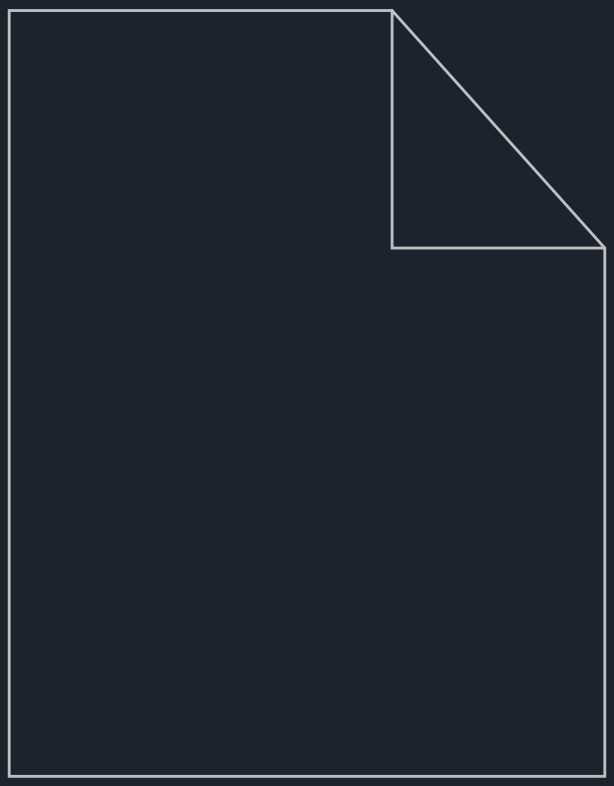
ALL ABOUT MPLS TRAFFIC ENGINEERING



A PACKET DESIGN E-BOOK

PACKET DESIGN™

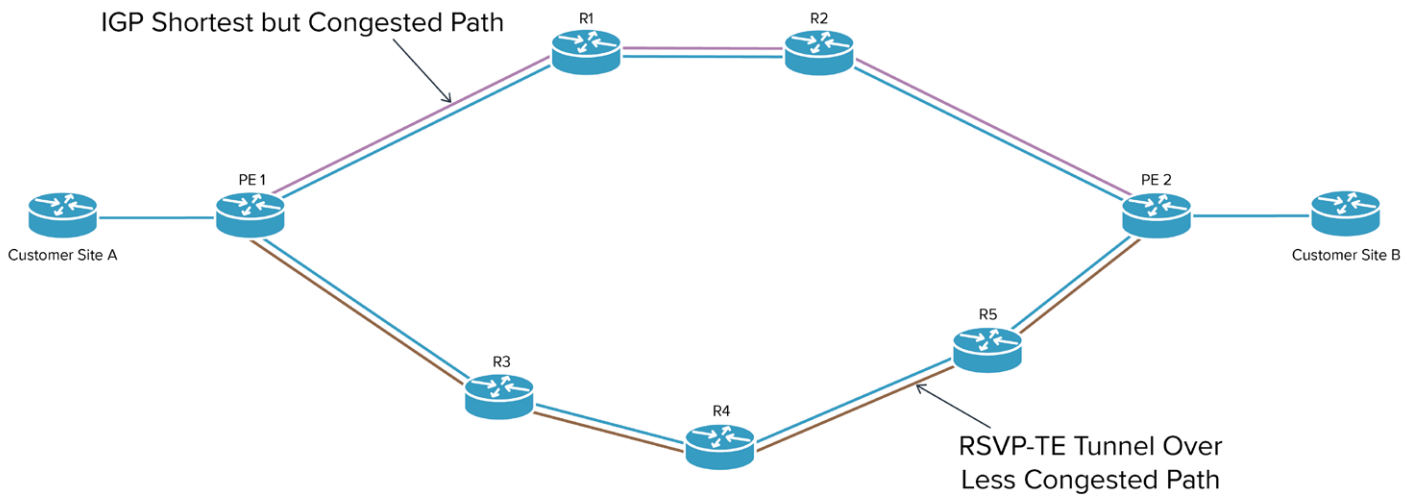
Index

What is MPLS Traffic Engineering?

Networks exist to deliver data packets between different endpoints. In a traditional IP-based network, data packets are forwarded on a per-hop basis. Each router between the source and destination performs a look-up of the packet's destination IP address in the routing table and selects the lowest cost path to forward the packets.

This method has a disadvantage: If one path is found to be optimal due to its low cost, every router in the network will prefer to use that path to forward packets even when there are several other routes available. This is because IGPs, the Interior Gateway Protocols, are designed to choose least-cost paths to forward packets. But when numerous forwarding routers prefer the same path, it can lead to over-utilization of the links along this path, resulting in congestion and packet drops. The preference for the shortest path holds true even when other under-utilized or idle paths are available. Thus, the per-hop routing approach to data transmission can affect data delivery and overall network performance in some networks.

