and lack of traffic engineering.

Importance

This lesson is a mandatory for the students planning to improve the usage of their network resources with MPLS traffic engineering.

Objectives

Upon completion of this lesson, the learner will be able to perform the following tasks:

- List the mechanisms that can be used to assign traffic to traffic trunks
- Describe the auto-route mechanism
- Describe the forwarding-adjacency feature
- Use static routes to assign traffic to traffic trunks
- Use static routes or auto-route toward next-hop routers in combination with exterior routing protocols

Learner Skills and Knowledge

To fully benefit from this lesson, you must have these prerequisite skills and knowledge:

- Cisco Certified Internetwork Professional (CCIP) level of knowledge or equivalent level of IP routing and Cisco IOS knowledge as well as solid understanding of MPLS and link state protocols (OSPF or Integrated IS-IS).

Mandatory Prerequisites:

- AMVS course

Optional prerequisites:

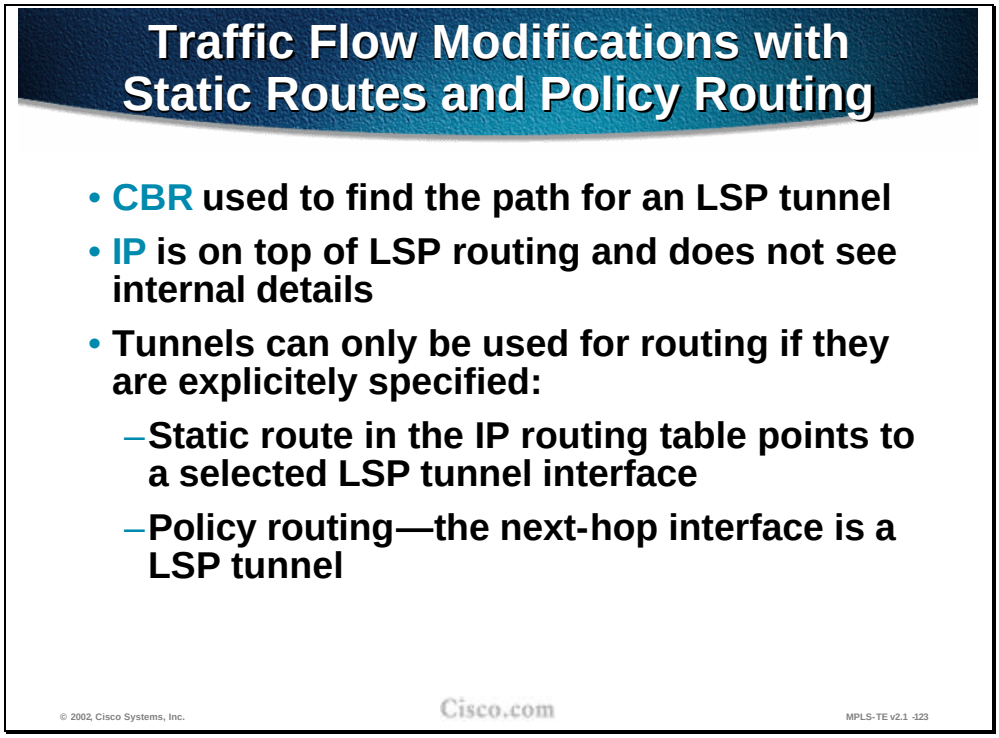
- CISIS course for students deploying MPLS TE in IS-IS environments

Outline

This lesson includes these sections:

- Overview
- IP Forwarding Database Modification with Static Routing and Policy Routing
- IP Forwarding Database Modification with Autoroute
- Autoroute Example
- Summary
- Lesson Review

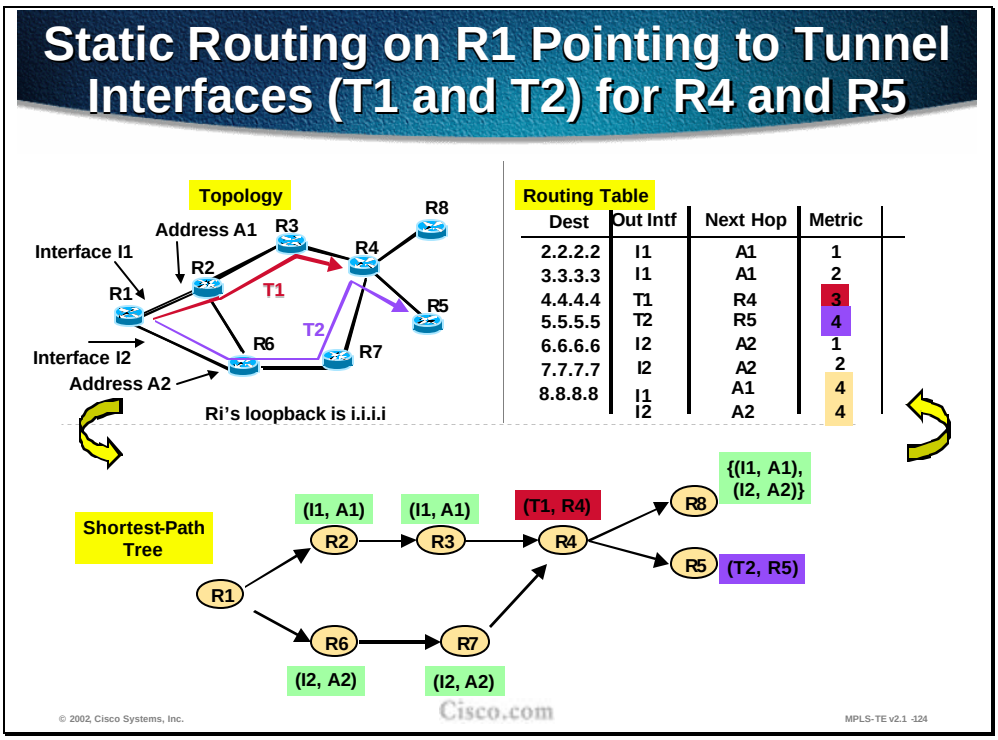
IP Forwarding Database Modification with Static Routing and Policy Routing



The Label Switched Path (LSP) is computed by the Constraint-Based Routing (CBR), which takes the resource requirements into consideration as well. When the LSP path is established for the trunk, the traffic can flow across it. From the IP perspective, an LSP path is a simple tunnel.

