Network Working Group
Request for Comments: 3346
Category: Informational

J. Boyle
PD Nets
V. Gill
AOL Time Warner, Inc.
A. Hannan
RoutingLoop
D. Cooper
Global Crossing
D. Awduche
Movaz Networks
B. Christian
Worldcom
W.S. Lai
AT&T

August 2002

Applicability Statement for Traffic Engineering with MPLS

Status of this Memo

Copyright Notice

Abstract

This document describes the applicability of Multiprotocol Label Switching (MPLS) to traffic engineering in IP networks. Special considerations for deployment of MPLS for traffic engineering in operational contexts are discussed and the limitations of the MPLS approach to traffic engineering are highlighted.

Table of Contents

1. Introduction

It is generally acknowledged that one of the most significant initial applications of Multiprotocol Label Switching (MPLS) is traffic engineering (TE) [1][2] in IP networks. A significant community of IP service providers have found that traffic engineering of their networks can have tactical and strategic value [2, 3, 4]. To support the traffic engineering application, extensions have been specified for Interior Gateway Protocols (IGP) IS-IS [5] and OSPF [6], and to signaling protocols RSVP [7] and LDP [8]. The extensions for IS-IS, OSPF, and RSVP have all been developed and deployed in large scale in many networks consisting of multi-vendor equipment.

This document discusses the applicability of TE to Internet service provider networks, focusing on the MPLS-based approach. It augments the existing protocol applicability statements and, in particular, relates to the operational applicability of RSVP-TE [9]. Special considerations for deployment of MPLS in operational contexts are discussed and the limitations of this approach to traffic engineering are highlighted.

2. Technical Overview of ISP Traffic Engineering

Traffic engineering (TE) is generally concerned with the performance optimization of operational networks [2]. In contemporary practice, TE means mapping IP traffic flows onto the existing physical network topology in the most effective way to accomplish desired operational objectives. Techniques currently used to accomplish this include, but are not limited to:

1. Manipulation of IGP cost (metrics)
2. Explicit routing using constrained virtual-circuit switching techniques such as ATM or Frame Relay SPVCs
3. Explicit routing using constrained path setup techniques such as MPLS

This document is concerned primarily with MPLS techniques. Specifically, it deals with the ability to use paths other than the shortest paths selected by the IGP to achieve a more balanced network utilization, e.g., by moving traffic away from IGP-selected shortest paths onto alternate paths to avoid congestion in the network. This can be achieved by using explicitly signaled LSP-tunnels. The explicit routes to be used may be computed offline and subsequently downloaded and configured on the routers using an appropriate mechanism. Alternatively, the desired characteristics of an LSP (such as endpoints, bandwidth, affinities) may be configured on a router, which will then use an appropriate algorithm to compute a path through the network satisfying the desired characteristics, subject to various types of constraints. Generally, the characteristics associated with LSPs may include:

- Ingress and egress nodes
- Bandwidth required
- Priority
- Nodes to include or exclude in the path
- Affinities to include or exclude in the path
- Resilience requirements

Affinities are arbitrary, provider-assigned, attributes applied to links and carried in the TE extensions for the IGPs. Affinities impose a class structure on links, which allow different links to be logically grouped together. They can be used to implement various types of policies, or route preferences that allow the inclusion or exclusion of groups of links from the path of LSPs. Affinities are unique to MPLS and the original requirement for them was documented in [2].

3. Applicability of Internet Traffic Engineering

As mentioned in [2] and [7], traffic engineering with MPLS is appropriate to establish and maintain explicitly routed paths in an IP network for effective traffic placement. LSP-tunnels can be used to forward subsets of traffic through paths that are independent of routes computed by conventional IGP Shortest Path First (SPF) algorithms. This gives network operators significant flexibility in controlling the paths of traffic flows across their networks and allows policies to be implemented that can result in the performance optimization of networks. Examples of scenarios where MPLS-based TE capabilities are applicable in service provider environments are given below. The applicability of MPLS is certainly not restricted to these scenarios.

3.1 Avoidance of Congested Resources

In order to lower the utilization of congested link(s), an operator may utilize TE methods to route a subset of traffic away from those links onto less congested topological elements. These types of techniques are viable when segments of the network are congested while other parts are underutilized.