With Traffic Engineering (TE), rather than per-hop routing decisions, the network operator's headend ingress router controls the path taken by traffic between a source and destination. Instead of sending traffic through a least-cost but congested path, TE directs designated traffic through under-utilized or idle paths, thereby distributing the bandwidth load in the network. TE tunnels also provide a mechanism for creating paths with certain quality of service (QoS) characteristics.



TE Tunnel includes more hops but is less congested.

For Traffic Engineering to work, the headend router that creates the TE tunnels must be aware of the latest network topology, traffic patterns, and the availability of resources, such as bandwidth on links. This information is collected with the help of IGPs, such as OSPF and IS-IS. In addition to traffic bandwidth requirements, TE tunnels may also factor Class of Service (CoS) requirements of the data to be forwarded. They then update this information to all the other routers within the IGP domain.

This data helps the headend ingress router in the provider network to compute multiple, distinct paths between the same source and destination edge routers to be used by the TE tunnels. Once the TE tunnels are created, packets are assigned MPLS labels and forwarded across the network based on the labels. Thus, this method of packet forwarding is referred to as MPLS Traffic Engineering.

Advantages of MPLS Traffic Engineering

The major advantages of MPLS Traffic Engineering are congestion avoidance and efficient use of existing network resources that can be idle or under-utilized. By driving traffic from over-utilized links to under-utilized ones, Traffic Engineering avoids network congestion caused by too much traffic on the least-cost link. This helps prevent packet drops and ensures data delivery.

Another advantage is the efficient use of resources, especially network bandwidth. With better utilization of available bandwidth, operators can save money and increase revenue. They can avoid spending capex on additional links with higher bandwidth to meet new traffic demands. The optimization of existing network resources also allows the network service provider to provide more services to customers, thus bringing in higher revenue.

Benefits of Traffic Engineering

- Minimizes the worst link utilization
  - Alleviates traffic congestion
  - Better/longer use of capital expenditures
- Routing traffic around congested links
  - Can use shortest as well as non-shortest paths
  - IGP metric tuning, RSVP-TE, segment routing

In addition to cost savings and helping with congestion avoidance, Traffic Engineering also helps with failover. When a primary path or tunnel between two endpoints in the network fails, Traffic Engineering drives traffic through idle links, thus providing Fast Reroute (FRR) on the TE tunnels.

Finally, Traffic Engineering also helps enable Constraint Based Routing (CBR), where the best possible path for traffic from a source to destination can be computed based on the resource availability and the demands of the traffic.

Evolution of Traffic Engineering

Network engineers have used Traffic Engineering for many years in many networks. Here's how Traffic Engineering has evolved over the years from an offline model, performed with the help of external network planning tools, to the on-device model in use today.

Offline Model to On-Device Traffic Engineering

The early implementations of Traffic Engineering used an offline model, which involved loading the current network topology and traffic demand matrix into a planning tool that calculated the best paths using optimization algorithms. But this model, though workable, could not handle the network changes efficiently enough due to several challenges.

The biggest issue came from having to add the network topology into the planning tool. This is easy in a network with only a handful of routers. However, the process of manually adding the topology of a network with several hundreds or even thousands of routers to a planning tool was a major challenge.

The difficulty became even greater when the network topology changed before the next traffic engineering adjustments. This meant that the next Traffic Engineering path computation required adding the updated topology and once again calculating the traffic demand matrix.

Another challenge was how long it took the optimization algorithms to calculate new paths, which typically ranged from a few hours to even a couple of days in very large networks. Consequently, when a link failure caused congestion, the network provider could not find alternate paths quickly enough for their customers resulting in dissatisfaction and possible attrition.

The drawbacks from the use of external planning tools led to the development of an on-device model of traffic engineering using Resource Reservation Protocol-Traffic Engineering (RSVP-TE) and the constraint based shortest path first (CSPF) algorithm. In this method, the routers, rather than external planning tools, managed the Traffic Engineering function. This method leverages the behavior of IGP routers, where they broadcast the available link bandwidth through the network. TE tunnels are then set up between source and destination routers, and the utilization on these tunnels is monitored to create a traffic demand matrix.

When increased utilization causes congestion on a tunnel, the headend router of the tunnel runs its CSPF algorithm to re-optimize the path. It then uses RSVP-TE to signal the new path to the other routers along the path. The CSPF algorithm used here is extremely fast, and because this model works from the routers themselves, the network topology is readily available. This allows the network to respond in real time and overcome the limitations of offline traffic engineering. This is especially useful in alleviating congestion caused by link failures.

Related Reading:

Traffic Engineering Evolves Blog Series: Offline to On-Device

Traffic Engineering Mechanisms

Now that we have covered what Traffic Engineering is and the basics of how it works, let us look at the different TE mechanisms.

The most widely used mechanism is RSVP-TE which we referred to in the Evolution of Traffic Engineering section. The other is Segment Routing, an emerging technology that is beginning to be adopted in many service provider networks.