These engineered tunnels can only be used for the IP routing if the tunnels are explicitly specified for routing:

- Via static routes that point to the tunnel
- Via policy routing that sets a next-hop interface to the tunnel



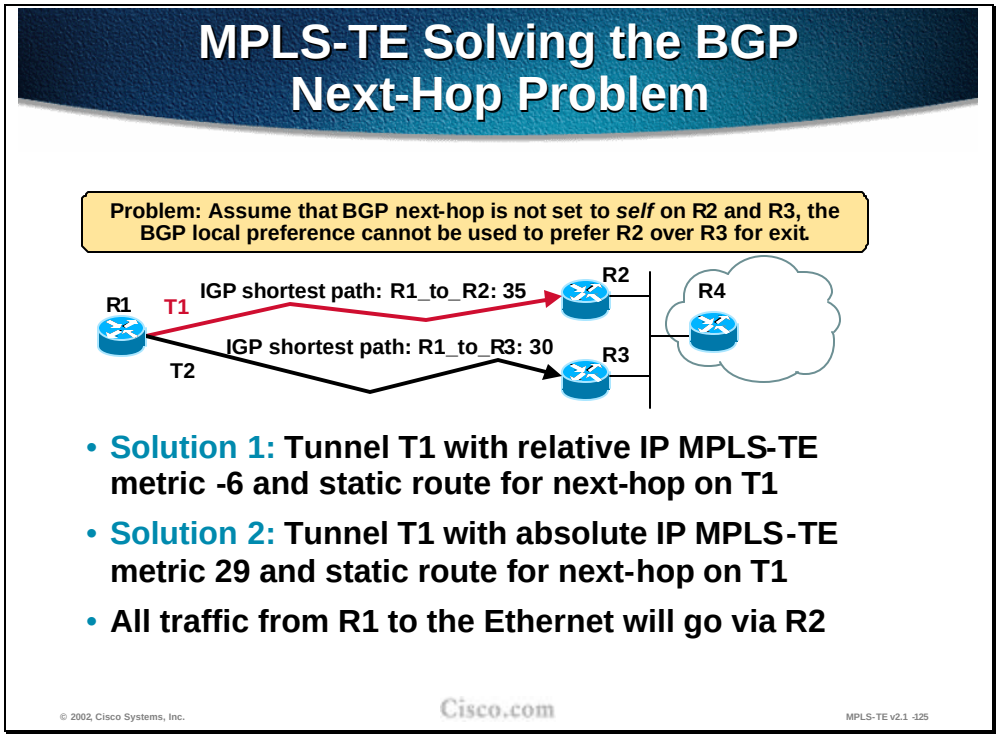
The example topology shows two engineered tunnels: T1 (between R1 and R4) and T2 (between R1 and R5). The loopback addresses on each router are in the form i.i.i.i where i is the router number (e.g. R5's loopback address is 5.5.5.5). The metric on each of the interfaces is set to 1.

R1 has two physical interfaces: I1 and I2, and two neighboring routers (next hops) with addresses A1 and A2 respectively.

The routing table lists all eight loopback routes and associated information. Only the statically configured destinations (R4 and R5) list tunnels as their outgoing interfaces. For all other destinations the normal IGP routing is used and results in physical interfaces (along with next hops) as the outgoing interfaces towards these destinations. The metric to the destination is normal IGP metric.

Note Even for the destination that is behind each of the tunnel endpoints (R8), the normal IGP routing is performed if there is no static route to the traffic engineered tunnel.

The SPF calculates paths to destinations in its usual way with the exception of the paths for the tunnels where a constraint-based computation is performed.



There are several interesting problems associated with MPLS-TE routing. In the example in the figure, there are two internal routers serving exits to external BGP destinations – R4 (including the Ethernet between R2, R3 and R4). These two routers are also the endpoints of two MPLS-TE tunnels. The BGP next-hop *self* is not configured on R2 and R3 (the R4 remains the next hop), which prevents the use of local preference to prefer R2 to R3 as an exit to external destinations. The best IGP metric is used instead, and R3 becomes the preferred exit.

There are two possible solutions to the problem. Both solutions use the static routes to MPLS-TE tunnels and some metric modifications (relative or absolute metrics):

- The metric for tunnel T1 is set to a relative value decreased by 6 (35-6=29) and the static route for the Ethernet is configured to the tunnel T1
- The metric for tunnel T1 is set to an absolute value of 29 and the static route for the Ethernet is configured to the tunnel T1

In both cases all the traffic for the Ethernet will flow via R2 since its metric less than 30

Practice

Q1) What is the major drawback of the static assignment of traffic to MPLS-TE tunnels?

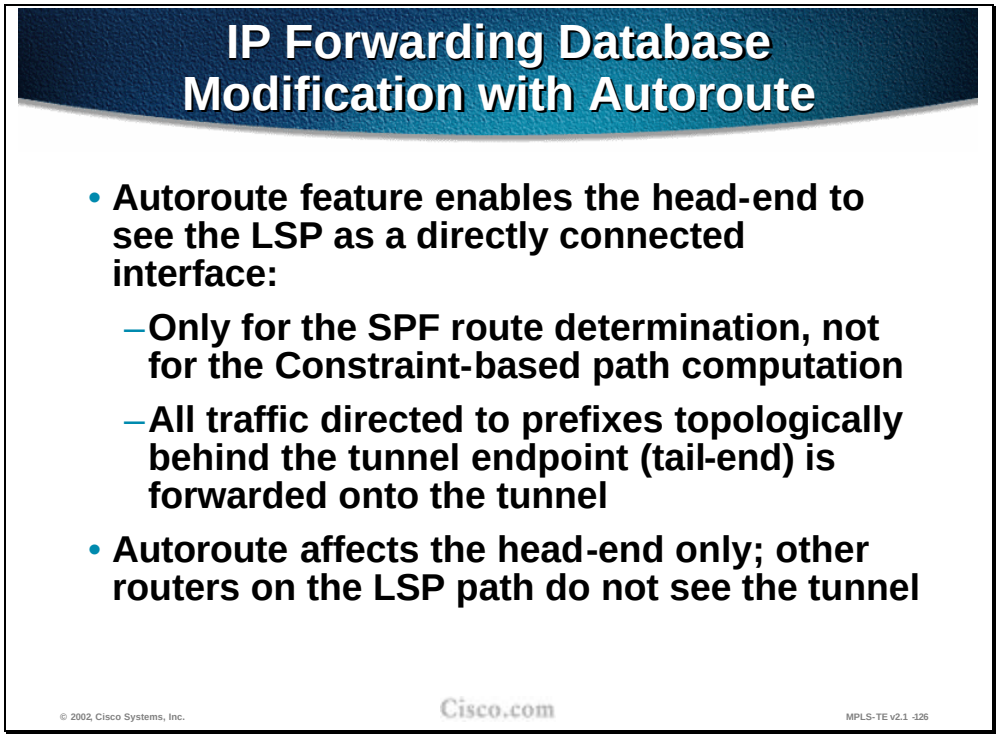
A) The static routes can be used only for destinations that are attached to the tunnel end-point.

