

# Traffic Engineering With Multiprotocol Label Switching

## Executive Summary

The strength of the Internet has been its immense scalability and adaptability to accommodate a seemingly unceasing portfolio of applications. In this tradition of adaptability, Multiprotocol Label Switching (MPLS) is the latest technique being implemented in the Internet core. MPLS provides a virtual path capability between packet (label) switches to efficiently carry differentiated services across the Internet. Additionally, MPLS has been enhanced with the capability to precisely engineer traffic tunnels to avoid congestion and more fully utilize all available bandwidth. The greatest strength of MPLS is its seamless coexistence with IP traffic and its reuse of proven IP routing protocols.

The white paper provides a problem definition providing the motivation for MPLS and traffic engineering. The core value of MPLS is then described followed by a discussion of three MPLS applications: shortcut routing, tunnel restoration, and integration of MPLS with Quality of Service. The last section provides a detailed analysis of MPLS signaling and the application of traffic engineering. The use of ATM for traffic engineering is also discussed and contrasted with the MPLS approach.

Avici™ Systems Inc.  
101 Billerica Ave  
North Billerica, MA 01862  
978-964-2000  
[www.avici.com](http://www.avici.com)



# Table of Contents

<b>EXECUTIVE SUMMARY .....</b>	<b>1</b>
<b>TRAFFIC ENGINEERING .....</b>	<b>3</b>
DEFINING THE PROBLEM .....	3
UTILIZING ALL AVAILABLE BANDWIDTH .....	3
<b>MULTI-PROTOCOL LABEL SWITCHING.....</b>	<b>4</b>
CORE VALUE OF MPLS .....	4
<b>MPLS APPLICATIONS .....</b>	<b>6</b>
SHORT CUT ROUTING .....	7
<i>BGP NEXT HOP</i> .....	7
<i>A MPLS Short Cut to BGP NEXT HOP</i> .....	7
TUNNEL RESTORATION .....	8
<i>Planned Head End Reroute Capability</i> .....	9
<i>Fast Reroute</i> .....	10
<i>Avici Failure Affinity Groups</i> .....	11
INTEGRATING MPLS AND QOS.....	11
<i>Integrated Services</i> .....	12
<i>Differentiated Services</i> .....	12
<i>IntServ Meets DiffServ and MPLS at the Internet Core</i> .....	13
<b>TRAFFIC ENGINEERING WITH MPLS AND ATM .....</b>	<b>16</b>
WHAT IS TRAFFIC ENGINEERING WITH MPLS? .....	16
<i>Traffic Engineering Policy</i> .....	16
<i>Enhancements to IGP and the Role of the TE-LSDB</i> .....	20
<i>Signaling for Label Distribution</i> .....	22
CURRENT SOLUTIONS TO TRAFFIC ENGINEERING .....	25
<i>Traffic Engineering with ATM</i> .....	25
<i>Drawbacks to ATM for Traffic Engineering</i> .....	26
<b>CONCLUSIONS.....</b>	<b>27</b>
<b>REFERENCES.....</b>	<b>28</b>

## Table of Figures

Figure 1 BGP Topology .....	4
Figure 2 MPLS Shim Headers .....	5
Figure 3 Short Cut Tunnel .....	8
Figure 4 Head end Reroute with Shared Explicit.....	9
Figure 5 Bypass Tunnel Concept .....	10
Figure 6 DiffServ Field in IPv4 Header .....	12
Figure 7 Administratively Specified Explicit Paths .....	17
Figure 8 Resource Affinity with Color Groups .....	17
Figure 9 Multi-LSP Load Balancing .....	19
Figure 10 Role of TE-LSDB.....	21
Figure 11 RSVP-TE PATH Messages.....	23
Figure 12 RSVP-TE RESV Message Flow .....	24
Figure 13 Example Label Assignment.....	25
Figure 14 Forwarding Labeled Packets.....	25
Figure 15 Layer 3 Topology with ATM.....	26
Figure 16 IP Layer 3 Overlay.....	27

## Traffic Engineering

### *Defining the Problem*

The Internet core is constantly expanding to meet the growing demands for bandwidth. However the growth in demand for bandwidth can at times overtake core providers' ability to add infrastructure. The demand for bandwidth in the Internet core is doubling roughly every four months to five months. The core provider has several options to meet this demand.

- Increase the number of circuits
- Increase the bandwidth of existing circuits
- Increase the capacity of the core routers
- Add more core routers
- Wait until QoS becomes ubiquitous and leverage it for admission control
- Better utilize all available bandwidth in the core

In fact, most core providers are doing all of the above. New fiber is being lit up with multiple wavelengths to support hundreds of gigabytes of bandwidth. New terabit routers, like the Avici TSR™ system are being delivered into the market to keep pace with the deployment high speed, high-density optical switching. Better utilizing all Internet core bandwidth and implementing QoS are discussed in this paper.

### *Utilizing All Available Bandwidth*

The network topologies at the Internet core are designed to provide substantial path diversity between backbone nodes in the provider's points of presence (POPs). Within the POPs backbone, nodes are connected to public and/or private peering (with other ISPs), provider data centers (for high volume web hosting), and large (sophisticated) corporate customers. Determination of routing between ISPs at peering points is controlled by external BGP (E-BGP) route admission policy. The BGP peering between service providers is called external and the BGP peering within a provider's network is called internal. Within the provider's core, internal BGP has rigid requirements for exchanging routing information amongst its peers. Internal BGP routers must peer with each other in a complete mesh<sup>1</sup> (the details are beyond the scope of this paper).

Figure 1 depicts the relationship between external and internal BGP. The dotted lines represent the shortest path used for both the internal peering relationships (over TCP) and data path for traffic between the core provider's customers. Within the core provider's network, the shortest paths are determined with either OSPF or IS-IS depending on the carrier.

---

<sup>1</sup> The restriction for a complete mesh has been obviated by the use of route reflectors and BGP Confederations (RFC 1966 BGP Route Reflection, Bates & Chandra, June 1996)

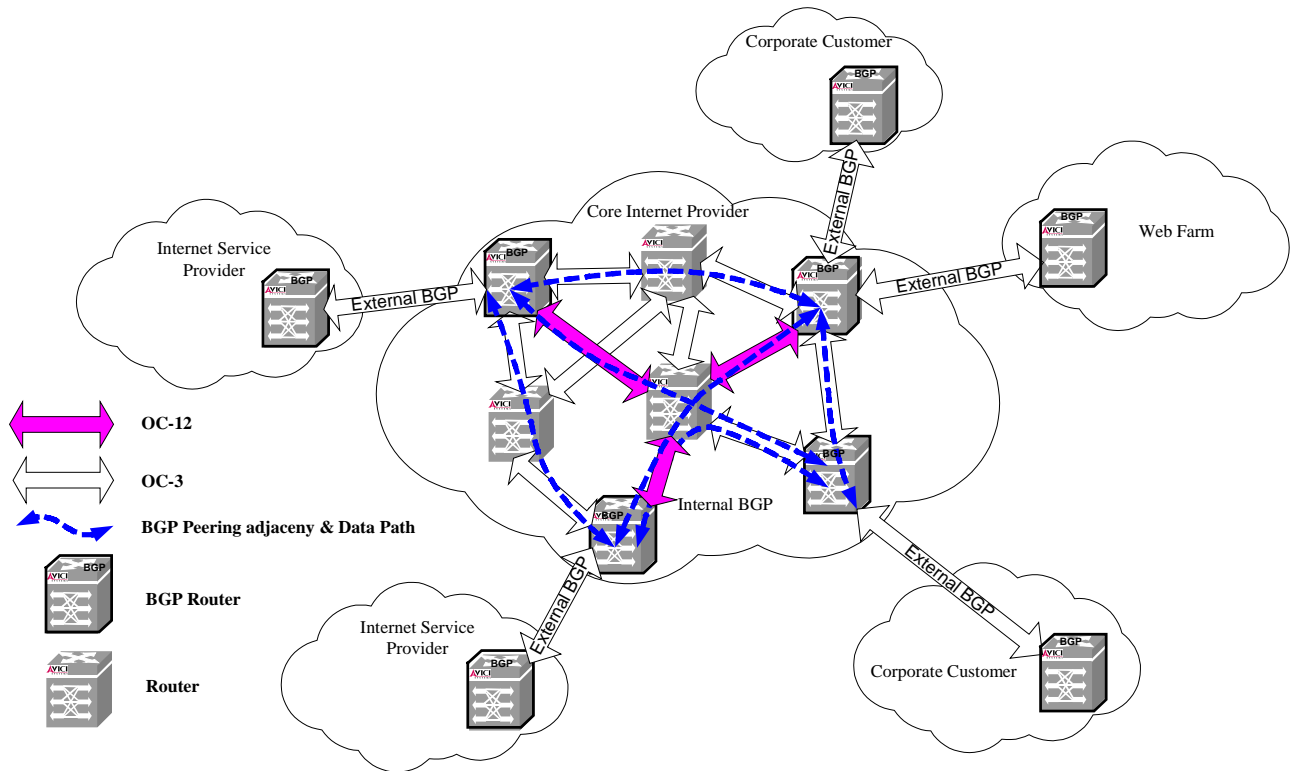


Figure 1 BGP Topology

Figure 1 also depicts two classes of bandwidth in the core provider's network, OC-3 and OC-12. The network engineer has the latitude to provide each link with a cost, which is (roughly), inversely proportional to the link's bandwidth. When the routers calculate shortest paths to all destinations the costs are considered in algorithm. The shortest path to a destination is biased towards links with the lowest total cost (summed in the outgoing direction from root to destination).

Unfortunately the cost metric associated with OSPF and IS-IS is too simplistic a modality for "traffic engineering." Using shortest path with costs alone can lead to significant imbalances in path loading. In other words, shortest path with costs is insufficient to utilize all available bandwidth in the core network.

