

## Traffic Engineering for the New Public Network

*Traffic engineering is the prevailing technique for mapping traffic flows reliably and cost-efficiently which enables ISPs to deliver premium services over optimally-utilised bandwidth and with superior flexibility for tailored service offerings. This solution is becoming ever more crucial as ISP networks grow larger and as customer demands become greater.*

*This article describes the benefits of a traffic engineering architecture that uses a combination of Multi Protocol Label Switching (MPLS) forwarding, Interior Gateway Protocol (IGP) extensions, Constrained Shortest Path First (CSPF) path selection, and Resource Reservation Protocol (RSVP) signaling.*

### INTRODUCTION

A fundamental challenge facing Internet Service Providers (ISPs) is enhancing customer satisfaction while scaling their networks to sustain high growth rates. ISPs must deploy a physical topology that meets their customers' needs, while provisioning multiple circuits of varying bandwidths. After deploying the network, ISPs must map customer traffic flows onto the physical topology.

Traffic engineering is the prevailing technique for mapping traffic flows reliably and cost-efficiently which enables ISPs to deliver premium services over optimally-utilised bandwidth and with superior flexibility for tailored service offerings. This solution is becoming ever more crucial as ISP networks grow larger and as customer demands become greater.

This article describes the benefits of a traffic engineering architecture that uses a combination of Multi Protocol Label Switching (MPLS) forwarding, Interior Gateway Protocol (IGP) extensions, Constrained Shortest Path First (CSPF) path selection, and Resource Reservation Protocol (RSVP) signaling.

### TRAFFIC ENGINEERING

IGPs can contribute to network congestion since bandwidth availability and traffic characteristics are not taken into account when forwarding tables are built. Traffic engineering resolves this issue by providing ISPs precise control over traffic flows within their routed domains, thereby enabling them to redirect traffic away from the shortest path and onto a potentially less congested path. This powerful tool can balance and optimise their network traffic load so that links are neither over-utilised nor under-utilised. ISPs can thus exploit the economies of bandwidth provisioned across their entire network.

With the unprecedented growth in demand for network resources, the mission-critical nature of Internet Protocol (IP) applications, and the increasingly competitive nature of the Internet marketplace, traffic engineering has become vital to ISPs' success. Traffic engineering enables ISPs to maximise operational efficiency and minimise operational costs. In so doing, they can provide more options, lower costs, and better ser-

vice to their customers. Following are a few examples of traffic engineering capabilities.

- Routes primary paths around congestion in the network.
- Provides precise control over how traffic is rerouted when a primary path fails.
- Provides efficient use of available aggregate bandwidth and long-haul fibre by optimally using all segments.
- Enhances traffic performance by minimising packet loss, by minimising prolonged periods of congestion, and maximising throughput.
- Enhances statistically bounded performance characteristics (such as loss ratio, delay variation, and transfer delay) that are required to support multiservices offerings.

### DETERMINING THE ARCHITECTURE

In order to determine traffic engineering architecture, solutions used in both traditional routed cores and in IP-over-ATM topologies need to be evaluated.

Traditional routed cores have limited aggregate bandwidth and packet-processing capabilities. IGP route calculations are based on topology and on a simple metric, such as hop count or an administrative value. One limitation to this solution is that metric manipulation is not scalable; as networks become fully meshed or more redundant, metric adjustments in one part of the network often cause unforeseen problems in other parts. Furthermore, since bandwidth availability and traffic characteristics are not distributed, traffic load is not accounted for when forwarding tables are calculated, often resulting in inefficient use of expensive resources, as well as the uneven distribution traffic that might lead to congestion.

IP-over-ATM topologies offer high-speed interfaces, deterministic performance, and traffic engineering through explicitly routed Permanent Virtual Circuits (PVCs) but there are also many limitations. Firstly, the overhead increases significantly since ISPs must manage and operate two different networks: an ATM infrastructure and a logical IP overlay. With separate net-