B) The tunnel is used only for explicit routes that are statically defined.

C) When assigning a static route pointing to a tunnel it doesn't end up in CEF cache.

D) Floating static routes to tunnels are not supported.

IP Forwarding Database Modification with Autoroute

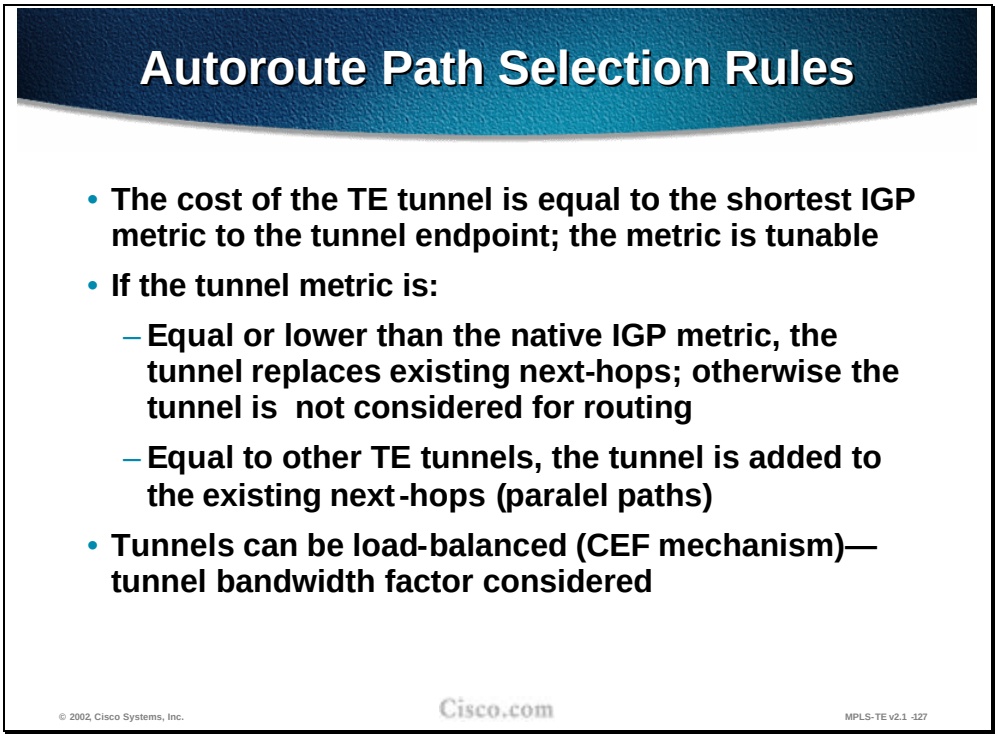


To overcome the problems resulting from static routing configuration onto the MPLS-TE tunnels, the autoroute feature of Cisco IOS was introduced. The autoroute feature enables the head-end routers to see the MPLS-TE tunnel as a directly connected interface and use it in its modified SPF computations.

The MPLS-TE tunnel is only used for normal IGP route calculation (at the head-end only) and is not included in any constraint-based path computation.

The autoroute feature results in all the prefixes topologically behind the MPLS-TE tunnel endpoint (tail-end) to be reachable via the tunnel itself (unlike with static routing where only statically configured destinations were reachable via the tunnel).

The autoroute feature affects the head-end router only and has no effect on intermediate routers. These routers still use normal IGP routing for all the destinations.



Since the autoroute feature includes the MPLS-TE tunnel into the modified SPF path calculation, the metric of the tunnel plays a significant role. The cost of the tunnel is equal to the best IGP metric to the tunnel endpoint regardless of the LSP path. The tunnel metric is tunable using either relative or absolute metrics.

When installing the best paths to the destination, the tunnel metric is compared to other existing tunnel metrics and to all the native IGP path metrics. The lower metric is better and if the MPLS-TE tunnel has an equal or lower metric than the native IGP metric, it is installed as a next hop to the respective destinations.

If there are tunnels with equal metrics they are installed in the routing table and provide for load balancing. The load balancing is done proportionally to the configured bandwidth of the tunnel.

Practice

Q1) Which path is preferred when using the autoroute feature for the destinations behind the tunnel endpoints?

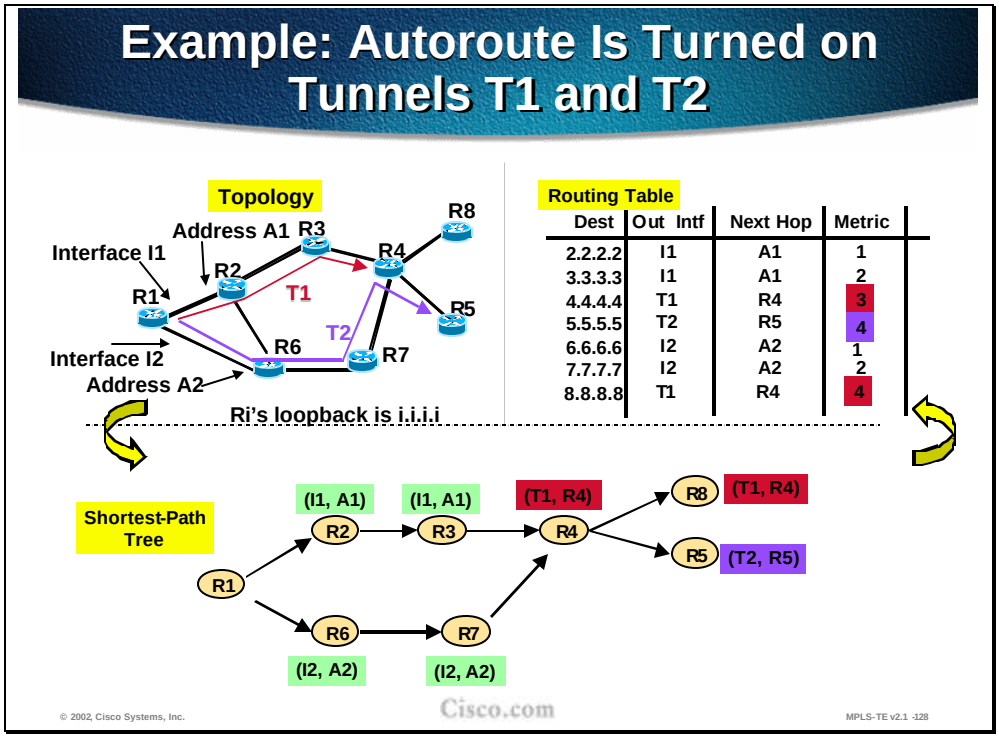
A) The tunnel if its metric is equal than the native IGP metric.