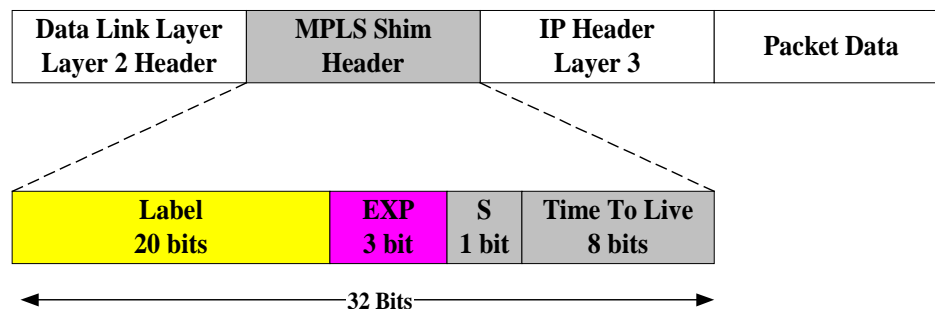
## Multi-Protocol Label Switching

### Core Value of MPLS

ATM is used by many Internet core provider as a means to traffic engineer paths between ATM attached IP routers. The section titled "What is Traffic Engineering with MPLS?" on page 16 provides a detailed assessment of ATM applied to traffic engineering. ATM is not an optimal solution for the Internet core due to line rate limitations, excessive overhead, and high administrative cost of ownership. Therefore new solutions are necessary to provide ATM-like functionality but with full Layer 3

participation and a single set of signaling protocols. The latest approach to traffic engineering comes as an enhancement to what was originally termed as tag switching. Tag (or label) switching has its origins in proprietary vendor implementations. The fundamental idea of tag switching is to leverage layer 3 interior routing protocols (OSPF and IS-IS) to calculate shortest paths to all possible destinations, but then assign a sequence of labels/tags along each path. As IP packets enter the label switched network, they are encapsulated in a labeled envelope. The label, and not the IP destination address, determines the next hop. When a labeled packet arrives at a label switching router (LSR) the incoming label identifies the packet's trajectory through the network. The incoming label is removed and replaced with the appropriate outgoing label. When the packet arrives at the periphery of the label switched network, the IP packet is de-encapsulated and routed normally. The goal of early (and proprietary) label switching was to simplify the determination of the next hop. In the absence of label encapsulation the determination of the next hop requires an expensive longest prefix match in a voluminous routing table. A longest prefix lookup against a table with "FULL" Internet routing (currently at 62,000 networks) is very taxing on the previous generation of Internet core routers. Ironically the latest generation of Internet core routers (including the Avici TSR system) can perform a longest prefix match against tables 4 to 5 times larger than the existing Internet routing table with no performance penalty.

The IETF draft for label switching is based on a 32-bit shim header that goes between the data link layer and the IP layer. The shim header is shown in Figure 2



- The **Label Field** is self explanatory
- The **EXP Field** is "Experimental" though it is proposed use is to indicate Per Hop Behavior of labeled packets traversing Label Switching Routers
- The **Stack (S) Field** indicates the presence of a label stack
- The **Time to Live Field** is decremented at each LSR hop and is used to throw away looping packets

Figure 2 MPLS Shim Headers

By itself, label switching does not satisfy the need for traffic engineering. The IETF has been working on standardizing an enhanced Label switching framework that includes a traffic engineering policy component. Underlying the MPLS framework is the ability to

create a tunnel or label switched path<sup>2</sup> through an IP network by applying a series of traffic engineering constraints. The constraints can include a specified path, bandwidth, acceptable network resources, and/or a specified quality of service. Once the tunnel is created, labels are assigned and distributed accordingly. IP packets matching the desired destination prefixes at the head end of a given tunnel are encapsulated in a labeled envelope and sent into the tunnel.

MPLS introduces new terminology into the familiar layer 3 landscape. A router or switch capable of supporting MPLS is referred to as a label switching router (LSR). The MPLS tunnel is referred to as a label switched path (LSP) and is constructed from a series of unidirectional links called Label Switched Hops. The entry point into the tunnel is called the ingress or head end and the last router in the MPLS label switched path is called the egress. MPLS terminology includes a name for the LSR that precedes the egress LSR, this node is known as the penultimate LSR. Label switching routers along a given label switched path, between the ingress and egress nodes, are referred to as midpoint label switching routers.

The creation of MPLS label switched paths involves multiple components. There is a traffic engineering policy component that provides a user interface for selecting the traffic paths associated constraints. There is an IGP component which is composed of traffic engineering extensions to IS-IS and OSPF (IS-IS-TE and OSPF-TE respectively) routing protocols. There is a signaling component which is based on traffic engineering extensions to RSVP (RSVP-TE) or an entirely new protocol called Constraint Based Label Distribution Protocol (CR-LDP). Lastly, there is the data-forwarding component, which is based on a label-swapping scheme (see "Traffic Engineering with MPLS and ATM" on page 16).

Traffic Engineering uses MPLS to construct paths through the Internet core. The actual path selection is performed using a specialized database contained on each Label Switching Router called the traffic engineering link state database (TE-LSDB). The TE-LSDB contains the network topology of the Internet core (bounded by a single IGP area). After the traffic engineer inputs the constraints for path selection (egress, desired path, bandwidth, and inclusion/exclusion of label switching router interfaces), the TE-LSDB is pruned of non-compliant links and the shortest label switched path is selected.

## MPLS Applications

The first applications of MPLS will most likely be bounded to single IP Autonomous System. In the Internet core the most cogent application for MPLS and traffic engineering is short cut routing. Short cut routing is accomplished by applying engineered paths between core Internet exterior BGP routers. Applying fast reroute techniques and Quality of Service awareness to the engineered paths can then further refine and differentiate the short cut route. This section describes the application of short cut routing, label switched path restoration techniques, and the intersection of QoS and MPLS.

---

<sup>2</sup> The terms tunnel and label switched path (or LSP) are used interchangeably throughout this paper.

## ***Short Cut Routing***

### **BGP NEXT HOP**

Referring to Figure 1 it can be noted that BGP is used exclusively for peering between different Internet providers and sometimes between providers and corporate customers. From a backbone provider's perspective BGP is used to acquire routes from the backbone provider's peers/subscribers. The backbone provider then decides which subscriber routes to advertise and to whom. Each route advertised to a peer/subscriber contains a BGP NEXT HOP record. The BGP NEXT HOP record is typically the IP address of an interface on the remote end of the link between the backbone provider's router and its peer/subscriber router. A single backbone provider BGP router can peer with multiple subscribers. Therefore the links to its peers lead to multiple BGP NEXT HOPS. As IP packets enter a backbone provider network they are routed by IGP toward the BGP NEXT HOP.

### **A MPLS Short Cut to BGP NEXT HOP**

A common early implementation of MPLS and traffic engineering will be the establishment of tunnels between all external BGP routers within a single provider's core backbone (bounded by a single AS). The tunnels will be engineered away from the IGP SPF and areas of congestion on the backbone. After a tunnel is successfully signaled between two BGP routers, the ingress side must map the prefixes destined for the tunnel into its forwarding table. The selection of prefixes, in the application of short cut routing, is referred to as the FEC (Forward Equivalence Class) and is predetermined by the BGP NEXT HOP routes directly accessible by the egress router. Figure 3 depicts two short tunnels, tunnel 1 services all BGP NEXT HOPS in ISP1, and tunnel 2 services both ISP 3 and 4. Figure 3 provides a simplified forwarding table for the Ingress LSR (router A). From the forwarding table it can be observed that IP packets (arriving at the ingress router A) with destination IP prefixes intended for ISP 1 will be placed into LSP1. IP packets with destination IP prefixes intended for ISP3 and ISP4 will be placed into LSP2. Arriving packets with destination IP prefixes for ISP 2 will be forwarded to router B on route to router E, along the shortest path.

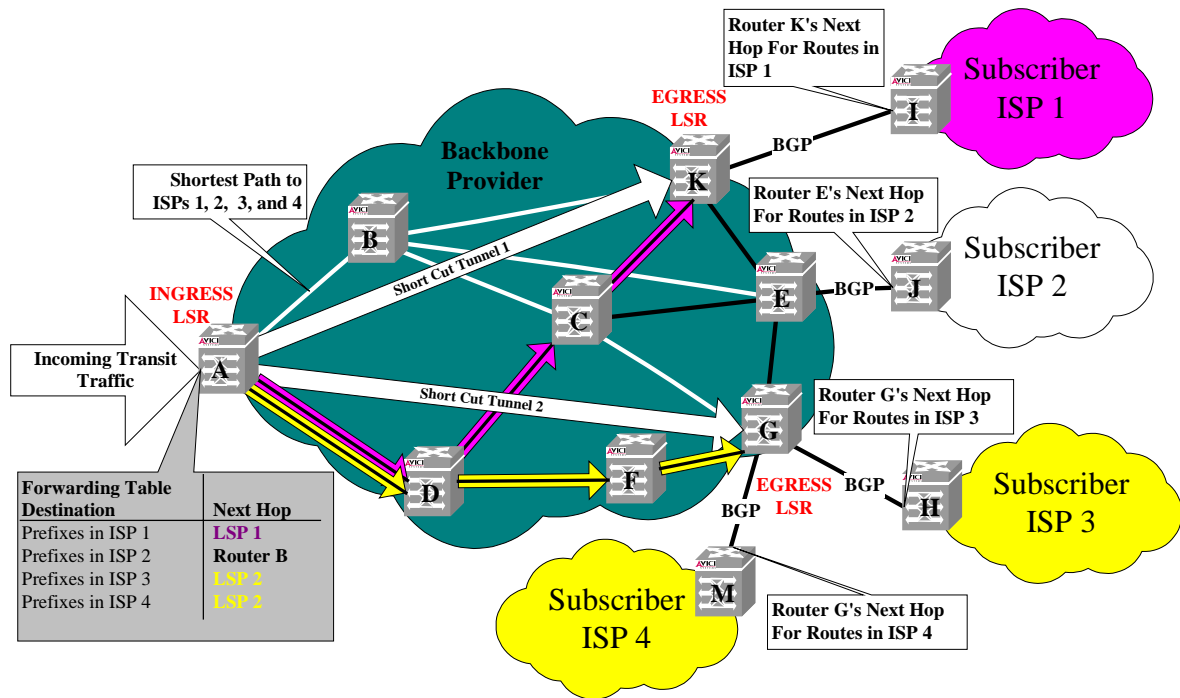


Figure 3 Short Cut Tunnel

It is interesting to note that without the ability to filter routes from the forwarding table a shortcut tunnel will, by default, carry prefixes for all BGP NEXT HOPS adjacent to the egress label switching router. Avici will provide route filtering within BGP access lists to further tune which prefixes are admitted into the shortcut tunnel.