***AD ENEA OSE***

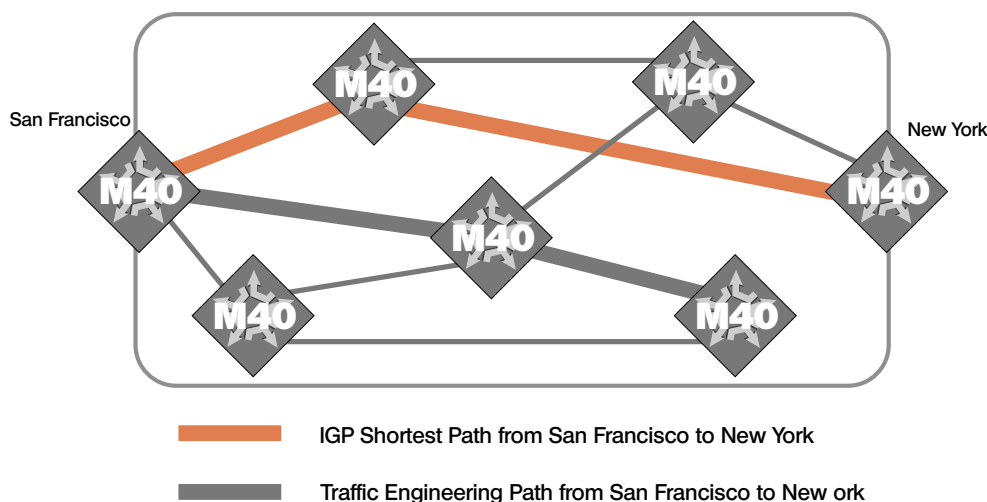


Figure 1. Traffic Engineering Path versus IGP Shortest Path.

works, it is difficult and more expensive to integrate routing functions (executed on the routers) and traffic engineering (executed on the ATM switches). Furthermore, a cell tax is introduced when packet-oriented protocols are carried over an ATM infrastructure. For example, assuming a 20-percent overhead for ATM when accounting for framing and realistic distribution of packet sizes, a 2.488-Gbps OC-48/STM-16 link requires 498 Mbps (almost a full OC-12/STM-4) for the ATM overhead, leaving only 1.99 Gbps available for customer data.

In assessing these models, it can be determined that traffic engineering architecture must achieve the following goals:

- Overcome the limitations of the traditional router cores (scalability issues, unevenly distributed traffic, and under-utilised links).
- Overcome the limitations the IP-over-ATM model (eliminate the extraneous cell tax, and the complexity and expense of co-ordinating and managing two separate networks).
- Provide a level of functionality equivalent to the current IP-over-ATM model (high-speed optical interfaces, deterministic performance, and PVC capabilities).
- Automate the traffic engineering process so that ISPs can provide enhanced customer service and reliability while reducing operational costs.

## IMPLEMENTATION

Traffic engineering implementation needs to combine relatively simple and easily deployable technologies to create a robust solution capable of scaling with the growth of the optical Internet. Companies should also be actively involved in the Internet Engineering Task Force (IETF) MPLS and related working groups, and incorporate existing IETF solutions so that development is relatively simple and unambiguous, and solutions can be implemented with minimal risk.

Traffic engineering strategy should comprise of four

functional components: packet forwarding, information distribution, path selection, and signaling. The architecture should separate each of these individual components with clean interfaces for additional flexibility.

## MPLS PACKET FORWARDING

MPLS directs a flow of IP packets along a pre-determined path known as a Label-Switched Path (LSP). LSPs are similar to ATM PVCs in that traffic flows in one direction from ingress to egress. Duplex traffic requires two LSPs: one LSP to carry traffic in each direction.

An LSP is created by concatenating one or more Label-Switched Routers (LSRs - routers that support MPLS-based forwarding, such as the Juniper M40 Internet Optimised Router). Packets are then forwarded from one LSR to the next across the LSP.

When an ingress LSR receives an IP packet, it assigns a label to the packet, encapsulates the packet in a 4-byte MPLS header, and forwards the packet to the next LSR in the LSP. MPLS provides a tremendous amount of flexibility in the way that an IP packet can be assigned to an LSP. For example, in Juniper Network's traffic engineering implementation, all packets arriving at an interface are forwarded along an LSP, or all packets arriving that are destined to exit at the same egress LSR are forwarded along the same LSP.

At each transit router, label swapping occurs. The transit LSR uses the packet's label as an index into its MPLS forwarding table, thus avoiding the need to evaluate the packet's IP header. Based on the lookup, the LSR replaces the old label with a new one and forwards the packet to the next LSR in the LSP. This process is repeated until the packet reaches the egress LSR.