B) The tunnel if its metric is lower than the native IGP metric.

C) The tunnel is always preferred over the native IGP path.

D) The tunnel if its metric is equal or lower than the native IGP metric.

Autoroute Example



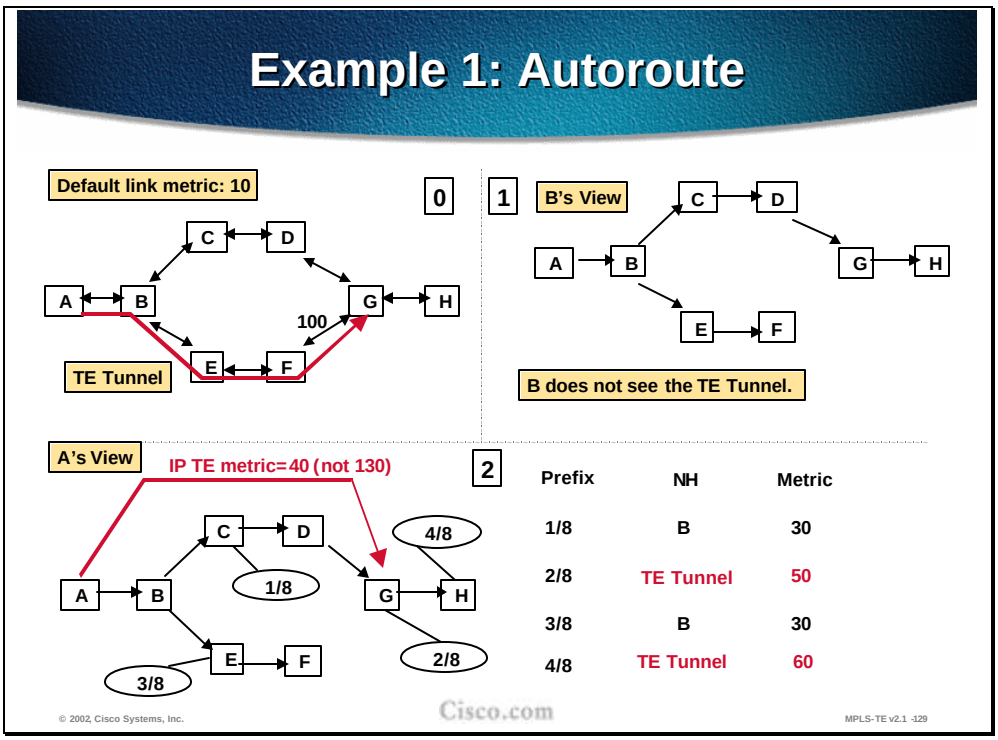
The example topology shows two engineered tunnels: T1 (between R1 and R4) and T2 (between R1 and R5). The loopback addresses on each router are in the form i.i.i.i where i is the router number (e.g. R5's loopback address is 5.5.5.5). The metric on each of the interfaces is set to 1.

R1 has two physical interfaces,I1 and I2, and two neighboring routers (next hops) with addresses A1 and A2 respectively.

The routing table lists all eight loopback routes and associated information. The autoroute feature is turned on for both tunnels (T1 and T2) at their head-end (router R1).

The routing table shows all destinations at the endpoint of the tunnel and behind it (R8) as reachable via the tunnel itself. The metric to the destination is normal IGP metric.

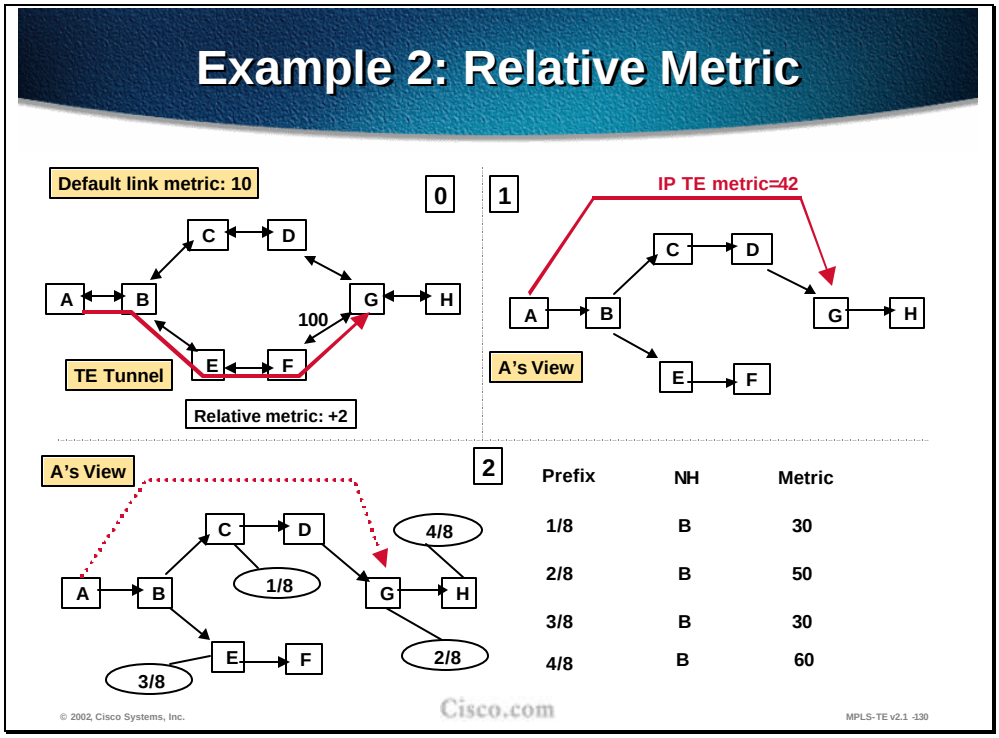
Note Unlike the static route configuration with autoroute feature the destinations behind the tunnel endpoints (R8 in this example) are reachable via the tunnel as well.