### Tunnel Restoration

Path protection is an essential element in designing high-speed networking. Path protection can be provided at multiple layers in a communications protocol stack.

- Optical Layer: Performed with full or partial mesh of the optical switching systems.
- Physical Layer: Performed with a variety of Automatic Protection Switch strategies exploiting SONET line or path switching.
- MPLS Layer: Performed with head end reroute procedures (see below).
- IP Layer: Performed by BGP and IGP convergence algorithms.
- Avici composite link technology: (see Avici Composite Links on page 18) If a member of a composite fails, traffic is rerouted over surviving members in under 45ms. MPLS and IP can be provisioned to use composite links as a method for protecting POS links between adjacent routers/label switching routers.

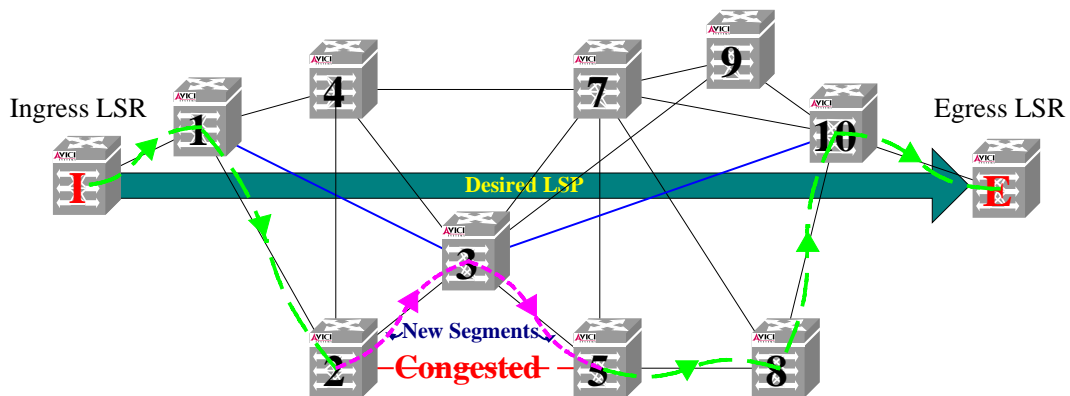
## Planned Head End Reroute Capability

The head end can reroute tunnels (that it signaled) that are not sustainable for a variety of reasons:

- Administrative change to label switched path entity
- Congestion along a given label switched path
- Failed link
- Failed node

The stimulus for a head end reroute can come from one of three sources: the command-line interface, notification by the IGP of a change in topology, or notification from RSVP-TE that the path can not be sustained. When the head end determines that it must perform a reroute the following events occur:

The head end label switching router constructs a new constrained SPF to egress after pruning the TE-LSDB from link faults and new areas of congestion. The head end then signals a new tunnel by creating label switched paths with the shared-explicit reservation style. The shared explicit reservation style allows new label switching paths to be created over acceptable links that are already carrying label switched paths from the prior tunnel without double counting reserved bandwidth. Thus a new tunnel can be overlaid atop portions of the original tunnel without reapportioning link bandwidth (see Figure 4). The head end uses a different LSP-ID for each new label switched path, effectively masquerading as a different sender. When the new tunnel is completed, the head end will modify its forwarding table and redirect the prefixes accordingly. After the reroute is completed, the original label switched path is torn down to return resources to the network.



Path Message contains request for "Shared Explicit" Reservation Style.  
Same Tunnel ID but new LSP\_ID

—→ Path Messages Traversing existing LSP segments toward Egress

- - -→ Path Messages Traversing new LSP segments toward Egress

Figure 4 Head end Reroute with Shared Explicit

Head end rerouting is best suited for path re-optimization as opposed to fault recovery. The reconvergence time for a head end reroute around a failure link is nondeterministic based on the complexity of the path. Head end rerouting can be optimized by preprovisioning an alternative tunnel or by performing load balancing of a FEC across multiple tunnels. In both cases, once the ingress is notified of a fault it will immediately modify its forwarding table and reroute affected prefixes.

### Fast Reroute

An alternative to head end reroute is local or fast reroute. Fast reroute requires the creation of a backup label switched path for each primary label switched hop within a given tunnel. The backup label switched path can provide protection against a failed link or a failed label switching router. The establishment of the backup Label Switched Path is the responsibility of a midpoint label switching router. But the constraints expressed for the original tunnel are only articulated to the ingress label switching router. Therefore RSVP signaling (see “Signaling for Label Distribution” on page 22) must be extended to carry information about required bandwidth and administrative colors to each midpoint label switching router. Assuming the midpoint label switching router is informed of the original constraints, it can then determine backup paths consistent with the traffic engineering policy expressed at the Ingress Label Switching Router.

The Internet Draft [draft-swallow-rsvp-bypass-label-00.txt](#) proposes the use of bypass tunnels to protect multiple primary tunnels. Bypass tunnels are set up using standard signaling. A separate bypass tunnel is required for each label switch hop and must intersect the protected tunnel(s) at two points referred to as the Point of Local Repair (PLR) and the Merge Point (MP). If a fault occurs anywhere on the primary tunnel the affected traffic is rerouted over the bypass tunnel at the PLR node and will rejoin the primary at the MP LSR.

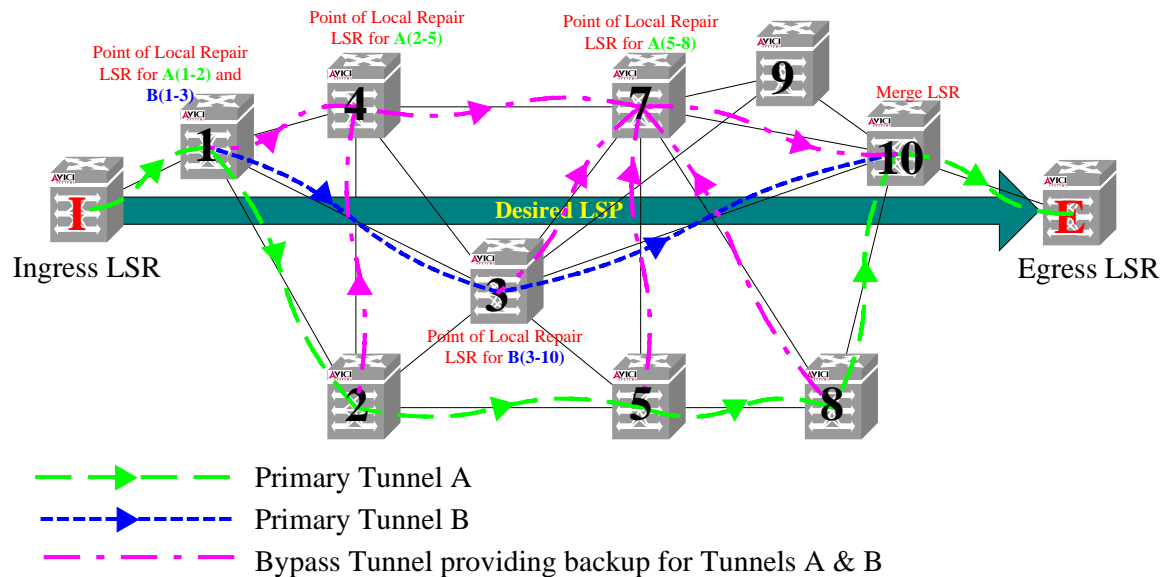


Figure 5 Bypass Tunnel Concept

Figure 5 depicts a series of bypass tunnels which backup two primary Label Switched Paths, **A** and **B**. A key requirement for the bypass tunnel, to provide protection to multiple Label Switched Paths, is the use of label stacking. In referring to Figure 5, if the link between LSR 1 and LSR 2 should fail, LSR 1 (the PLR) must redirect labeled packets over the bypass tunnel. The [draft-swallow-rsvp-bypass-label-00.txt](#) proposes label stacking to solve the problem of LSR 10 (the merge node) correctly associating labeled packets arriving from LSR 7 (over the bypass tunnel) as belonging to the original tunnel (**I-1-2-5-8-10-E**). LSR 1 must label packets intended for tunnel **A** with the label that LSR 10 (the merge node) expects to be receiving from LSR 8<sup>3</sup>. LSR 1 must then push a second label onto the packet intended for LSR 4 (the next hop of the bypass tunnel). Packets for Tunnel **A** are then shuttled over the bypass tunnel using normal label switching. At label switching router 10 a label stack "POP" occurs exposing the label that LSR 10 is expecting from LSR 8, thus closing the circuit around the failure.

### Avici Failure Affinity Groups

Provisioning protected paths is dependent upon the level of diversity available in the network. To be effective the backup label switched path must not utilize overlapping physical resources with the primary. In support of this restriction, Avici has introduced the notion of Failure Affinity Groups. A Failure Affinity Group is defined on each label switching router and contains network resources that share a common vulnerability. Examples of failure affinity group members include:

- Interfaces that connect to fiber in a common conduit
- Interfaces that connect to a common transit POP
- Interfaces that connect to a common next hop label switching router

When label switched paths are created with local protection each midpoint label switching router must signal a protected label switched path for each primary label switch hop. The midpoint label switching router consults its local TE-LSDB to determine an optimal CR-SPF to protect the primary label switched path and rejoin the primary path. The TE-LSDB on the midpoint treats links with common failure affinity as a constraint and prunes them accordingly. Thus protected paths are created to maximize diversity and minimize failure affinity.

### Integrating MPLS and QoS

It has been amply demonstrated that MPLS is suited for traffic engineering. It has also been noted that the first application for MPLS will be shortcut routing in the Internet core. When compared to normal SPF routing or ATM PVC's, shortcut routing as an application of MPLS, is a significant achievement. But much more can be accomplished when MPLS is tied to Quality of Service. For a more complete discussion on MPLS and QoS

---

<sup>3</sup> The [draft-swallow-rsvp-bypass-label-00.txt](#) proposes that the RSVP-TE Record Route Object (RRO) be enhanced to include the labels assigned to each router hop along the signaled path. Therefore the PLR node (LSR 1) depicted in Figure 5 has advanced knowledge of the label LSR 10 is expecting from LSR 8 and will push the label on the packets destined to Tunnel A's egress.

refer to [Delivering Internet Quality of Service](#) on the [www.avici.com](http://www.avici.com) website. There are two proposals for adding QoS capabilities to today's best effort Internet: Integrated Services and Differentiated Services.