Operators who do not make extensive use of LSP-tunnels may adopt a tactical approach to MPLS TE in which they create LSP-tunnels only when necessary to address specific congestion problems. For example, an LSP can be created between two nodes (source and destination) that are known to contribute traffic to a congested network element, and explicitly route the LSP through a separate path to divert some traffic away from the congestion. On the other hand, operators who make extensive use of LSP-tunnels, either for measurement or automated traffic control, may decide to explicitly route a subset of the LSPs that traverse the point of congestion onto alternate paths. This can be employed to respond quickly when the bandwidth parameter associated with the LSPs does not accurately represent the actual traffic carried by the LSPs, and the operator determines that changing the bandwidth parameter values might not be effective in addressing the issue or may not have lasting impact.

There are other approaches that measure traffic workloads on LSPs and utilize these empirical statistics to configure various characteristics of LSPs. These approaches, for example, can utilize the derived statistics to configure explicit routes for LSPs (also known as offline TE [10]). They can also utilize the statistics to set the values of various LSP attributes such as bandwidths, priority, and affinities (online TE). All of these approaches can be used both tactically and strategically to react to periods of congestion in a network. Congestion may occur as a result of many

factors: equipment or facility failure, longer than expected provisioning cycles for new circuits, and unexpected surges in traffic demand.

3.2 Resource Utilization in Network Topologies with Parallel Links

In practice, many service provider networks contain multiple parallel links between nodes. An example is transoceanic connectivity which is often provisioned as numerous low-capacity circuits, such as NxDS-3 (N parallel DS-3 circuits) and NxSTM-1 (N parallel STM-1 circuits). Parallel circuits also occur quite often in bandwidth-constrained cities. MPLS TE methods can be applied to effectively distribute the aggregate traffic workload across these parallel circuits.

MPLS-based approaches commonly used in practice to deal with parallel links include using LSP bandwidth parameters to control the proportion of demand traversing each link, explicitly configuring routes for LSP-tunnels to distribute them across the parallel links, and using affinities to map different LSPs onto different links. These types of solutions are also applicable in networks with parallel and replicated topologies, such as an NxOC-3/12/48 topology.

3.3 Implementing Routing Policies using Affinities

It is sometimes desirable to restrict certain types of traffic to certain types of links, or to explicitly exclude certain types of links in the paths for some types of traffic. This might be needed to accomplish some business policy or network engineering objectives. MPLS resource affinities provide a powerful mechanism to implement these types of objectives.

As a concrete example, suppose a global service provider has a flat (non-hierarchical) IGP. MPLS TE affinities can be used to explicitly keep continental traffic (traffic originating and terminating within a continent) from traversing transoceanic resources.

Another example of using MPLS TE affinities to exclude certain traffic from a subset of circuits might be to keep inter-regional LSPs off of circuits that are reserved for intra-regional traffic.

Still another example is the situation in a heterogeneous network consisting of links with different capacities, e.g., OC-12, OC-48, and OC-192. In such networks, affinities can be used to force some types of traffic to only traverse links with a given capacity, e.g. OC-48.

3.4 Re-optimization After Restoration

After the occurrence of a network failure, it may be desirable to calculate a new set paths for LSPs to optimizes performance over the residual topology. This re-optimization is complementary to the fast-reroute operation used to reduce packet losses during routing transients under network restoration. Traffic protection can also be accomplished by associating a primary LSP with a set of secondary LSPs, hot-standby LSPs, or a combination thereof [11].

4. Implementation Considerations

4.1 Architectural and Operational Considerations

When deploying TE solutions in a service provider environment, the impact of administrative policies and the selection of nodes that will serve as endpoints for LSP-tunnels should be carefully considered. As noted in [9], when devising a virtual topology for LSP-tunnels, special consideration should be given to the tradeoff between the operational complexity associated with a large number of LSP-tunnels and the control granularity that large numbers of LSP-tunnels allow. In other words, a large number of LSP-tunnels allow greater control over the distribution of traffic across the network, but increases network operational complexity. In large networks, it may be advisable to start with a simple LSP-tunnel virtual topology and then introduce additional complexity based on observed or anticipated traffic flow patterns [9].