Resource Reservation Protocol – Traffic Engineering (RSVP-TE)

Resource Reservation Protocol (RSVP) is a protocol that is used to reserve resources along the end-to-end path of a traffic flow in an IP network. RSVP messages are sent by the headend router in a network to identify resource availability. An RSVP request consists of a FlowSpec that specifies the Quality of Service (QoS) requirement for the traffic flow and a FilterSpec that defines which flow must receive the QoS priority.

Once the necessary bandwidth is reserved along the path with RSVP, the application that made the request begins to transmit the traffic. RSVP is primarily used by real-time and multimedia applications to set up bandwidth reservations. RSVP thus communicates the requirements of specific traffic flows to the network. The RSVP signaling protocol was extended with MPLS features to support MPLS TE. This enabled RSVP to set up label switched paths (LSP) in an MPLS TE network.

Before we begin, these are some of the terms commonly used in MPLS and Traffic Engineering terminology:

Label Edge Router (LER) – An LER is a router that operates at the edge of an MPLS network and is the entry or exit point in the MPLS network.

Label Switching Router (LSR) – LSRs are routers that lie along a label switched path established by LERs between a source and destination pair. The function of an LSR is to perform MPLS label switching.

Label Switched Path (LSP) – An LSP is a path established between two routers (LERs) to route traffic in an MPLS network. An LSP is established over a sequence of LSRs. It is only after an LSP has been established that MPLS forwarding can occur.

Label Distribution Protocol (LDP) – An LDP is used to establish a label switched path from a source to a destination.

Headend Router – The upstream, transmit end of a tunnel – the router originates and maintains the traffic engineering LSP.

Tailend Router – The downstream, receive end of a tunnel – the router terminates the traffic engineering LSP originating from the headend router.

RSVP-TE primarily uses four messages to reserve a path for traffic. They are:

RSVP PATH message – This message is used to check the availability of resources along the TE path and acts as a reservation request. The ingress label switched router (LSR), which is the headend router, generates a PATH message. It traverses downstream to the tailend router through the future path of the TE tunnel and checks for the availability of the resources.

RSVP RESERVATION message – When the tailend router in the domain receives the PATH message, it generates an RSVP RESERVATION message to confirm the availability of resources requested.

RSVP Error messages – If a resource requested by the PATH message is not available, then the router that is unable to reserve the resource generates an RSVP Error message and sends it to the router from which it received the request or reply. There are two types of error messages.

1. PATH ERROR message: This error is triggered when a router receives a PATH message generated by the headend router but is unable to find the resources requested of it. PATH ERROR messages are sent to the upstream router from which the PATH message was received.

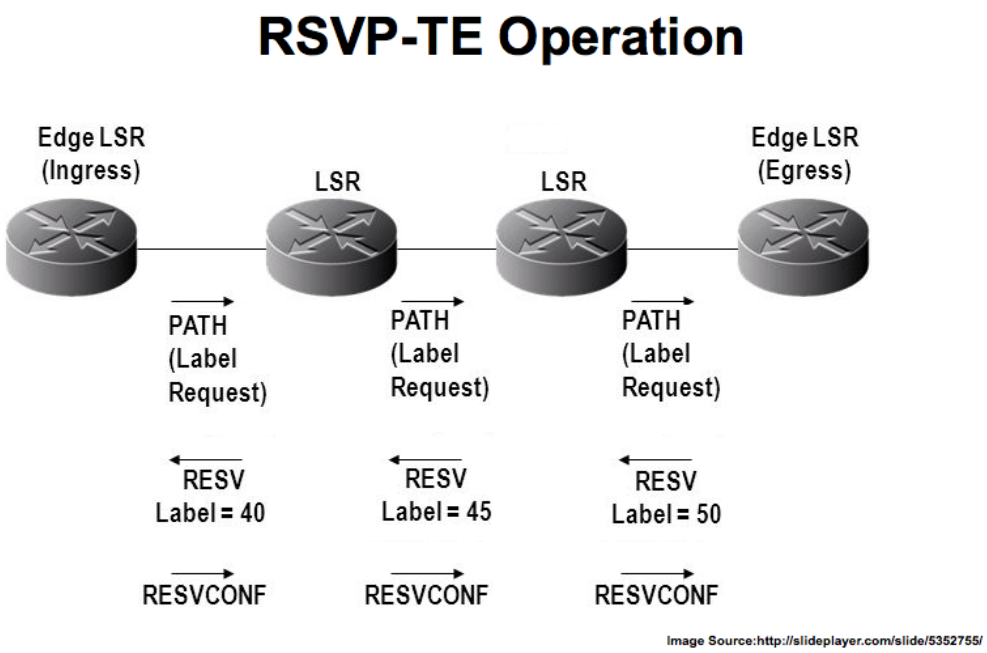
2. RESERVATION ERROR message: This error is generated if a router fails to find resources after an RSVP PATH message has reached the tailend router but before the RSVP RESERVATION message from the tailend router reaches it. RESERVATION ERROR (RESVERR) messages are sent to the downstream router.

RSVP Tear messages – This message is used to tear down a reservation made and release the resources so other TE tunnels can use them.