### Integrated Services

QoS can be flow based or class based. Flow based QoS assumes visibility to all flows in the Internet. Individual flows might be characterized by IP source-destination pairing, additionally the layer 4 header could be considered in a flow definition. Characterization of individual flows is typically a function of the Internet edge (close to flow origination). To maintain end to end QoS, flows must be signaled end-to-end through the Internet with a reservation protocol (RSVP). This approach is considered untenable within the Internet core due to the sheer number of flows coupled with the overhead associated with RSVP, especially for short-lived flows. Flow based QoS is described in RFC 1633 [Integrated Services in the Internet Architecture: an Overview](#). Integrated services or "IntSrv" defines two classes of service:

- **Guaranteed:** Bandwidth, bounded delay, and no-loss guarantees
- **Controlled load:** About as good as best effort traffic in a lightly loaded network

### Differentiated Services

Class based QoS involves marking the packets, within a flow, with a tag indicating the requested QoS treatment on a *PER HOP BASIS*. This approach is referred to as differentiated services (DiffServ) and provides Per Hop Behavior (PHB) as opposed to circuit oriented microflows with explicit reservation. Differentiated Services is considered highly scalable because the Per Hop Behavior can be provisioned in advance of the data flows on all routers in the path. PHB can be easily defaulted to predefined packet treatments for IP Precedence values (defined in RFC 1812 and RFC 791) or DiffServ codepoints (DSCP defined in RFC 2474). DSCP and IP Precedence are derived from the same byte in the IPv4 header. Within the DSCP byte, six bits are used as DSCP codepoints.

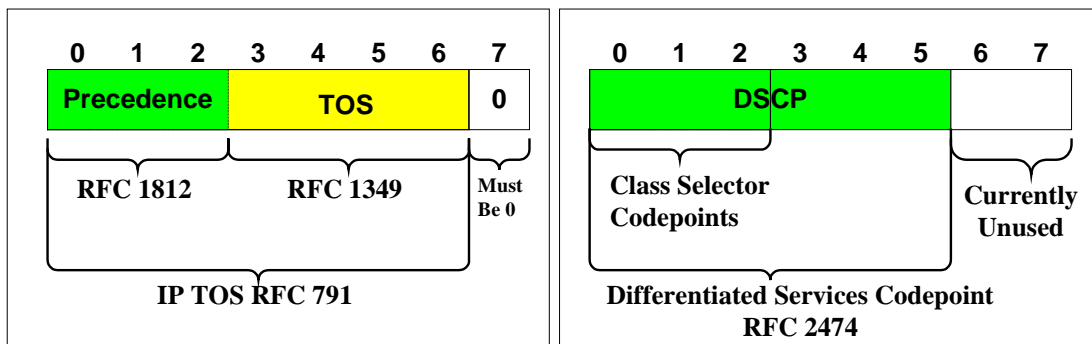


Figure 6 DiffServ Field in IPv4 Header

When packets arrive at the Internet core they can be classified and, based on DSCP bits, queued appropriately. DSCP settings can be aggregated into common classes and treated with identical PHBs.

Currently there are four proposed mappings of DSCP which together equal 22 unique code points:

- **DE:** Default, which is the best effort class (1 code point)
- **CS:** Class Selector, Backward Compatible with IP Precedence (8 code points)
- **EF:** Expedited Forwarding; approximates constant bit services, with controlled load (1 code point)
- **AF:** Assured Forwarding; defines four hierarchical classes and three drop precedences within each class (for a total 12 code points)

### IntServ Meets DiffServ and MPLS at the Internet Core

IntServ is a connection-oriented service that programs edge routers with microflows to maintain QoS contracts from the user/client to the application/server. The Internet has, until the advent of MPLS, been a completely connectionless system. The connection-orientated nature of MPLS differs significantly from IntServ in terms of aggregation. MPLS tunnels serve to aggregate flows into LSPs; IntServ serves to desegregate IP traffic into micro-flows. The two seemingly incongruous circuit oriented systems come together elegantly with DiffServ.

DiffServ and IntServ meet at the boundary between the customer edge and the provider edge. The microflows terminate at the provider edge router where packets are given DSCP settings that match the QoS requirements established within the microflow. Packets marked with DiffServ code points are then sent to the Internet core where they are classified by provider core routers and injected into MPLS LSPs that meet the QoS requirement of the DSCP. There are two proposed methods for aggregating DiffServ marked packets into MPLS tunnels for QoS: **L-LSPs** and **E-LSPs** (draft-ietf-mpls-diff-ext-04.txt).

#### *L-LSPs for Mapping DiffServ to MPLS*

L-LSP or label inferred LSPs associate the layer 3 DSCP with a specific layer 2.5 MPLS label. Stated another way, each label switching router that contributes a hop to the tunnel, built of L-LSPs, has packet scheduling that meets the QoS level defined by the DSCP provisioned at each hop. The ingress label switching router examines the DSCP in the IP header and selects a label switched path that has been provisioned for that QoS level. Each label switching router in the path examines the incoming label and determines the QoS treatment for the encapsulated packet. At the egress label switching router the last label is removed and the packet is sent to the next IP hop with its original DSCP. This method requires that an association of specific DiffServ code points to label switched paths be pre-established prior to traffic flow. There are still open issues regarding proposed extensions to RSVP-TE to signal the association of labels to packet scheduling, but RSVP was originally designed for this type of signaling (IntServ) thus it is not a stretch to add this functionality to RSVP-TE.

A critical requirement with the use of L-LSPs is maintaining DSCP to MPLS mapping at every label switch hop. The Avici TSR system is unique in its ability to fulfill this

requirement. The Avici TSR system performs all QoS packet processing in specialized Application Specific Integrated Circuits (ASIC). The TSR system also performs hop-by-hop label switching and conventional IP next hop processing in ASICs. Processing labeled packets at the egress label switching router is very hardware intensive because it involves a next hop label lookup leading to a “pop” operation and then layer 3 packet forwarding based on another lookup, all in a single step.

#### *Egress Label Switching Router Sophistication*

When packets arrive at the egress label switching router they must be provided the appropriate packet treatment to maintain end-to-end QoS, based on the incoming label. Prior to packet scheduling *two* lookups must occur:

- 1) The *first* lookup is used to determine the outgoing label (this is performed at each label switching router hop). In the egress label switching router case, the lookup for the outgoing label leads to a POP operation, (removing the top label), thus promoting the packet back into the layer 3 IP domain.
- 2) Once in the layer 3 IP domain a *second* lookup is necessary to determine the IP next hop.

The Avici TSR system is able to perform the two lookups concurrently with no performance penalty due to its advanced ASIC pipeline for label switching and processing next hops in the forwarding engine. This contrasts with currently available gigabit routers. The gigabit routers that reside in the Internet’s core must perform a pop operation on the label switching router *prior* to the Egress label switching router known as the “Penultimate LSR”. The “Penultimate LSR” must be provisioned, at label switched path establishment, to switch the incoming label with an outgoing “NULL” label before releasing the packet to the label switched path/tunnel Egress label switching router. The use of the NULL label in this fashion is referred to as “Penultimate Hop Popping”. When the packet is received by the egress label switching router with the NULL label it has, in effect, already been promoted to layer 3. Therefore only a single lookup is necessary to determine the IP layer 3 next hop. Penultimate hop popping is a clever workaround to the problem of performing two simultaneous lookups but it creates a glaring hole in end-to-end QoS because the packet’s output queuing cannot be differentiated.

#### *L-LSPs and Fast Reroute*

The strength of Label Inferred label switched paths is their relationship to fast reroute restoration services. Packets arriving at the ingress label switching router with DiffServ code points that dictate premium service (Expedited Forwarding) will be labeled for paths that are fast reroute capable. Other non-premium packets (Best Effort) with the same destination can be sent into a short cut tunnel that is not fast reroute capable. The determination for use of fast reroute services is made by service level agreements between the subscriber and the provider.

### *E-LSPs for Mapping DiffServ to MPLS*

The exclusive use of L-LSPs, as a means to aggregate packet treatments, is very resource intensive. DSCP has enough bits to provide 64 unique code points. It is doubtful that all 64 distinct QoS levels will ever be required, or for that matter, standardized. In any event, a 1:1 ratio of DSCP to Label inferred label switched path can potentially consume excessive labels and network resources. Therefore a second mapping of DSCP to MPLS has been proposed called E-LSPs or EXP (based on the use the “experimental” bits of the MPLS shim header) inferred label switched paths.

In referring to Figure 2, there are 3 bits in the Experimental Field of the MPLS Shim Header. The E-LSP is a direct mapping of up to 8 DiffServ code points into 8 possible EXP values. The mapping of DSCP to EXP is made by the Ingress label switching router. Once marked with an EXP setting the packet scheduling at each hop of the E-LSP tunnel mimics the non-MPLS IP class based QoS (per hop scheduling). E-LSP tunnels are much more frugal, in terms of label consumption, than their L-LSP counterparts. However a mapping of EXP bits cannot be easily associated with a fast reroute service, since all eight possible EXP markings share a common tunnel.

L-LSPs and E-LSPs each have a critical role to play in delivering flexible and effective Quality of Service to the Internet core. The mapping of DiffServ into labels and/or EXP bits is a tool to be exploited by the provider to meet even the most rigorous service level agreements.

## Traffic Engineering with MPLS and ATM

### *What is Traffic Engineering with MPLS?*

Today ATM is used extensively by network core providers to engineer paths between backbone routers. The combination of ATM and IP requires the establishment of two independent topologies (L2 PVC's and I-BGP Peering) see Traffic Engineering with MPLS and ATM Traffic Engineering with ATM on page 16. The first goal of traffic engineering with MPLS is to eliminate the signaling redundancy found in the ATM layer but preserve the notion of a circuit between backbone routers. The circuits carrying MPLS traffic are often referred to as traffic trunks and can be routed independently of the underlying link topology. Another important goal of traffic engineering is the creation of true differentiated IP services. When Quality of Service is combined with MPLS traffic engineered tunnels, incoming traffic can be classified and routed into MPLS tunnels that are engineered at each hop to deliver the requested QoS.