The following examples show the effect of the autoroute feature. In the first situation there is an MPLS-TE tunnel configured between A and G. The tunnel is seen for routing purposes only by the head-end (A). Intermediate routers do not see the tunnel nor do they take it into consideration for route calculations.

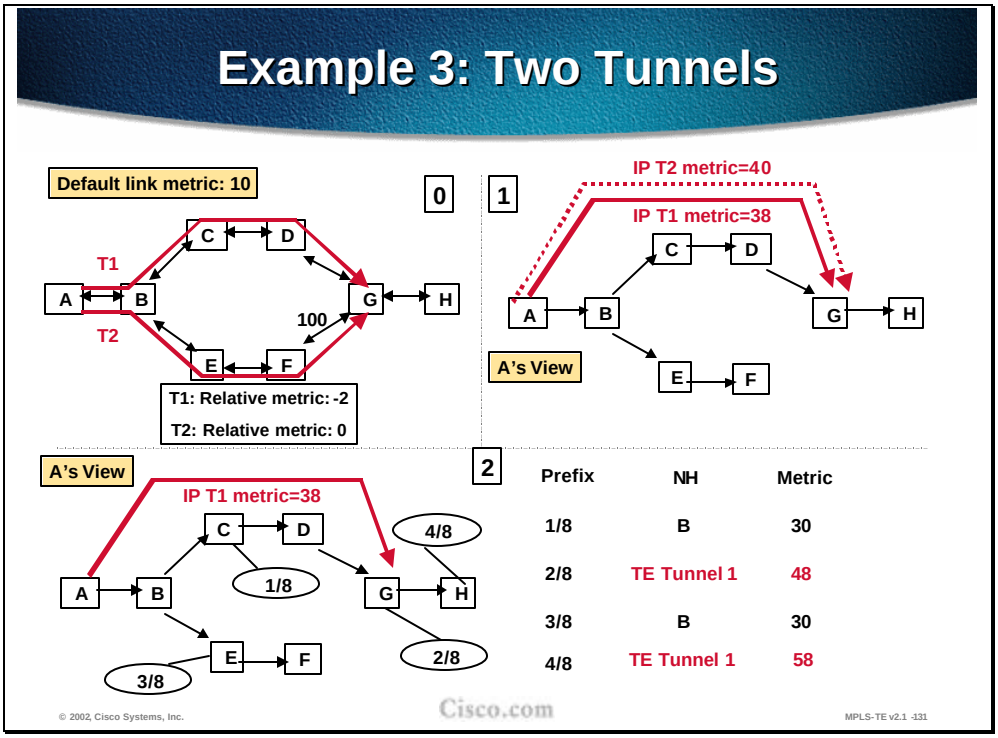
Although the LSP paths follow the path A-B-E-F-G, the tunnel cost is the best IGP metric to the tunnel endpoint. The link metric between F and G is 100. All other metrics are set to 10. Although the LSP path passes the F-G link, the overall metric of the tunnel is 40 (the sum of metrics on the best IGP path A-B-C-D-G).

In the routing table all the networks topologically behind the tunnel endpoint (networks 2 and 4) are reachable via the tunnel itself since the MPLS-TE tunnel metric is equal to the native IGP metric, it is installed as a next hop to the respective destinations. This is the effect of the autoroute feature. The metrics to these two networks are the sums of the tunnel metric (40) and the native IGP metric from the tunnel endpoint to the respective networks.



The tunnel metrics can be tuned and either relative or absolute metrics can be used. In the second example, the LSP path still takes the same path (A-B-E-F-G), but the tunnel metric is set to *relative+2*. This setting results in a tunnel metric of 42.

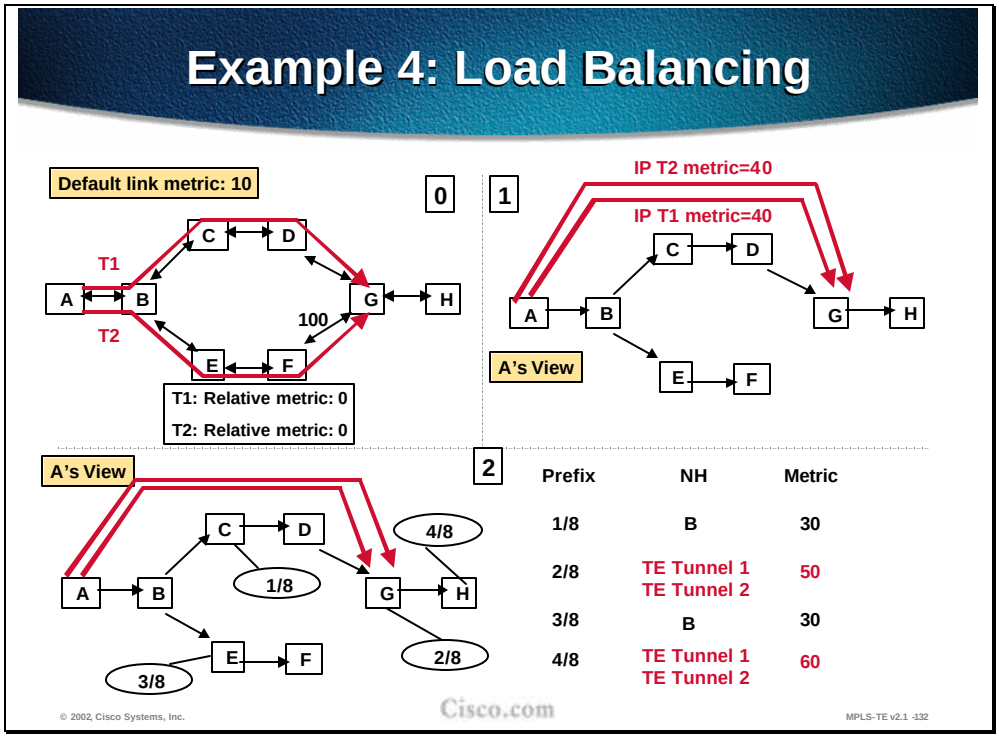
When the tunnel is considered in the IGP calculation the native IGP metric (40) is lower than the tunnel metric (42) for all the destinations topologically behind the tunnel endpoint. As a result, all the destination networks (1 to 4) are reachable via router B instead of via the TE tunnel as can be seen from the routing table.