Administrative policies should guide the amount of bandwidth to be allocated to an LSP. One may choose to set the bandwidth of a particular LSP to a statistic of the measured observed utilization over an interval of time, e.g., peak rate, or a particular percentile or quartile of the observed utilization. Sufficient over-subscription (of LSPs) or under-reporting bandwidth on the physical links should be used to account for flows that exceed their normal limits on an event-driven basis. Flows should be monitored for trends that indicate when the bandwidth parameter of an LSP should be resized. Flows should be monitored constantly to detect unusual variance from expected levels. If an unpoliced flow greatly exceeds its assigned bandwidth, action should be taken to determine the root cause and remedy the problem. Traffic policing is an option that may be applied to deal with congestion problems, especially when some flows exceed their bandwidth parameters and interfere with other compliant flows. However, it is usually more prudent to apply policing actions at the edge of the network rather than within the core, unless under exceptional circumstances.

When creating LSPs, a hierarchical network approach may be used to alleviate scalability problems associated with flat full mesh virtual topologies. In general, operational experience has shown that very large flows (between city pairs) are long-lived and have stable characteristics, while smaller flows (edge to edge) are more dynamic and have more fluctuating statistical characteristics. A hierarchical architecture can be devised consisting of core and edge networks in which the core is traffic engineered and serves as an aggregation and transit infrastructure for edge traffic.

However, over-aggregation of flows can result in a stream so large that it precludes the constraint-based routing algorithm from finding a feasible path through a network. Splitting a flow by using two or more parallel LSPs and distributing the traffic across the LSPs can solve this problem, at the expense of introducing more state in the network.

Failure scenarios should also be addressed when splitting a stream of traffic over several links. It is of little value to establish a finely balanced set of flows over a set of links only to find that upon link failure the balance reacts poorly, or does not revert to the original situation upon restoration.

4.2 Network Management Aspects

Networks planning to deploy MPLS for traffic engineering must consider network management aspects, particularly performance and fault management [12]. With the deployment of MPLS in any infrastructure, some additional operational tasks are required, such as constant monitoring to ensure that the performance of the network is not impacted in the end-to-end delivery of traffic. In addition, traffic characteristics, such as latency across an LSP, may also need to be assessed on a regular basis to ensure that service-level guarantees are achieved.

Obtaining information on LSP behavior is critical in determining the stability of an MPLS network. When LSPs transition or path changes occur, packets may be dropped which impacts network performance. It should be the goal of any network deploying MPLS to minimize the volatility of LSPs and reduce the root causes that induce this instability. Unfortunately, there are very few, if any, NMS systems that are available at this time with the capability to provide the correct level of management support, particularly root cause analysis. Consequently, most early adopters of MPLS develop their own management systems in-house for the MPLS domain. The lack of availability of commercial network management systems that deal specifically with MPLS-related aspects is a significant impediment to the large-scale deployment of MPLS networks.

The performance of an MPLS network is also dependent on the configured values of bandwidth for each LSP. Since congestion is a common cause of performance degradation in operational networks, it is important to proactively avoid these situations. While MPLS was designed to minimize congestion on links by utilizing bandwidth reservations, it is still heavily reliant on user configurable data. If the LSP bandwidth value does not properly represent the traffic demand of that LSP, over-utilization may occur and cause significant congestion within the network. Therefore, it is important to develop, deploy, and maintain a good modeling tool for determining LSP bandwidth size. Lack of this capability may result in sub-optimal network performance.

4.3 Capacity Engineering Aspects

Traffic engineering has a goal of ensuring traffic performance objectives for different services. This requires that the different network elements be dimensioned properly to handle the expected load. More specifically, in mapping given user demands onto network resources, network dimensioning involves the sizing of the network elements, such as links, processors, and buffers, so that performance objectives can be met at minimum cost. Major inputs to the dimensioning process are cost models, characterization of user demands and specification of performance objectives.

In using MPLS, dimensioning involves the assignment of resources such as bandwidth to a set of pre-selected LSPs for carrying traffic, and mapping the logical network of LSPs onto a physical network of links with capacity constraints. The dimensioning process also determines the link capacity parameters or thresholds associated with the use of some bandwidth reservation scheme for service protection. Service protection controls the QoS for certain service types by restricting access to bandwidth, or by giving priority access to one type of traffic over another. Such methods are essential, e.g., to prevent starvation of low-priority flows, to guarantee a minimum amount of resources for flows with expected short duration, to improve the acceptance probability for flows with high bandwidth requirements, or to maintain network stability by preventing performance degradation in case of a local overload.