RSVP-TE at Work

To reserve a path via RSVP-TE, the headend router sends an RSVP PATH message that checks the availability of requested resources on all the LSRs along the path on which the TE tunnel is to be created. Upon receiving the PATH message, the tailend router in the path then confirms the reservation with an RSVP RESERVATION message, which confirms the assignment of an LSP to a TE tunnel. This message is then propagated upstream to the headend router through all the LSRs along the future TE tunnel path.

After all the LSRs in the path accept and confirm the LSP, the MPLS TE LSP is operational. With this, the headend router can then direct traffic through new tunnels based on the resource requirements of the traffic being transmitted. The traffic engineered MPLS network using RVSP-TE is ready.



Related Reading:

Configuring RSVP-TE on Cisco ASR 9000

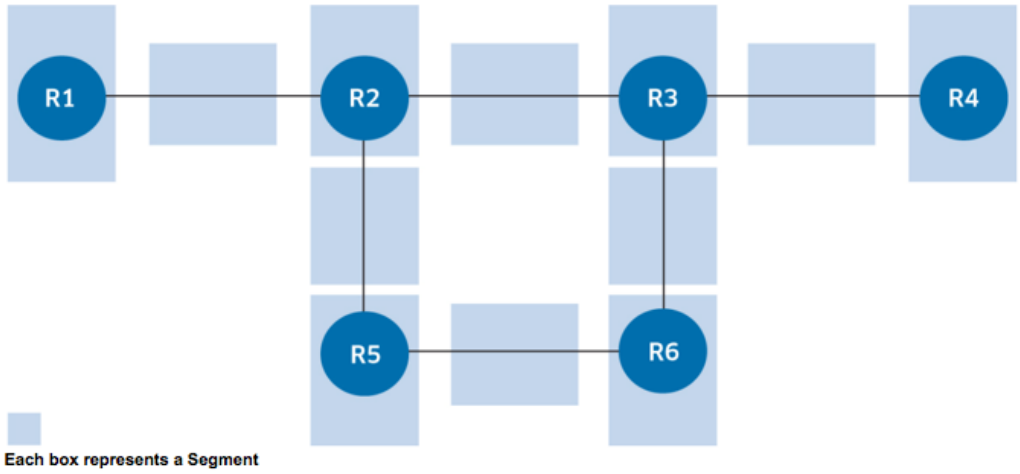
Segment Routing

Segment Routing (SR) is not a new technology, but all the major network equipment vendors have recently embraced it. Here's why. SR is a packet forwarding technology where the source node defines the path for traffic, which is then sent through specific nodes and forwarding paths called segments. An SR path is not dependent on hop-by-hop signaling, LDP, or RSVP. Instead, it uses segments for forwarding.

Segment Routing, like MPLS, uses label switching to forward packets, but in SR terminology the labels are called segments. There are two types of segments. A node segment identifies the shortest path to a destination node (e.g. R1 to R3 in the below image). The node segment is advertised throughout the network and all remote nodes install the segment in their respective MPLS data planes. The other type is an adjacency segment, which represents a link between two adjacent nodes (e.g. R2 to R5).

The network operator gives each node in an SR domain a globally unique identifier, known as the segment identifier (SID). The SR node then allocates a local SID for each of its adjacency segments which are stored only by that specific node in its data plane.

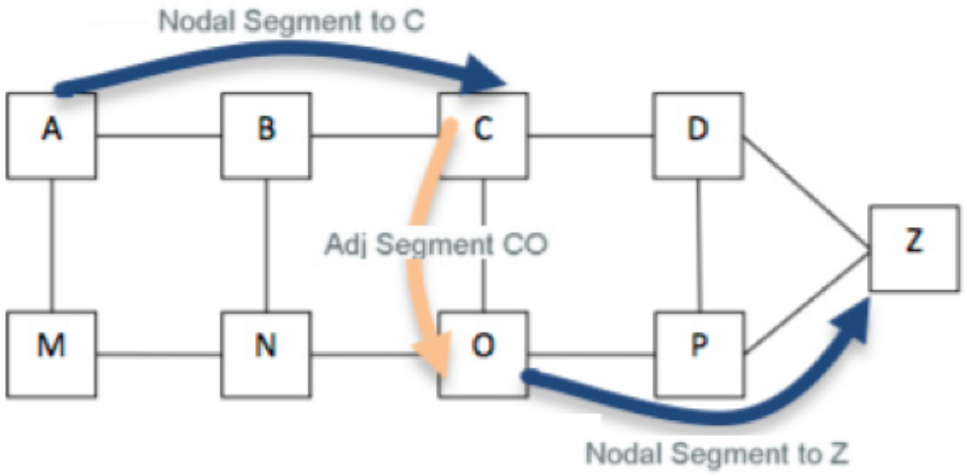
A Segment Routing Domain Represented in Segments



Source: https://insight.nokia.com

When a Segment Routing tunnel is created it either contains a single segment that represents a path to a destination, or a segment list, which is a set of segments that the tunnel will encode to reach its destination. In the SR domain in the image above, traffic from node A to Z traverses a node segment that carries traffic to C, an adjacency segment CO that represents the link from node C to O, and another node segment that carries traffic from O to Z.