In the third example of the autoroute feature there are two configured MPLS-TE tunnels: T1 (following the LSP path A-B-C-D-G) and T2 (following the path A-B-E-F-G). The metric of T1 is tuned to relative-2. This setting results in a T1 metric of 38. The T2 metric is unchanged and is set to the best IGP metric to the tunnel endpoint (40).

Both tunnel metrics are equal to or less than the native IGP metric to the tunnel endpoints (40). Therefore both the tunnels are used: T1 as a primary tunnel and T2 as a secondary tunnel. All the destinations behind the T1 endpoint are reachable via the tunnel itself due to the autoroute feature. As seen from the routing table, the networks 2 and 4 are reachable via T1 and their respective metrics are sums of the tunnel metric and the native IGP metric from the tunnel endpoint to the respective networks.

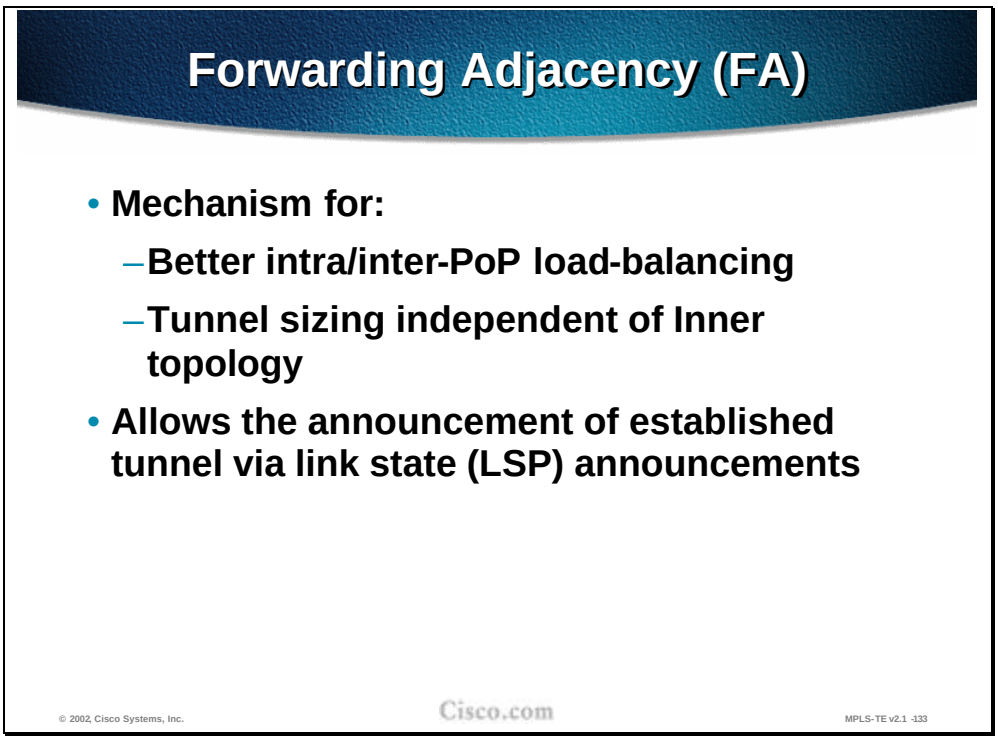
The secondary tunnel (t2) is used as a backup and provides for a fast transition from the primary tunnel in a case of failure. The drawback is that T2 reserved the bandwidth, which cannot be used by other tunnels.



The last example of the autoroute feature shows two configured tunnels: T1 (following the LSP path A-B-C-D-G) and T2 (following the path A-B-E-F-G). The relative tunnel metrics are unchanged and equal to the native IGP metric to the tunnel endpoints. In both cases, the tunnel metric is 40.

Since both tunnel metrics are equal to the native IGP metrics, the tunnels are preferred routing paths for all the destinations behind the tunnel endpoints (networks 2 and 4) and thus both tunnels appear in the routing table.

The load balancing across the parallel paths is done in proportion to the configured bandwidth on the tunnel.

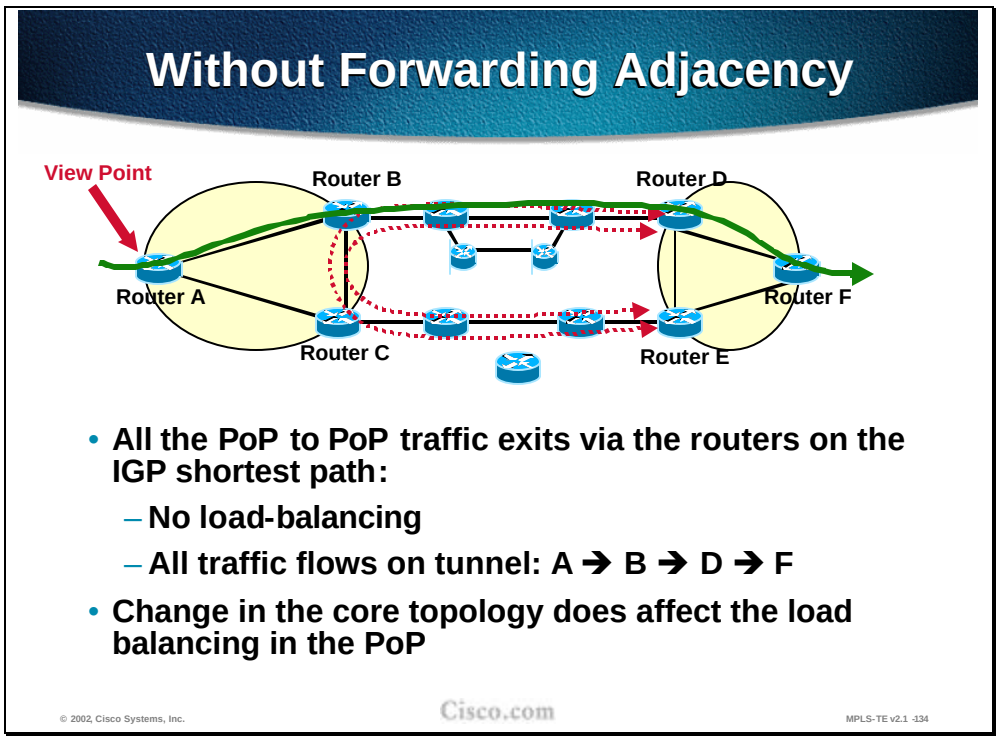


The MPLS TE Forwarding Adjacency feature allows a network administrator to handle a traffic engineering, label-switched path (LSP) tunnel as a link in an Interior Gateway Protocol (IGP) network based on the Shortest Path First (SPF) algorithm. A forwarding adjacency can be created between routers regardless of their location in the network