As previously stated, Traffic Engineering with MPLS has four components:

- User interface for articulating traffic engineering policy in terms of constraints to conventional SPF
- IGP component which is composed of traffic engineering extensions to IS-IS and OSPF
- Signaling component which is based on traffic engineering extensions to RSVP or CR-LDP

### Traffic Engineering Policy

Traffic engineering policy provides the traffic engineer with a rich set of knobs to define paths through the Internet. Path Selection is a key deliverable of MPLS. The degree of control in path selection is an area for vendor differentiation. The goals for MPLS in general and path selection in particular are partially articulated in [RFC 2702, Requirements for Traffic Engineering Over MPLS, D. Awduche, September 1999](#). The requirements cited in RFC 2072 are paraphrased below:

#### *Administratively Specified Explicit Path Selection*

Path selection can range from a strict hop by hop delineation to a loose path definition containing only the desired ingress and egress. The Avici traffic engineering component allows completely strict, completely loose, and any combination of strict and loose, see Figure 7.

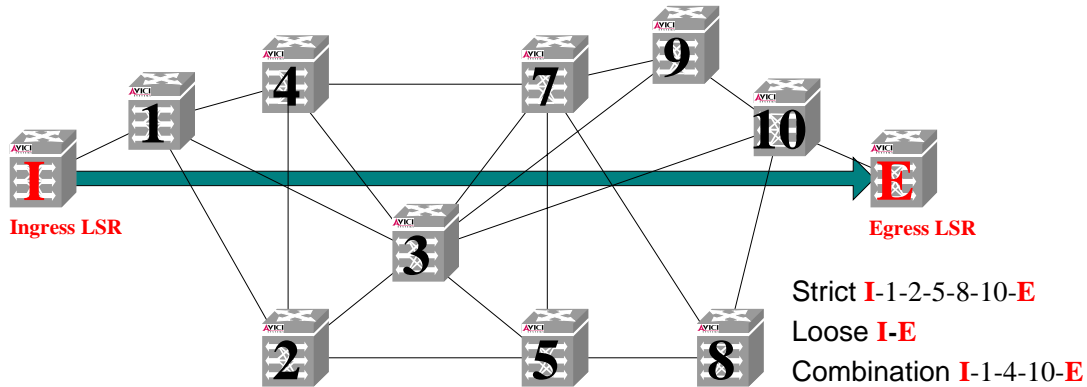


Figure 7 Administratively Specified Explicit Paths

### Resource Class Affinity

Resource classes are defined as static assignment of label switching router interfaces to groups referred to as colors. For example, premium label switching router trunks could be assigned to the blue group, while slower OC-3 trunks could be assigned to the yellow group. Additionally resource class affinity can be used to identify preferential links for protection schemes or for bulk transport. In Figure 8, blue (solid) links provide premium OC-48c trunking. Red (dashed) links provide bulk OC-12c service, and links that are in the (semi-dashed) green affinity group have some form of protection (optical or physical layer).

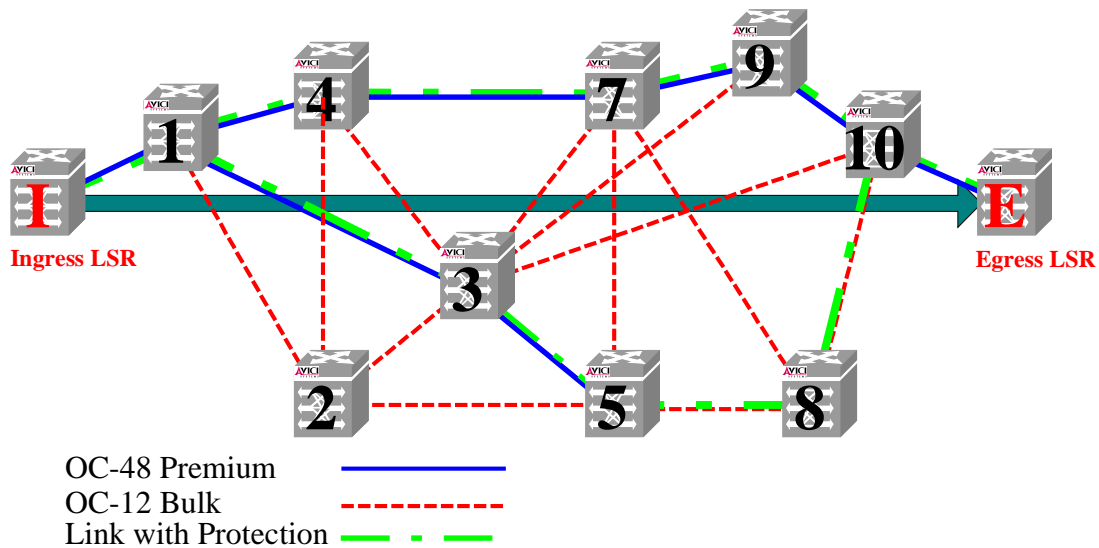


Figure 8 Resource Affinity with Color Groups

The traffic engineer can then use the predefined affinities to tailor the path selection accordingly. If a path requires bulk protection, the tunnel creation command would include the color green.

### *Label Switched Path Adaptivity*

Adaptivity is a measure of how responsive an existing label switched path is to changes in the underlying network topology. An adaptive label switched path will dynamically reroute in a make-before-break fashion to provide a more optimum path. A reroute can be triggered by topology changes caused by network faults or congestion within an MPLS tunnel. Other sources of topology change include the addition of new physical links between label switching routers or manual changes to existing resource affinities (thereby increasing the capacity of an existing tunnel within the original constraints). Adaptivity controls the degree of “re-optimization” of a given MPLS tunnel or label switched path. Re-optimization is a re-route operation of one or more segments of an established label switched path. Adaptivity is expressed as an attribute used with the tunnel creation command. The adaptivity attribute is a binary value either permitting or denying re-optimization.

Closely associated with the adaptivity attribute is the adaptivity timer. The adaptivity timer is used to dampen the label switched path reroute signaling in the presence of a network fault or congestion along the label switched path. Dampening preserves network stability in presence of short term loading along label switched paths that are provisioned for adaptability.

### *Label Switched Path Bandwidth Reservation Priority*

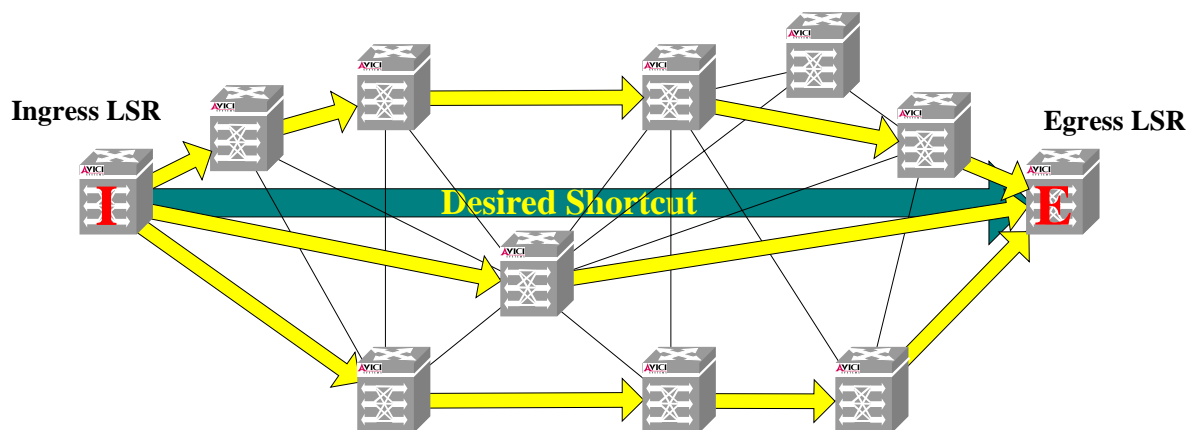
MPLS label switched paths can be provisioned to reserve bandwidth from either actual link bandwidth or from some amount of over subscribed bandwidth (per link). In initial implementations of MPLS, tunnels are not matched to weighted queues necessary for precisely allocating and policing bandwidth. Instead, bandwidth granted to a label switched path is deducted from the available link bandwidth. Link bandwidth is allocated into 7 holding priorities, 0 being the highest. The desired holding value is included in the tunnel creation command. The tunnel will be created only if there is sufficient bandwidth at the requested holding priority. For example, if a tunnel is signaled with a request for 5 Mbits of reserved bandwidth at a holding priority of two and there is no bandwidth left at that priority the tunnel will not be created.

### *Avici Composite Links*

Avici uniquely supports the ability to interconnect two TSR systems with multiple SONET links, which look to the IP routing process like a single PPP link. Avici refers to this capability as composite links. The TSR system supports up to 64 member links. Load is balanced across the links in a weighted fashion permitting a 4:1 mismatch in link bandwidth (OC-48c can be mixed with OC-12c). Packet ordering is maintained by hashing the source and destination IP addresses of data sent over the composite link. MPLS fully utilizes the composite link. As additional member links are added, the MPLS path selection process (on all originating ingress label switching routers that utilize the affected composite link) is notified (see Enhancements to IGP and the Role of the TE-LSDB on page 20) of the additional bandwidth.

### *Load Distribution Across Parallel Label Switched Paths*

Parallel label switched paths are defined as sharing the same ingress and egress in conjunction with similar traffic engineering constraints. To achieve load distribution across multiple label switched paths, incoming packets must be hashed using source and destination IP addresses and then labeled in a weighted round robin fashion. The distribution of packets onto label switched paths is weighted by the percentage of the total bandwidth assigned to each path. Packet ordering is preserved but the effectiveness of the load distribution is dependent of the diversity of the traffic flows.