Source Routing using Segments



Source: Cisco Segment Routing Intro

By using segments, SR can ignore the shortest or lowest cost path preference of IGP, which may be congested, and encode any path in the MPLS network to route traffic to its destination. This gives the operator the ability to steer traffic over different paths based on its requirements and the state of the network. For example, an operator can send voice traffic over a low latency path and bulk data over a high latency path using Segment Routing.

Segment Routing has a few advantages over RSVP-TE. One is that it provides more granular control to the network operator over routes and traffic. Because SR supports Class of Service-based TE (CoS), a provider can define per-flow CoS policies and encode a segment to fulfill the CoS demands. RSVP-TE has failed to provide this level of granular control due to scalability issues.

The second advantage is that Segment Routing removes the needs for protocols such as LDP or RSVP and uses fewer labels compared to an LDP or RSVP-TE deployment. This reduces the overhead on the network. SR also does not use signaling, allowing it to scale significantly better than RSVP-TE while simplifying the network.

A third reason is its automated and native FRR capabilities that greatly reduce convergence time to sub 50 milliseconds in any topology. Segment Routing can also work with Path Computation Element (PCE) to enable an agile WAN-SDN (a.k.a., Carrier SDN). SR with an SDN application can be used to provision TE tunnels automatically and provide value-added services, such as bandwidth management, bandwidth calendaring, and bandwidth on-demand.

These Segment Routing capabilities enable seamless MPLS networks while providing all the TE functions that RSVP-TE does.

Related Reading:

Configuring Segment Routing on Cisco IOS XE

Enabling Segment Routing capabilities with OpenDaylight, Cisco IOS-XR 6.0.0, and the Packet Design SDN Platform

VIDEO – Segment Routing: A Control Plane Simplification

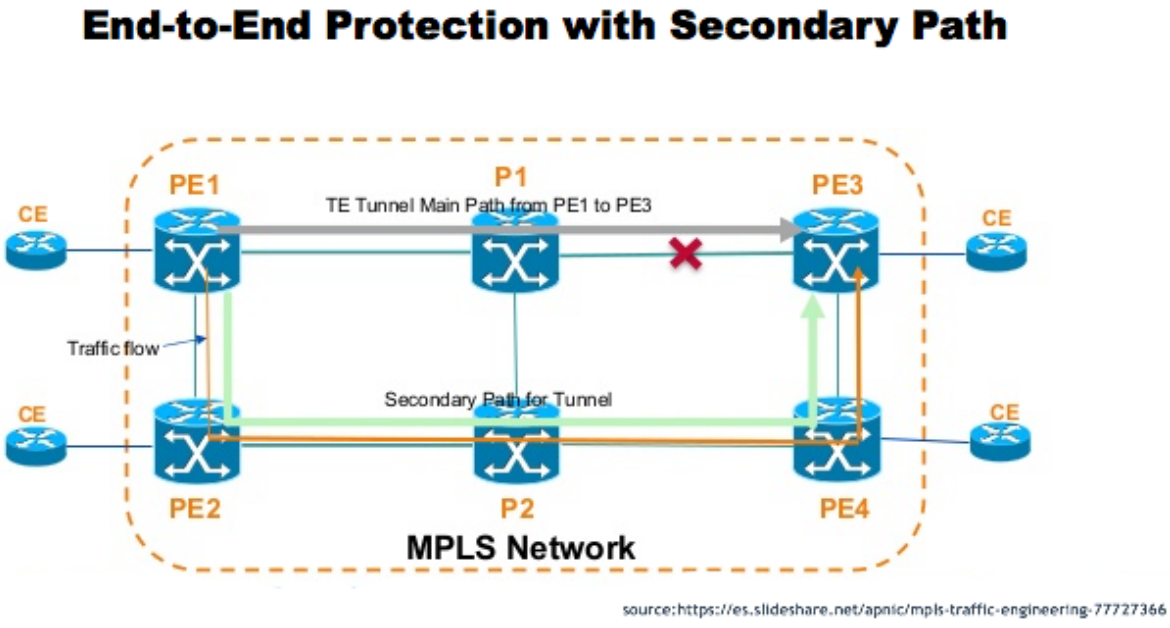
MPLS-TE Tunnel Protection Mechanisms

A failure can hit any network, including traffic engineered ones, resulting in disruption of services offered. Resiliency mechanisms are a must, even in TE networks. There are two mechanisms that can be used in MPLS-TE networks for quick recovery from failure: End-to-End Protection (path protection enabled by using a secondary path), and local protection enabled by MPLS Fast Reroute (FRR).