4.4 Network Measurement Aspects

Network measurement entails robust statistics collection and systems development. Knowing *what* to do with these measurements is often where the secret-sauce is. Examples for different applications of measurements are described in [13]. For instance, to ensure that the QoS objectives have been met, performance measurements and performance monitoring are required so that real-time traffic control

actions, or policy-based actions, can be taken. Also, to characterize the traffic demands, traffic measurements are used to estimate the offered loads from different service classes and to provide forecasting of future demands for capacity planning purposes. Forecasting and planning may result in capacity augmentation or may lead to the introduction of new technology and architecture.

To avoid QoS degradation at the packet level, measurement-based admission control can be employed by using online measurements of actual usage. This is a form of preventive control to ensure that the QoS requirements of different service classes can be met simultaneously, while maintaining network efficiency at a high level. However, it requires proper network dimensioning to keep the probability for the refusal of connection/flow requests sufficiently low.

5. Limitations

Significant resources can be expended to gain a proper understanding of how MPLS works. Furthermore, significant engineering and testing resources may need to be invested to identify problems with vendor implementations of MPLS. Initial deployment of MPLS software and the configurations management aspects to support TE can consume significant engineering, operations, and system development resources. Developing automated systems to create router configurations for network elements can require significant software development and hardware resources. Getting to a point where configurations for routers are updated in an automated fashion can be a time consuming process. Tracking manual tweaks to router configurations, or problems associated with these can be an endless task. What this means is that much more is required in the form of various types of tools to simplify and automate the MPLS TE function.

Certain architectural choices can lead to operational, protocol, and router implementation scalability problems. This is especially true as the number of LSP-tunnels or router configuration data in a network increases, which can be exacerbated by designs incorporating full meshes, which create O(N^2) number of LSPs, where N is the number of network-edge nodes. In these cases, minimizing N through hierarchy, regionalization, or proper selection of tunnel termination points can affect the network’s ability to scale. Loss of scale in this sense can be via protocol instability, inability to change network configurations to accommodate growth, inability for router implementations to be updated, hold or properly process configurations, or loss of ability to adequately manage the network.

Although widely deployed, MPLS TE is a new technology when compared to the classic IP routing protocols such as IS-IS, OSPF, and BGP. MPLS TE also has more configuration and protocol options. As such, some implementations are not battle-hardened and automated testing of various configurations is difficult if not infeasible. Multi-vendor environments are beginning to appear, although additional effort is usually required to ensure full interoperability.

Common approaches to TE in service provider environments switch the forwarding paradigm from connectionless to connection oriented. Thus, operational analysis of the network may be complicated in some regards (and improved in others). Inconsistencies in forwarding state result in dropped packets whereas with connectionless methods the packet will either loop and drop, or be misdirected onto another branch in the routing tree.

Currently deployed MPLS TE approaches can be adversely affected by both internal and external router and link failures. This can create a mismatch between the signaled capacity and the traffic an LSP-tunnel carries.

Many routers in service provider environments are already under stress processing the software workload associated with running IGP, BGP, and IPC. Enabling TE in an MPLS environment involves adding traffic engineering databases and processes, adding additional information to be carried by the routing processes, and adding signaling state and processing to these network elements. Additional traffic measurements may also need to be supported. In some environments, this additional load may not be feasible.

MPLS in general and MPLS-TE in particular is not a panacea for lack of network capacity, or lack of proper capacity planning and provisioning in the network dimensioning process. MPLS-TE may cause network traffic to traverse greater distances or to take paths with more network elements, thereby incurring greater latency. Generally, this added inefficiency is done to prevent shortcomings in capacity planning or available resources path to avoid hot spots. The ability of TE to accommodate more traffic on a given topology can also be characterized as a short-term gain during periods of persistent traffic growth. These approaches cannot achieve impossible mappings of traffic onto topologies. Failure to properly capacity plan and execute will lead to congestion, no matter what technology aids are employed.