Forwarding Adjacency is a mechanism to allow the announcement of established tunnels via IGP to all nodes within an area.

By using forwarding-adjacency you can achieve

- A better load balancing when creating POP-to POP tunnels
- Use tunnels from any upstream node independent of the inner topology of the network
- Use tunnels independent of topology changes within the tunneled network area



Before looking into the real benefits of this feature we need to clearly see the limitations of the autoroute command in certain network topologies.

In this example we have established tunnels from B to D, from B to E, from C to E and from C to D preferring the tunnels B to D and C to E.

The path metric from B to D for ISIS is 30 (assuming default metric)

The path metric from C to E is 20.

But traffic is entering at router A. And router A has no knowledge about the existence of tunnels between B and D and C and E, so all he has is its IGP information, telling him, that the better path to F leads via router B and D.

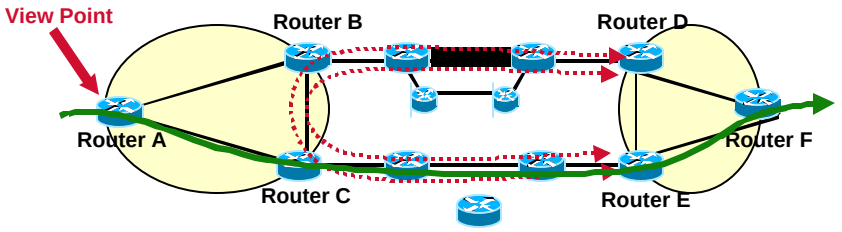
The results are

- There will be no load balancing
- All traffic will flow via B and D

Any change in the core topology will affect the path metric and thus any load balancing for POP to POP traffic.

Note You can theoretically also prevent this problem by creating tunnels from any router to any router but this design does not scale in medium and big networks

Without Forwarding Adjacency (Cont.)



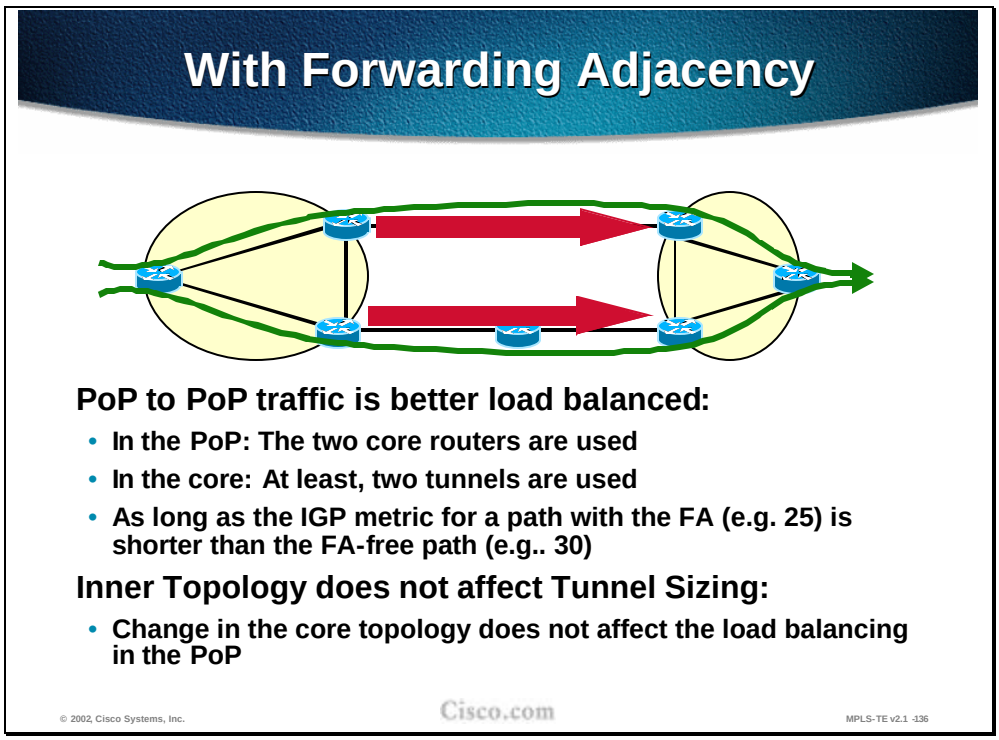
- All the PoP to PoP traffic exits via the routers on the IGP shortest path
- Change in the core topology does affect the load balancing in the PoP:
  - Normal state: All traffic flows A ➔ B ➔ D ➔ F
  - Link failure: All traffic flows A ➔ C ➔ E ➔ F

What happens in case a link on the path of our tunnel between B and D gets broken?

Even though we possibly have a rerouting in place the IGP metric for the complete path A – B – intermediates – D – F will change.

In our example the change in the metric will result in a possibly unplanned switchover of the traffic from A to F from the upper to the lower path.

This may result in a possible congestion on the path from C to E , whereas the protected path from B to D gets idled out.