End-to-End Protection

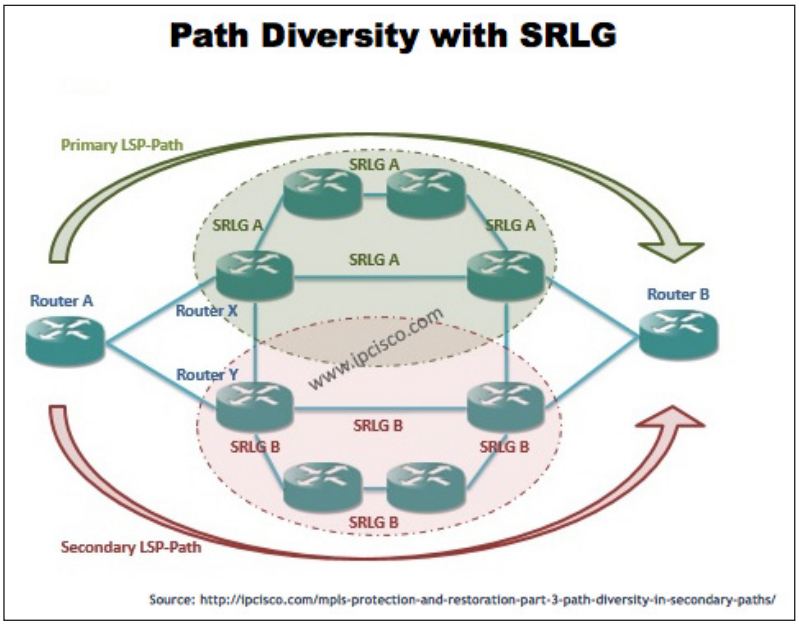
End-to-End Protection, provides failure recovery for the entire LSP that is carrying a TE tunnel's traffic. This is achieved using two LSPs – the primary LSP, which is the active LSP carrying the TE traffic, and a secondary path that acts as a standby path, ready to take over when the primary path fails.

In this mechanism, the secondary LSP is configured and established in advance. When a primary LSP fails, the headend router along the tunnel is alerted of the path failure using failure detection mechanisms that leverage RSVP signaling or IGP. The headend router then immediately switches the MPLS-TE tunnel's traffic from the primary to the secondary LSP. Once the primary LSP recovers, traffic is switched back to it.



When enabling secondary paths for path protection, it is also necessary to ensure that both the primary and secondary LSPs use different paths. This provides resiliency by ensuring that both LSPs have no single point of failure. This method is referred to as path diversity and can be achieved using full strict hop LSP paths, Shared Risk Link Group (SRLG), or Admin groups.

With a full strict hop LSP, the exact path to be taken by the LSP is identified and no label switched routers (LSR) are allowed to overlap between the two LSPs. Here, the exact order of the LSRs through which the RSVP messages are sent is specified. This way, the primary and the secondary LSPs take two unique paths, and failure of an LSR along one path does not affect the other path. This method also brings with it more configuration overhead and adds to the complexity, especially in large networks.



The other mechanism for path diversity is SRLG. An SRLG is a group of links that share a common physical attribute (a common fiber) and are considered to carry the same risk. If one link in a group fails, other links in the group may fail too, thus putting them all in a group with shared risks and thus, Shared Risk Link Group or SRLG. MPLS path diversity is achieved over SRLGs by ensuring that the primary and secondary LSPs do not use links from the same SRLG. This ensures that when a primary LSP fails, the secondary LSP also does not fail.

Admin groups work similarly to SRLG. Links are assigned to different Admin groups, and the primary and secondary LSPs are configured to not use links in the same Admin groups.

MPLS FRR (Fast Reroute)

Having a secondary LSP pre-established to act as the backup path is much faster than having the headend router dynamically compute new LSPs when the need arises. But there is a faster mechanism for MPLS-TE protection. This mechanism, known as MPLS FRR, protects MPLS TE tunnels from link and node failures and is also referred to as local protection.

With Fast Reroute, MPLS-TE LSPs are protected from link or node failures by bypassing the local point of failure until the headend router establishes a new end-to-end LSP. It is because the protection happens close to the point of failure rather than for the entire end-to-end path that FRR is referred to as local protection.

The advantage of FRR is that it provides recovery in less than 50 milliseconds during a failure with minimal packet loss. It also does not incur the overhead of end-to-end protection when an entire backup LSP must be created.