6. Conclusion

The applicability of traffic engineering in Internet service provider environments has been discussed in this document. The focus has been on the use of MPLS-based approaches to achieve traffic engineering in this context. The applicability of traffic engineering and associated management and deployment considerations have been described, and the limitations highlighted.

MPLS combines the ability to monitor point-to-point traffic statistics between two routers and the capability to control the forwarding paths of subsets of traffic through a given network topology. This makes traffic engineering with MPLS applicable and useful for improving network performance by effectively mapping traffic flows onto links within service provider networks. Tools that simplify and automate the MPLS TE functions and activation help to realize the full potential.

7. Security Considerations

This document does not introduce new security issues. When deployed in service provider networks, it is mandatory to ensure that only authorized entities are permitted to initiate establishment of LSP-tunnels.

8. References

1 Rosen, E., Viswanathan, A. and R. Callon, "Multiprotocol Label Switching Architecture," RFC 3031, January 2001.

2 Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M. and J. McManus, "Requirements for Traffic Engineering Over MPLS," RFC 2702, September 1999.

3 X. Xiao, A. Hannan, B. Bailey, and L. Ni, "Traffic Engineering with MPLS in the Internet," IEEE Network, March/April 2000.

4 V. Springer, C. Pierantozzi, and J. Boyle, "Level3 MPLS Protocol Architecture," Work in Progress.

5 T. Li, and H. Smit, "IS-IS Extensions for Traffic Engineering," Work in Progress.

6 D. Katz, D. Yeung, and K. Kompella, "Traffic Engineering Extensions to OSPF," Work in Progress.

7 Awduche, D., Berger, L., Gan, D.H., Li, T., Srinivasan, V. and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," RFC 3209, December 2001.

8 Jamoussi, B. (Editor), "Constraint-Based LSP Setup using LDP," RFC 3212, January 2002.

9 Awduche, D., Hannan, A. and X. Xiao, "Applicability Statement for Extensions to RSVP for LSP-Tunnels," RFC 3210, December 2001.

10 Awduche, D., Chiu, A., Elwalid, A., Widjaja, I. and X. Xiao, "Overview and Principles of Internet Traffic Engineering", RFC 3272, May 2002.

11 W.S. Lai, D. McDysan, J. Boyle, M. Carlzon, R. Coltun, T. Griffin, E. Kern, and T. Reddington, "Network Hierarchy and Multilayer Survivability," Work in Progress.

12 D. Awduche, "MPLS and Traffic Engineering in IP Networks," IEEE Communications Magazine, December 1999.

13 W.S. Lai, B. Christian, R.W. Tibbs, and S. Van den Berghe, "A Framework for Internet Traffic Engineering Measurement," Work in Progress.

9. Acknowledgments

The effectiveness of the MPLS protocols for traffic engineering in service provider networks is in large part due to the experience gained and foresight given by network engineers and developers familiar with traffic engineering with ATM in these environments. In particular, the authors wish to acknowledge the authors of RFC 2702 for the clear articulation of the requirements, as well as the developers and testers of code in deployment today for keeping their focus.

10. Authors' Addresses

Jim Boyle
Protocol Driven Networks
Tel: +1 919-852-5160
EMail: jboyle@pdnets.com

Vijay Gill
AOL Time Warner, Inc.
12100 Sunrise Valley Drive
Reston, VA 20191
EMail: vijay@umbc.edu

Alan Hannan
RoutingLoop Intergalactic
112 Falkirk Court
Sunnyvale, CA 94087, USA
Tel: +1 408-666-2326
EMail: alan@routingloop.com

Dave Cooper
Global Crossing
960 Hamlin Court
Sunnyvale, CA 94089, USA
Tel: +1 916-415-0437
EMail: dcooper@gblx.net

Daniel O. Awduche
Movaz Networks
7926 Jones Branch Drive, Suite 615
McLean, VA 22102, USA
Tel: +1 703-298-5291
EMail: awduche@movaz.com

Blaine Christian
Worldcom
22001 Loudoun County Parkway, Room D1-2-737
Ashburn, VA 20147, USA
Tel: +1 703-886-4425
EMail: blaine@uu.net

Wai Sum Lai
AT&T
200 Laurel Avenue
Middletown, NJ 07748, USA
Tel: +1 732-420-3712
EMail: wlai@att.com

11. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.