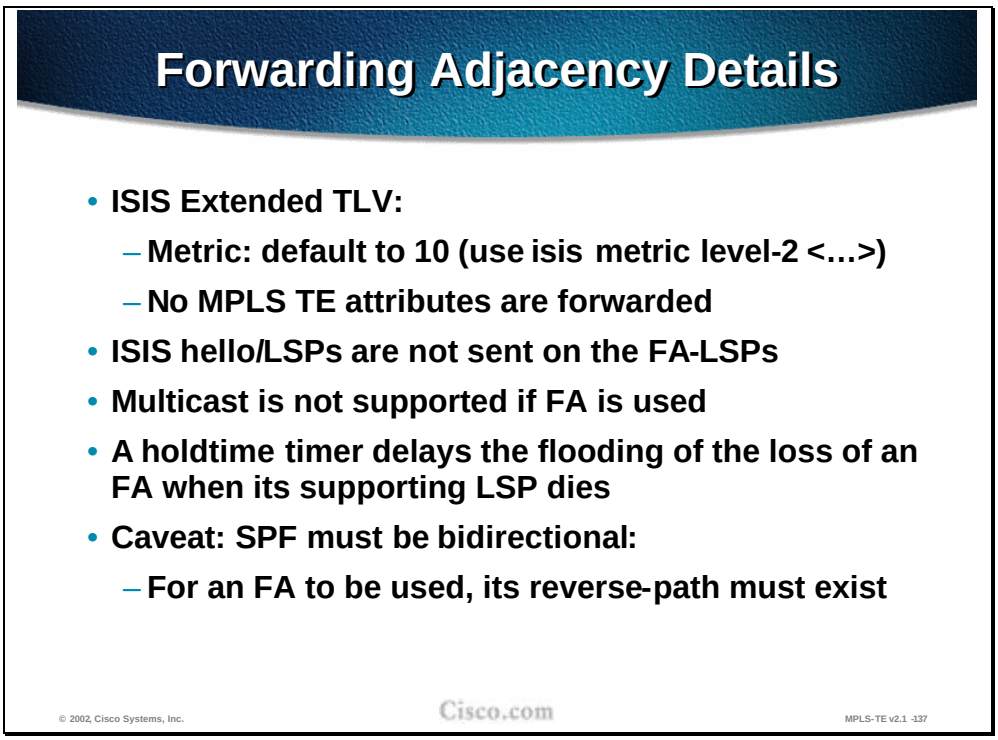
By using forwarding adjacency you can create POP-to POP tunnels where traffic paths and load balancing can be designed independent of the inner (core) topology of the network and independent of link failures.

To advertise a TE tunnel as a link in an IGP network, use the **tunnel mpls traffic-eng forwarding-adjacency** command in interface configuration mode.

tunnel mpls traffic-eng forwarding-adjacency {**holdtime** *value*}

Syntax Description

holdtime *value*	(Optional) Time in milliseconds (ms) that a TE tunnel waits after going down before informing the network. The range is 0 to 4,294,967,295 ms. The default value is 0**.**



Caveats when Using Forwarding-Adjacency

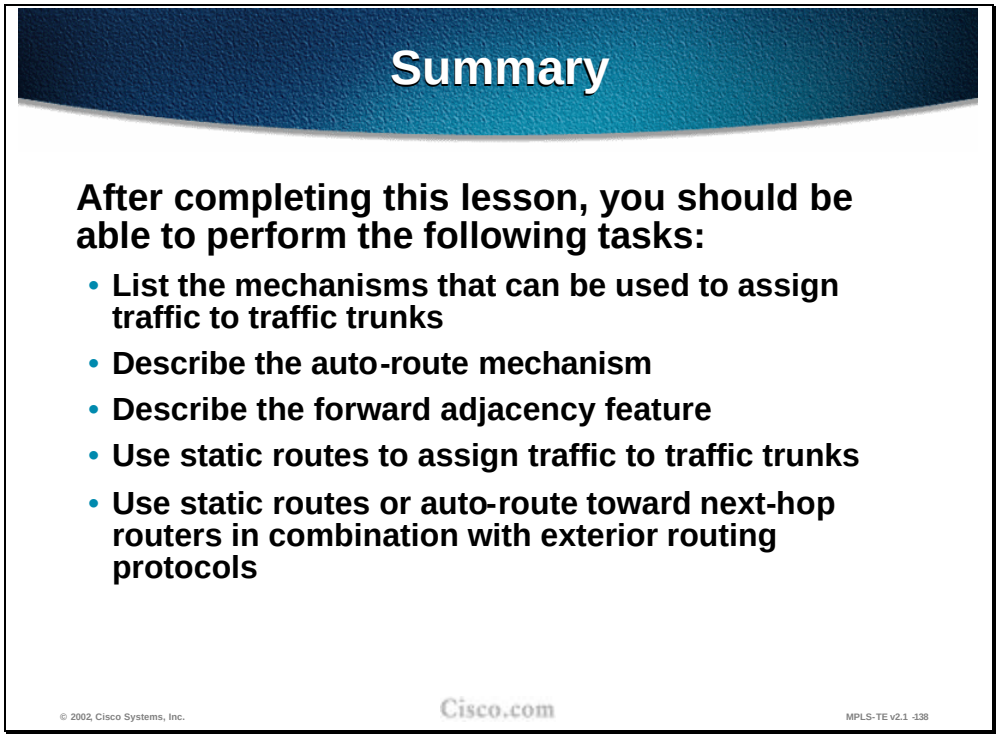
Caveats when using forwarding-adjacency are:

- Using the MPLS TE Forwarding Adjacency feature increases the size of the IGP database by advertising a TE tunnel as a link.
- The MPLS TE Forwarding Adjacency feature is supported by Intermediate System-to-Intermediate System (IS-IS). Open Shortest Path First (OSPF) support will be available in a future release.
- When the MPLS TE Forwarding Adjacency feature is enabled on a TE tunnel, the link is advertised in the IGP network as a Type Length Value (TLV) 22 without any TE sub-TLV.
- MPLS TE forwarding adjacency tunnels must be configured bidirectionally.
- No forwarding of MPLS-TE attributes for this link
- No support for Multicast as this is a unidirectional link (UDL) and the reverse path check (RPF) will fail.

Note You must configure a forwarding adjacency on two LSP tunnels bidirectionally, from A to B and B to A. Otherwise, the forwarding adjacency is advertised, but not used in the IGP network.

Summary

This section summarizes the key points discussed in this lesson.



Lesson Review

Instructions

Answer the following questions:

1. Explain the drawbacks of the static assignment of the traffic to MPLS-TE tunnels.
2. What are the benefits of using the autoroute feature in MPLS-TE?
3. Which path is preferred when using autoroute feature for the destinations behind the tunnel enpoints?
4. How is load-balancing done on two euqal-cost MPLS-TE tunnels?
5. What are the reasons for turning on forwarding-adjacency?

Summary

This section summarizes the key points discussed in this module.