**Figure 9 Multi-LSP Load Balancing**

Avici uses the term composite paths to describe our implementation of load balancing. Composite paths can be used for fast restoration of link failures. If one of the composite path members experiences a link failure, packets hashed to that label are rerouted by the ingress label switching router over a surviving trunk. The ingress is notified either by the IGP or by RSVP Signaling.

Additionally the Internet Draft for MPLS-Optimized Multi-Path (draft-villamizar-mpls-omp-01.txt) presents a dynamic load distribution algorithm based on OMP enhanced IGP flooding of link utilization along label switched paths. MPLS-OMP is ideally suited to best effort traffic.

### *Label Switched Path Preemption*

Order is important in the establishment of MPLS tunnels. Each successfully signaled tunnel removes available bandwidth from each link in its path at a given holding priority. Eventually bandwidth resources will become strained and limit tunnel creation. Label switched path preemption introduces a setup (or reservation) priority attribute for use in conjunction with the aforementioned holding priority. If the setup priority of a requested tunnel exceeds the holding priority of an existing tunnel and resources at that holding priority are exhausted, then the new label switched path will preempt the existing tunnel. The existing tunnel will be destroyed, freeing its resources and the new tunnel will then be established. If there are multiple tunnels at a subordinate holding priority the decision for selection the preempted tunnel is a vendor implementation detail.

### *Label Switched Path Resilience*

There are two dimensions to label switched path resilience, the first is concerned with label switched path establishment, and the second is concerned with responsiveness to hard faults in label switched paths. In the first case it can be observed that there are numerous options (knobs) available to the traffic engineer for tunnel creation. The abundance of knobs is necessary for flexibility but can frustrate the underlying goal of creating label switched paths that overlay atop the Internet core. Avici has provided an innovative way to allow maximum flexibility in applying constraints without losing sight of the goal of creating a feasible tunnel. This is accomplished with the application of policy constraint attribute prioritization. In other words key attributes, including explicit path, requested bandwidth, and inclusion or exclusion of administrative colors can be assigned a priority. When the ingress router processes the request for a new tunnel, the constraints are now prioritized relative to one another. The user can therefore bias the constraints towards required bandwidth, network resource, and/or explicit path selection.

The second dimension to label switched path resilience comes after the tunnel has been successfully established and data flow is activated. Resilience in this context is concerned with fast recovery after a loss of a link carrying one or more label switched paths. Tunnel resilience is realized with a “local protection” option used at tunnel creation. With local protection, each label switching router creates a protected label switched path (preferably using diverse links) for each primary label switched path. The topic of “fast reroute” is discussed in Fast Reroute on page 10.

### **Enhancements to IGP and the Role of the TE-LSDB**

OSPF and IS-IS are used as Interior Gateway Protocols (IGP) within the Internet core. Both OSPF and IS-IS are simply referred to as IGP. IGP utilizes a link state database (LSDB) on every participating router. IGP floods link state information, which advertise router adjacencies, link costs, and changes to link state to all routers in the IGP domain (or area). Each router runs a Dijkstra algorithm against the LSDB producing the shortest paths to all destination prefixes with the local router as the root. The information flooded by IGP is insufficient for traffic engineering. Thus IGP has been enhanced (referred to as OSPF-TE and IS-IS-TE) to include three new flooded messages (referred to as Type, Variable Length messages –TLV):

- Reservable Bandwidth at each priority (0-7)
- Link Color assignments
- Traffic engineering assigned metrics

Please note that these new flooded announcements come from traffic engineering provisioning performed on each label switching router. Information related to neighbor adjacencies is also provided. Label switching routers use the IGP extensions to create a new link state database specifically for traffic engineering (referred to as the TE-LSDB). The TE-LSDB contains the network topology of interconnected label switching routers within a single MPLS domain (usually synonymous with the IGP domain or area).

As label switched paths are successfully established bandwidth is naturally consumed. The TE-LSDB is then updated by virtue of IGP flooding, with the available bandwidth (at

each holding priority) for every link in the MPLS domain. Information about link colors and traffic engineering metrics is also flooded but remains somewhat static by comparison.

The TE-LSDB in the head end router plays a pivotal role in tunnel creation. The signaling for tunnel establishment is controlled from the Ingress router. Therefore the commands for tunnel creation (including the articulation of constraints) are executed on the head end label switching router.

Unlike the IP LSDB, the TE-LSDB is not used to produce an exhaustive set of shortest paths. After the traffic engineer issues the tunnel creation command, the constraints (expressed on the command line) are used to prune links from the TE-LSDB. The Dijkstra algorithm is then applied to the remaining graph to produce a list of constrained shortest paths (CR-SPF) from the head end to the tail end label switching router (see Figure 10).

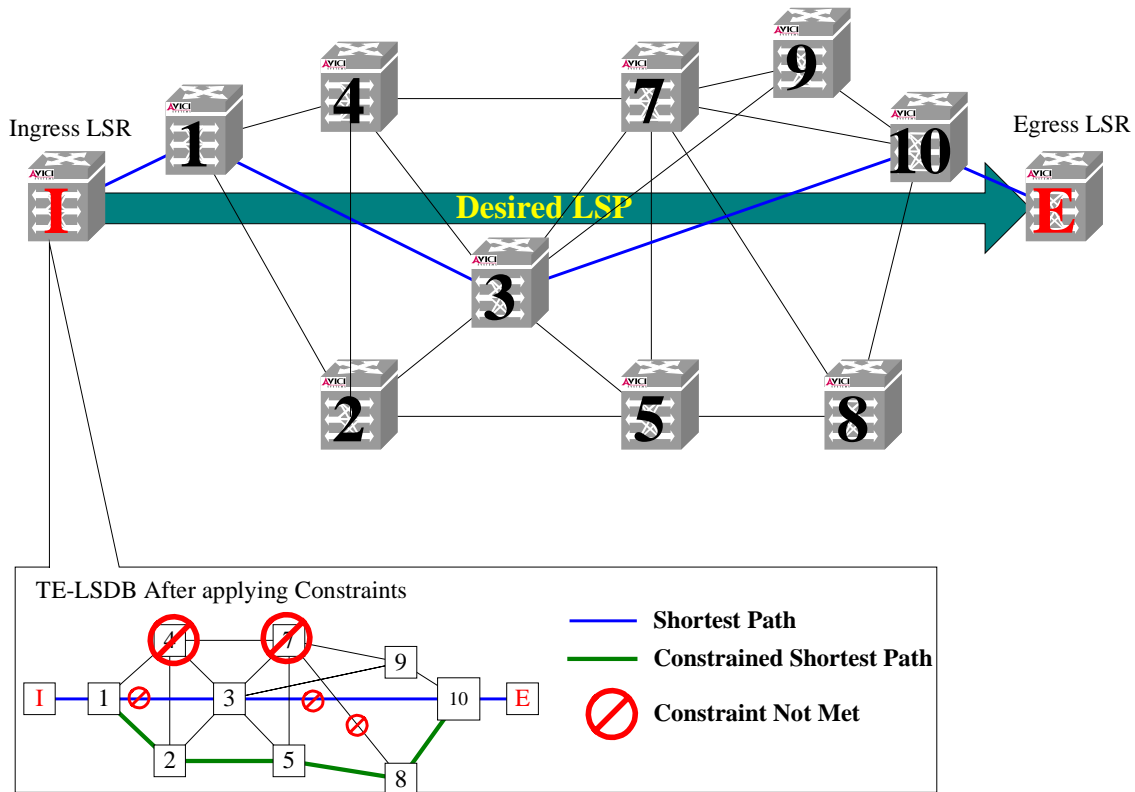


Figure 10 Role of TE-LSDB

Each time the constraints are changed the CR-SPF must be recalculated. It must then be observed that label switched paths are created in serial fashion since only a single CR-SPF can be determined at a time.

## Signaling for Label Distribution

After the CR-SPF has been determined for the desired tunnel two tasks remain before data is admitted into the tunnel: Labels must be distributed to every label switching router in the tunnel's path and the desired prefixes must be mapped to the appropriate labels. The prefixes directed into a tunnel are collectively referred to as the Forward Equivalence Class (FEC).

There are three signaling protocols defined for label distribution: Label Distribution Protocol (LDP), Constraint Based LDP (CR\_LDP), and ReSerVation Protocol for Traffic Engineering (RSVP-TE)<sup>4</sup>. LDP simply distributes labels along all SPF paths, thus producing label switched paths that mirror the IGP SPF. LDP does not provide traffic-engineered paths. LDP uses both UDP and TCP for signaling. CR-LDP is a set of protocol extensions to LDP which incorporate user defined constraints and thus enable traffic engineering. RSVP-TE is based on the original RSVP specification; intended for QoS based routing of microflows, but with significant enhancements tailored for the Internet core<sup>1</sup>.

### *RSVP PATH Messages*

After the TE-LSDB has been pruned of TE constraints (see Figure 10) the head end label switching router will begin the process of signaling the desired path. Signaling commences with the transmission of RSVP PATH messages. PATH messages are layered directly on IP. The PATH message contains the destination address of the Egress Router, but is locally processed by each router along the selected path (see Figure 11).

The PATH message contains special objects that are used by each label switching router along the PATH and for indicating to the egress label switching router the desired reservation style. The Path message contains the following objects:

- **Label Request Object**
- **Explicit Route Object ERO:** Identifies the route from ingress to egress
- **Record Route Object RRO:** Creates a path list of label switching routers visited by the PATH message
- **Session Object:** Assigns a global label switched path tunnel ID
- **Session Attribute:** Controls label switched path setup priority, preemption, reservation style<sup>5</sup> (Fixed Filter or Shared Explicit) and fast reroute

---

<sup>4</sup> Avici is delivering RSVP-TE signaling based on the installed base of MPLS and customer demand for interoperability with incumbent equipment providers. Therefore this paper will detail the basics of RSVP signaling for MPLS label distribution.

<sup>5</sup> The "Ingress node may reroute bit" is set in the Session\_Attribute Object to indicate the requirement of Shared Explicit reservation style.