The egress LSR removes the MPLS header and forwards the packet based on the destination address in the packet's IP header.

A key benefit of MPLS packet forwarding is that the LSP's physical path is not limited to the shortest path that the IGP would choose to reach the destination IP

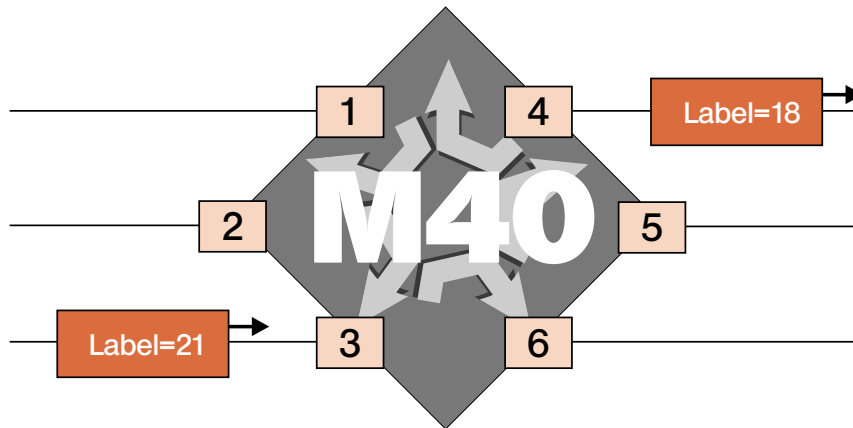


Figure 2. LSR Forwarding a Packet.

address. Another key benefit is that the IP header analysis is performed only once at the ingress LSR, rather than repetitively at each hop in the route. This procedure ensures a single forwarding algorithm can be used for multiple services and traffic types. Therefore, future services can be migrated easily to operate over the common MPLS forwarding infrastructure simply by changing the way that packets are assigned to an LSP. For example, packets could be assigned to an LSP based on a combination of the destination subnetwork and application type, the source and destination subnetworks, a specific Quality of Service (QoS) requirement, an IP multicast group, or a Virtual Private Network (VPN) identifier.

## FLEXIBLE LSP CALCULATION, CONFIGURATION, AND MANAGEMENT

Maintaining a flexible strategy for traffic engineering over MPLS means that a number of different ways to route and configure an LSP are supported. For instance, ISPs can calculate the full path for the LSP offline and statically configure the ingress LSR with the full path. The ingress LSR then uses RSVP as a dynamic signaling protocol to install forwarding state in each LSR along the LSP. They can also configure any number of LSPs as dynamic, hot-standby backups for the primary LSP, as well as constraint-based routing to enhance performance requirements for both the primary and backup LSPs.

Additional flexibility includes the ability for ISPs to modify LSP attributes on the fly, reroute LSPs, and enable re-optimization of an LSP. They can gain access to per-LSP accounting and traffic statistics to use as input to network plans and to identify potential bottlenecks. As well, they can determine whether one LSP can preempt another LSP from a given physical path. Moreover, ISPs can specify

- Partial or complete hops for LSP establishment,
- Multiple and ordered secondary paths for backing up an LSP,
- Policy constraints that LSPs must follow,

- A priority order for LSP establishment, and
- Multiple parallel LSPs for load balancing.

Since LSPs are similar to connection-oriented virtual circuits, fitting LSPs into the existing offline network planning and analysis tools is relatively straightforward. The output of these tools can be converted into the configuration necessary to establish the physical paths for LSPs.

## IGP EXTENSIONS FOR INFORMATION DISTRIBUTION

Since traffic engineering requires detailed knowledge about the network topology, as well as dynamic information about network loading, precise information distribution is critical. This component is easily implemented by defining simple extensions to the IGP so that link attributes are included in each router's LSA. IS-IS extensions are supported by the definition of new Type Length Values (TLVs), while Open Shortest Path First (OSPF) extensions are implemented with Opaque Label Switched Applications (LSAs). The standard flooding algorithm used by the link-state IGP ensures that link attributes are distributed to all routers in the domain.

Each LSR maintains network link attributes and topology information in a specialised Traffic Engineering Database (TED). This database is used exclusively for calculating explicit LSPs across the physical topology, and is maintained so that the traffic engineering computation is independent of the IGP and its link-state database. The IGP continues its operation without change, performing the traditional shortest-path calculation based on information contained in the router's link-state database.