MPLS FRR is classified into two categories – link protection and node protection. But first, there are two router roles to know:

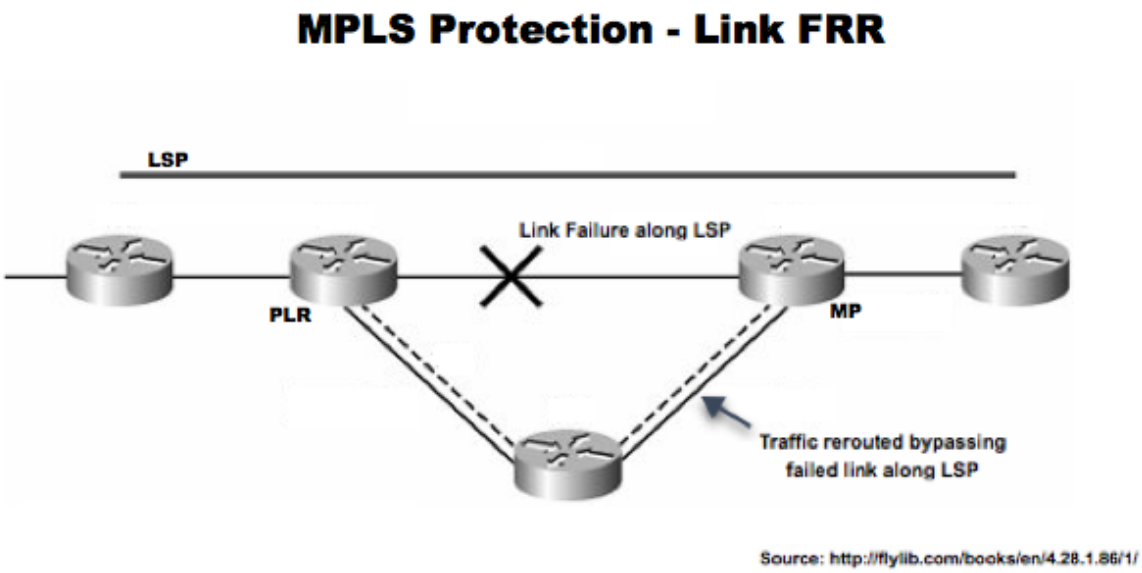
Point of Local Repair (PLR): The router where the backup tunnel originates after the failure of the downstream link or node. This router forwards the traffic along the alternate path and notifies the headend router that the primary LSP has an issue.

Merge Point (MP): The point where the alternate (backup) path terminates and merges into the original LSP.

Now here are the two types of FRR mechanisms.

Link Protection:

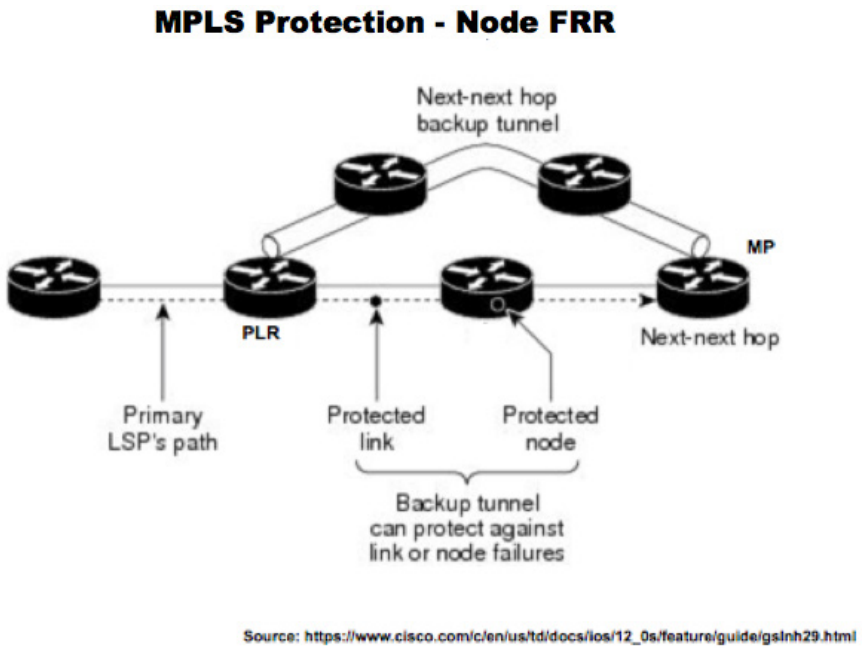
In this mechanism, when a link along an LSP fails, traffic is rerouted to the next hop through backup tunnels that bypass only the failed link in the LSP. These backup tunnels created are referred to as next-hop (NHOP) backup tunnels because they terminate at the next hop after the point of failure.



Here, when a link fails, the PLR swaps the MPLS label and pushes the backup label. This reroutes the traffic along the backup path until the backup terminates at the MP and traffic rejoins the primary LSP. The PLR also sends path error messages to the headend router to notify of the LSP failure.

Node Protection:

This mechanism protects the TE tunnel when the next downstream router fails. Here, backup tunnels that bypass the next-hop node are created to carry traffic. The backup tunnels terminate at a node after the next-hop node (two hops away) and so are called next-next-hop (NNHOP) backup tunnels.



Unlike link protection that provides resiliency only in case of link failure, node protection can provide protection against both link and node failures.

Link and node protection FRR have two further categorizations. The first is One-to-One protection where each LSP is protected with a separate backup LSP. This One-to-One protection tunnel is referred to as a detour.

The second is Many-to-One or facility backup (1:N) where a single backup tunnel is used for multiple LSPs. This mechanism is referred to as a bypass and allows many LSPs to be bound by a single bypass tunnel.