- **Sender\_TSPEC**: Transmission Specification sent to egress to indicate desired reservation characteristics (QoS) for the sender's traffic

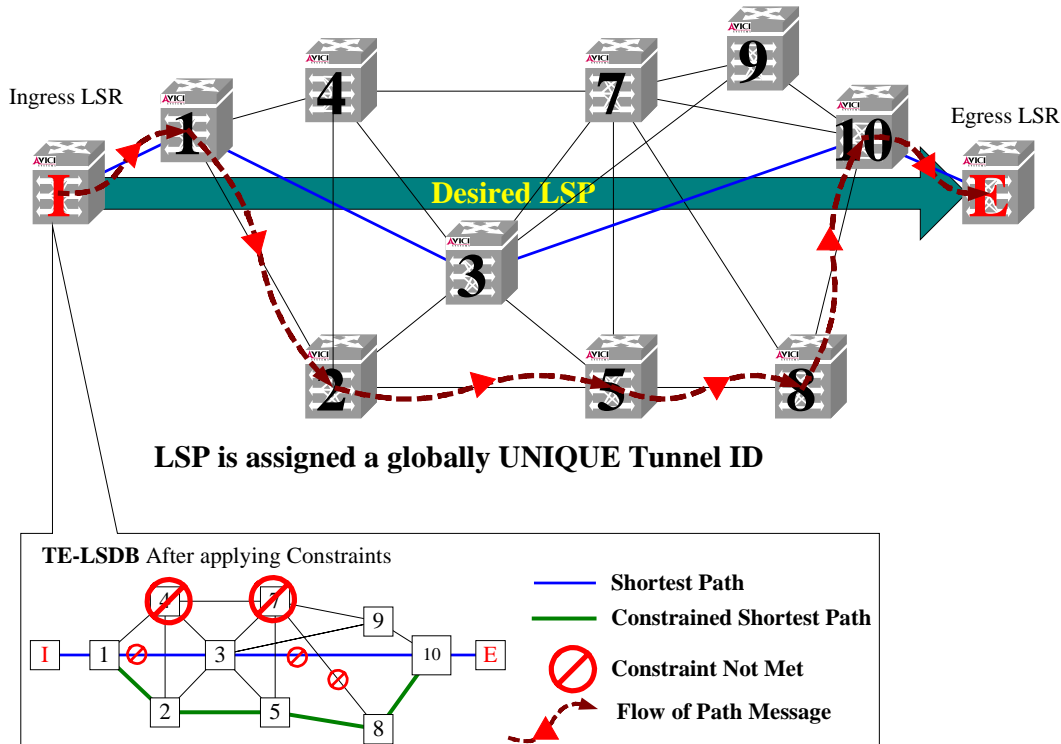


Figure 11 RSVP-TE PATH Messages

The PATH message follows the Explicit Route Object, which was defined in the policy definition for the desired tunnel. At each hop the Route Record Object (RRO) is updated with the IP address of the visited label switching router. The RRO is useful for loop detection and for network management tasks.

The Session Attribute contains the setup and holding priorities for the label switched path. As mentioned previously, if the requested bandwidth is unavailable at the included holding priority and the included setup priority is greater than an existing label switched path then preemption can occur. The preempted label switched path will be torn down (The Ingress LSR that signaled the preempted LSP would then attempt to reestablish the affected LSP). Also contained in the session attribute a semaphore that identifies the reservation style. There are two options for reservation styles implemented:

- Fixed Filter
- Shared Explicit

Multiple label switched paths established with the Fixed Filter reservation style, along a common link, will reserve separate bandwidth. Multiple label switched paths established with the Shared Explicit reservation, along a common link, will share the same bandwidth. The latter reservation style is necessary for path reoptimization and restoration and is typically the default setting.

## RSVP-TE RESV Messages

Upon receipt of the PATH message the egress label switching router (tail end of the tunnel) now has implicit knowledge that the requested path is feasible, and is aware of the desired reservation style. The egress label switching router must now initiate the Label Distribution process in the opposite direction to that of the PATH message, see Figure 12. The egress label switching router issues a RESV message addressed to the first downstream router in the RRO received in the PATH message. The RESV message contains the following objects:

- **Label Object:** Contains the actual label value used to associate the label with the signaled path.
- **Record Route Object:** Creates a path list of label switching routers used to route RESV messages back to the Ingress.
- **Session Object:** Provides a global label switched path tunnel ID copied from the PATH message.
- **Style Object:** controls label reservation style (Fixed Filter or Shared Explicit) Copied from the PATH message.

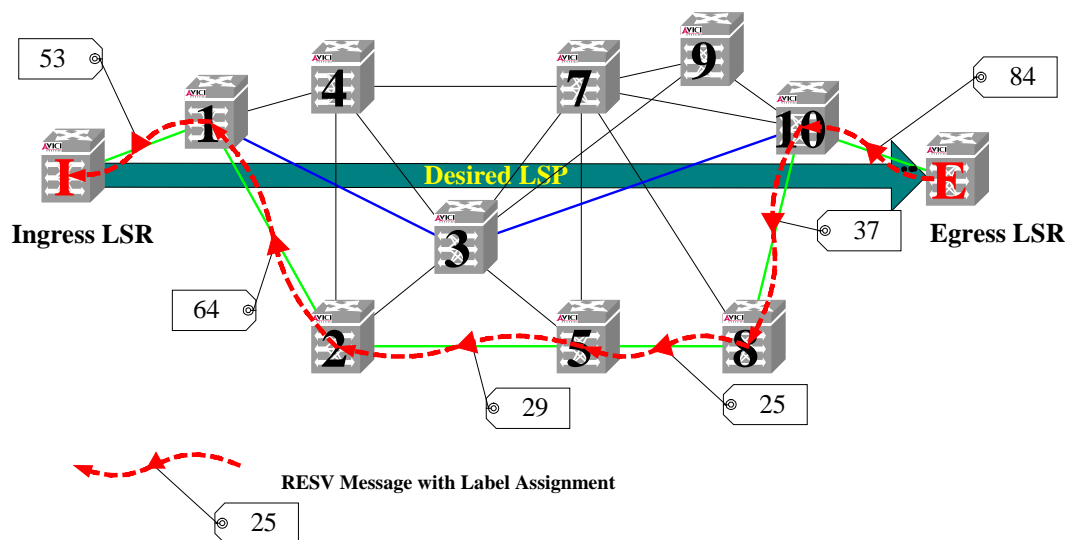


Figure 12 RSVP-TE RESV Message Flow

At each label switching router, the RESV message is used to identify and assign the label value to the incoming interface. Each label switching router must then allocate a local label for the next downstream label switching router (identified by the RRO). The label value is applied to the appropriate local interface and then copied into the Label Object for delivery to the next downstream label switching router. This process continues until the Ingress receives the final RESV message. Figure 13 depicts the label assignments based on the labels distributed in the above example.

LSR 1				LSR 1				LSR 2				LSR 5				LSR 8				LSR 10				LSR E			
LSP Tunnel ID 3				LSP Tunnel ID 3				LSP Tunnel ID 3				LSP Tunnel ID 3				LSP Tunnel ID 3				LSP Tunnel ID 3				LSP Tunnel ID 3			
In Interface	In Label	Out Interface	Out Label	In Interface	In Label	Out Interface	Out Label	In Interface	In Label	Out Interface	Out Label	In Interface	In Label	Out Interface	Out Label	In Interface	In Label	Out Interface	Out Label	In Interface	In Label	Out Interface	Out Label	In Interface	In Label	Out Interface	Out Label
-	-	2	53	1	53	3	64	2	64	4	29	1	29	4	25	1	25	3	37	1	37	3	84	1	84	-	-

Figure 13 Example Label Assignment

After the labels are distributed to all LSRs along the primary path the LSP is enabled for forwarding labeled IP packets. Each LSR forwards the incoming-labeled packet to appropriate next hop based the association of the incoming label with outgoing label and its assigned interface (see Figure 14)

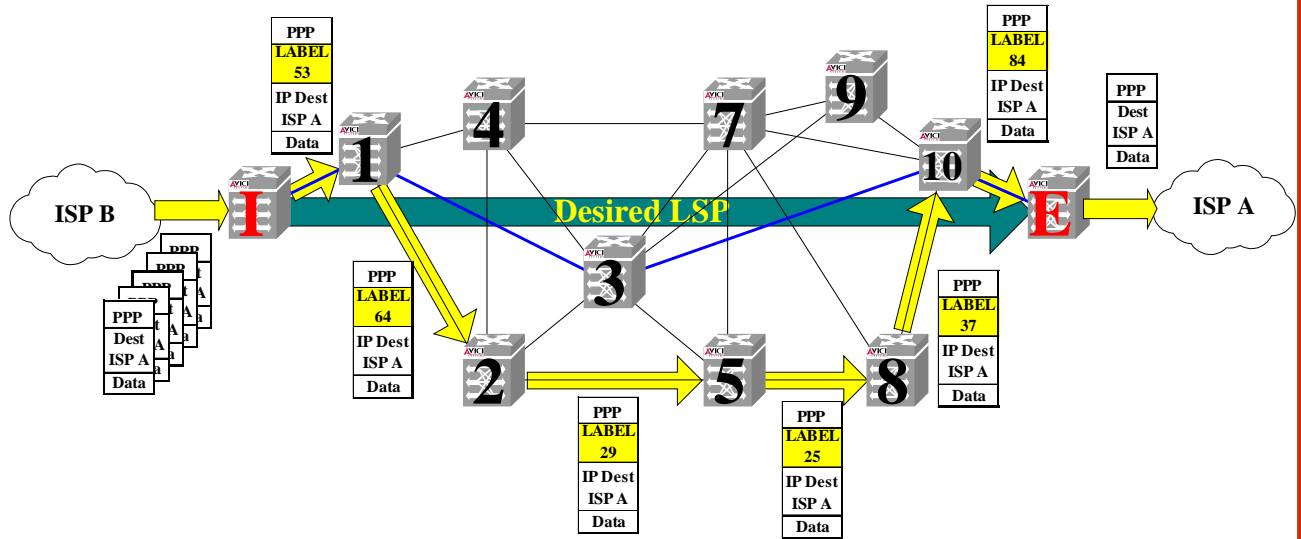


Figure 14 Forwarding Labeled Packets

## Current Solutions to Traffic Engineering

### Traffic Engineering with ATM

Prior to the introduction of MPLS as a means to achieve traffic engineering the only approach was ATM as a layer 2 transport. ATM has evolved to become a stable connection oriented transport that currently operates, ATM switch to ATM switch up to OC-48 line rate. ATM delivers traffic engineering with many advanced features including:

- PVC creation from any ingress to any egress in a given ATM backbone
- Sophisticated signaling to simplify path creation and reroute around failures
- QoS features for bandwidth reservation, constant bit rate, variable bit rate, and unspecified bit rate services, applied to the cell

## Drawbacks to ATM for Traffic Engineering

ATM is a mature technology with many advantages noted above, however there are also numerous drawbacks. One issue with ATM is overhead. ATM consumes 10% of available bandwidth with a 5 byte cell header for each 48 byte payload cell, plus an additional 5% is needed for the adaption layer for IP over ATM (RFC 1483 describes ATM Adaption Layer 5 AAL5). For example, an ATM OC-48 link requires 494Mbit/sec for overhead. Compounding the bandwidth issues is ATM's limited scalability at higher link rates. ATM switches have only recently delivered OC-48 interface rates and it is questionable whether OC-192 is feasible considering the overhead associated segmentation and reassembly, wasted bandwidth, and other inefficiencies of pushing 53bytes across 10Gbit/sec links. Today the fastest IP router ATM interface is OC-12, which create a bottleneck with the advent of OC-192 capable transport systems.