## CSPF PATH SELECTION

After network link attributes and topology information are flooded by the IGP and placed in the TED, each ingress LSR uses the TED to calculate the paths for its own set of LSPs across the routing domain. The ingress LSR determines the physical path for each LSP by applying a CSPF algorithm to the information in the

TED. CSPF is a modified shortest-path-first algorithm that accounts for specific restrictions when calculating the shortest path across the network. Input into the CSPF algorithm includes topology link-state information, network state attributes, and administrative attributes required to support traffic traversing the proposed LSP.

As CSPF considers each candidate node and link for a new LSP, it either accepts or rejects a specific path component based on resource availability or on whether policy constraints are violated. The output is an explicit route consisting of a sequence of LSR addresses that provides the shortest path through the network, meeting the constraints. This explicit route is then passed to the signaling component, which establishes forwarding state in the LSRs along the LSP. The CSPF algorithm is repeated for each LSP that the ingress LSR generates. The explicit route can be either strict or loose, thus enabling the path selection process to be as controlled or as flexible as needed.

To globally optimise traffic engineering for the network, an offline calculation is required. The LSPs can be established in any order because each is installed following the rules for the globally optimised solution.

## RSVP SIGNALING

Signaling is responsible for establishing LSP state and for label distribution. Our implementation uses standard IETF extensions to the RSVP, which is an ideal protocol for establishing LSPs. Since RSVP is the standard resource reservation protocol for the Internet, there is no need to design yet another resource reservation protocol. RSVP is extensible, and is easily added to new functionality with new object types. RSVP offers other advantages, as well.

- RSVP's soft state can reliably establish and maintain LSPs in an MPLS environment.
- RSVP enables network resources to be explicitly reserved and allocated to a given LSP.
- RSVP enables the establishment of explicitly routed LSPs that provide traffic engineering and load balancing capabilities equivalent to those currently provided by ATM and Frame Relay.
- Edge-to-edge RSVP signaling across an MPLS domain scales since the number of LSPs is related to the number of edge LSRs in the domain, rather than to the number of entries in the routing table or the number of end system traffic flows.

## CONCLUSION

For high-performance, high-speed backbones, network growth and customer demands continue to grow at a rapid pace, while customer demands are ever increasing. ISPs planning to migrate to higher speeds should carefully examine the possible alternatives so bandwidth utilisation and operational costs do not constrain the future growth of their networks.

The traditional router cores have scalability issues, uneven traffic distribution, and under-utilised links. IP-over-ATM topologies are expensive and complex, imposing extraneous cell tax and requiring the co-ordi-

nation and management of two separate networks.

Successful traffic engineering solutions should provide ATM traffic management features while eliminating ATM's inefficient bandwidth use, network management and operation complications. Additionally, these solutions need to eliminate the associated scalability problems, thereby enabling ISPs to continue growing their networks to OC-48c/STM-16 speeds and beyond.

Clearly, traffic engineering is the basis for introducing differentiated services in the Internet. This model of tightly controlling traffic and optimizing network resources is extensible to premium services and to incorporating existing legacy network environments into a common traffic-engineered backbone. For example, preferred customers might be promised uncongested links and rapid recovery in case of network transmission failure. Alternatively, existing private line services and voice services could be carried over the Internet backbone. With traffic engineering as the baseline, the Internet backbone can be migrated to a multiservice, highly leveraged infrastructure ■

---

*Alan Taylor is European Technical Director for Juniper Networks. He joined Juniper Networks as a Consulting Engineer in Europe in April 1999, providing technical advice to Juniper's major service provider customers deploying large scale IP backbones. He became European Technical Director for Juniper in June 2000.*

*Prior to working for Juniper, he spent 10.5 years at 3Com in various technical roles. These included a six month spell as an Escalation Engineer for the European market, based at the 3Com HQ in California and 3 years as a Network Design Specialist in the UK and Europe. While acting as Network Design Specialist he played a leading role in several large-scale wide area networking projects in the retail and government sectors. For the last 3.5 years at 3Com, he was the European Product Marketing Manager for Wide Area Networking products.*

*He holds a Masters degree in Electronic and Electrical Engineering from Loughborough University in the UK.*