These mechanisms help ensure resiliency on MPLS-TE tunnels and ensure data delivery is not affected by downtime.

Related Reading:

Configuring MPLS TE Path Protection on Cisco IOS

Juniper MPLS FRR Overview

MPLS Protection and Restoration

SDN and Traffic Engineering

Software Defined Networking (SDN) has made its way into many areas of networking. Two areas in traffic engineering where SDN plays a role are addressing some of the challenges that come with traffic engineering and optimizing traffic engineering.

How SDN Addresses Traffic Engineering Challenges

In a TE network, a link failure can trigger race conditions when multiple headend routers try to compute new TE tunnels without considering the bandwidth requirements of other headend routers. This leads to some of the tunnels failing to find a path. In SDN-enabled TE networks, a centralized SDN controller takes care of computing the path and allocating the bandwidth while being aware of the entire network's bandwidth requirements. This approach enables the network to avoid triggering race conditions.

RSVP-TE creates another challenge by requiring tunnels to be formed between each router in the network to get the complete traffic matrix. This leads to what is known as the "n-squared" or "full-mesh" problem, where too many tunnels are created. For example, even a small service provider with 75 routers and 300 links can have 1,600 tunnels when creating a full mesh network. This number goes up in larger networks as the number of routers increases, making it extremely difficult for the network engineer to configure and manage the tunnels.

Because the SDN controller has a global network view, it allows for network-wide resource optimization. This overcomes the n-squared problem by creating tunnels only if needed, and only if they will have a positive impact.

The n-squared problem also creates another issue. IGP is used by TE mechanisms to collect bandwidth resource information. But the large number of tunnels results in IGP propagating more about the available bandwidth than its primary task of propagating the link up and down status. SDN takes the strain off IGP by instead leveraging NETCONF-enabled push-based telemetry, accessed via YANG models, and flow technologies such as NetFlow to get the traffic matrices.

Optimizing MPLS-TE Networks with SDN

The other area where SDN helps is the automation and optimization of a traffic-engineered network. Here are a few uses cases on how SDN can help with automation and optimization in MPLS-TE networks.

SDN with MPLS-TE allows network operators to make better use of network resources and run their networks hotter. Using SDN with RSVP-TE or Segment Routing, an operator can automate the balancing of network loads and shift traffic between different links. This results in better overall usage of network resources. It also allows them to run existing links at a higher capacity and shift traffic automatically to other links when there is congestion.

SDN can also help handle temporary spikes in network connection requests – e.g., a big game telecast. The usual process for such a request can take hours or weeks due to the level of planning involved. Using SDN, the operator can automate tunnel creation with the required bandwidth for the duration of the event and tear down after the event. This task cannot be performed by routers that respond only to current demands and require manual intervention.

Automation and optimization can even be extended to other scenarios such as rerouting traffic over a more suitable path to guarantee SLAs, providing additional bandwidth during a specific time for a customer, and many more.

Related Reading:

Packet Design SDN Traffic Engineering

SDN Traffic Engineering, A Natural Evolution

VIDEO – How SDN addresses Traffic Engineering Challenges

Automating Network Optimization

Packet Design Explorer Suite for Traffic Engineering:

Another important factor to consider in traffic-engineered networks is having the ability to monitor, automate, and optimize the performance of TE tunnels. This requires real-time network telemetry from the devices in IP/MPLS networks as well as from SDN controllers.

Packet Design Explorer Suite delivers visibility into the performance of both RSVP-TE and Segment Routing tunnels in traffic engineered networks. Packet Design's industry-leading route analytics technology provides a comprehensive traffic engineering management solution, including real-time TE tunnel monitoring, historical analysis, interactive network modeling, and capacity planning reports.

Network engineers can generate real-time and historical reports about TE tunnels, including their underlying routing topology, information about tunnel path changes and bandwidth utilization, detection of unprotected links, reports on SR metrics, and many other features. This data helps network teams ensure that the TE tunnels are performing as expected and aids in proactive troubleshooting when issues pop up. Network engineers can also use the Explorer Suite to model traffic engineering changes, such as the addition of new TE tunnels and failure scenarios, to visualize their impact on the network.

The Explorer Suite also functions as an analytics and automation SDN platform, enabling users to automatically define and optimize MPLS-TE tunnels. Using the platform, a turn-key SDN application assesses the current network state from its traffic matrices to determine if optimization is needed or not. Based on this, the application makes recommendations for new TE tunnels to be created and presents the impact that the tunnels will have. This allows the user to review the recommendations and, if satisfied, provision the tunnels in the network with a mouse click.

Further Reading:

MPLS Traffic Engineering is a complex subject and we hope this E-Book has helped explain some of its concepts and provided more insight into its workings, challenges, and benefits. For more reading on the topic, please see the following.

MPLS Technology Overview

Cisco Press MPLS Traffic Engineering

MPLS Traffic Engineering Technology

To learn more about Packet Design and the Explorer Suite, please visit www.packetdesign.com