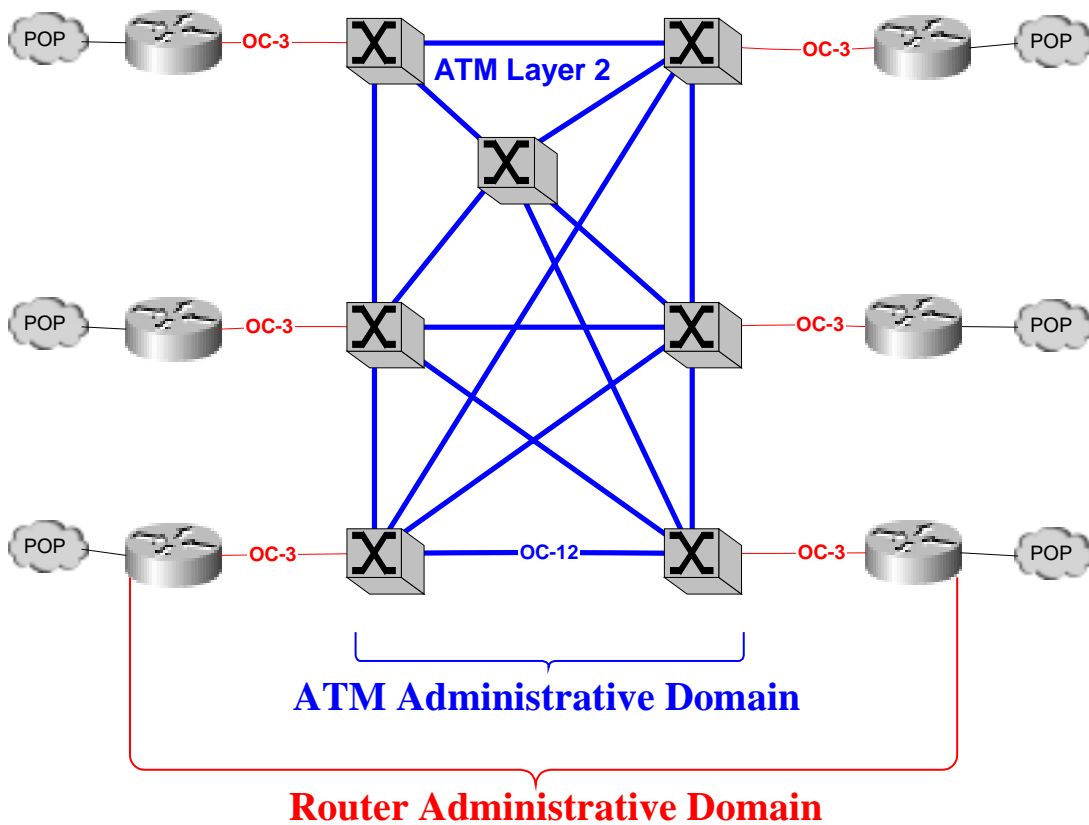


Figure 15 Layer 3 Topology with ATM

ATM in the Internet core also presents a significant administrative burden. ATM requires its own administrative domain distinct from IP at layer 3. Figure 15 depicts the two administrative domains. The ATM network elements must be interconnected in such a way to provide diversity and protection. The entire ATM topology is transparent to the constituent IP layer 3 topology. Therefore a second topology at layer 3 must be overlaid atop the ATM fabric (Figure 16). This is achieved by establishing PVC's between layer 3 routers. This creates three basic problems

- 1.) Two modalities for element management are required adding complexity and cost to network management.
- 2.) IP route exchange with an IGP, as noted earlier, requires direct peering/adjacency with all neighbors, therefore the number of PVC's required grow by a factor  $n^2$  where  $n$  is the number of internal IGP routers. For example, for 300 routers: 44,850 PVC's would be necessary to establish a complete mesh. If 4 more routers are added the PVC count jumps to 46,056 (an increase of 1206 PVC's). This represents a substantial network-provisioning problem. In the event of a router failure in this scenario, the surviving routers will issue IGP routing updates on the order of  $n^3$  (300 routers would issue 27 Million updates)
- 3.) ATM uses its own signaling protocol (PNNI) to establish PVC's. IP uses OSPF, IS-IS, and BGP as its signaling protocols. The two signaling layers operate independently and therefore complicate interworking between the layers. In the case of ATM based traffic engineering IP signaling protocols must run within the ATM PVCs.

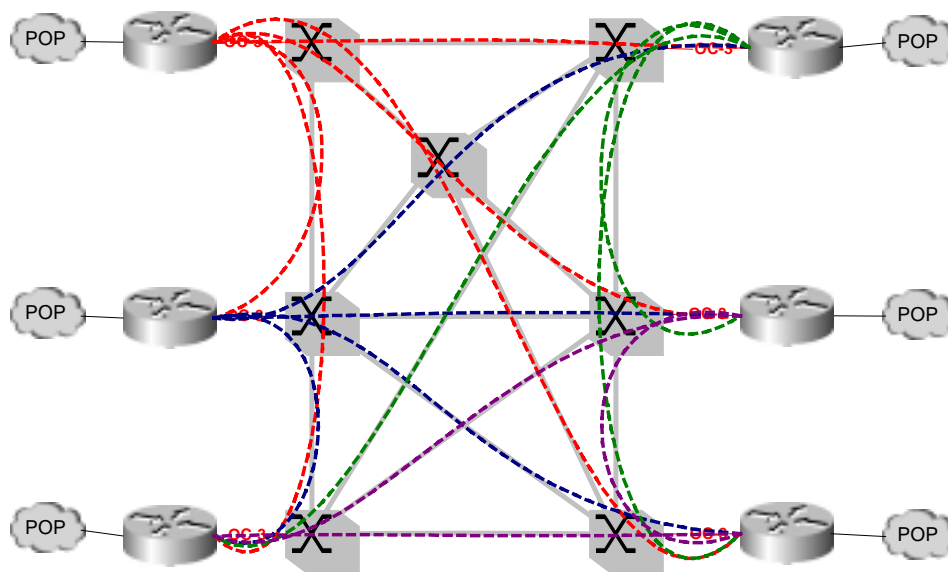


Figure 16 IP Layer 3 Overlay

## Conclusions

The resiliency and adaptability of the Internet is unparalleled in the history of communications. The Internet, with its growing suite of open and standardized protocols, is the clear winner in the inevitable convergence of private line, voice, video, and outsourced data services. MPLS is only the latest entrant in this remarkable evolution. MPLS when combined with traffic engineering deliver a formable tool for meeting the rigid requirements of differentiated services by leveraging the strengths of IP routing, the proven scalability of terabit routers, and the mechanisms for end to end QoS. Avici is a leader in combining the benefits of MPLS and traffic engineering on a next generation terabit routing platform.

## References

- 1.) "A Framework for MPLS", Ross Callon, George Swallow, N. Feldman, A. Viswanathan, P. Doolan, A. Fredette, 09/22/1999. (180569 bytes)
- 2.) "Multiprotocol Label Switching Architecture", Ross Callon, A. Viswanathan, E. Rosen, 08/27/1999. (145481 bytes)
- 3.) "MPLS Label Stack Encoding", Dino Farinacci, Tony Li, A. Conta, Y Rekhter, Dan Tappan, E. Rosen, G. Fedorkow, 09/13/1999. (46971 bytes)
- 4.) "Extensions to RSVP for LSP Tunnels", Der-Hwa Gan, Tony Li, George Swallow, Lou Berger, Vijay Srinivasan, Daniel Awduche, 09/29/1999. (105164 bytes)
- 5.) "MPLS Support of Differentiated Services", Bruce Davie, Pasi Vaananen, Liwen Wu, Francois Le Faucheur, Pierrick Cheval, Ram Krishnan, Shahram Davari, 10/11/1999.
- 6.) "Applicability Statement for Extensions to RSVP for LSP-Tunnels", Alan Hannan, Daniel Awduche, X Xiao, 10/05/1999. (17395 bytes)
- 7.) "MPLS Optimized Multipath (MPLS--OMP)", Curtis Villamizar, February 25, 1999
- 8.) "OSPF Optimized Multipath (OSPF-OMP)", Curtis Villamizar, 02/25/1999. (90622 bytes)
- 9.) "IS-IS Extensions for Traffic Engineering", Henk Smit, Tony Li, May 1999
- 10.) "OSPF Extensions for Traffic Engineering" Derek M. Yeung, February 1999
- 11.) "RSVP Label Allocation for Backup Tunnels", Robert Goguen, George Swallow, October 1999

Copyright © 2000 Avici Systems Inc. - All Rights Reserved

All information in this document is provided for informational use only and is subject to change without notice. Avici Systems Inc. assumes no responsibility for any errors that may be found in this document.

TSR is a registered trademark of Avici Systems Inc.

Avici and IPriori are trademarks of Avici Systems Inc.

All other service marks, trademarks, and registered trademarks mentioned herein are the property of their respective companies.