Brocade Products Solution Series

MPLS Traffic Engineering & RSVP Technology and Design Considerations

Larry Stinson, Technical Publications

Sept. 2012

BROCADE

Contents

Overview

The goal of this document is to provide a technology background for the various components of MPLS TE with RSVP, using functional case studies and discussion of design considerations for each stage of configuration that can be applied in the planning and deployment of MPLS traffic engineering in the field.

Introduction

Basic MPLS packet forwarding relies on an interior gateway protocol e.g. OSPF or ISIS to calculate routes which are in turn based on the shortest path first (SPF) algorithm. The standard IGP algorithm calculates the shortest path through the network based on sum of the link costs to reach the destination. While the IGP may select the shortest path to reach a destination, it does not take into account the amount of traffic being forwarded across a given link. Forwarding all traffic over the shortest path can lead to oversubscription on some links and under utilization of alternate (longer higher IGP cost) paths through the network.

IGP’s do have a capacity for load balancing traffic across equal cost paths (ECMP) however, there is an upper limit to the number of ECMP paths that are typically supported. Furthermore, equal cost paths may not provide the node redundancy needed to meet the high availability requirements of service provider core networks. In short, while the IGP SPF algorithm is efficient at best path selection, it does not provide flexibility for users to route traffic over any or all links in a given network results in inefficient use of network resources.

MPLS Traffic Engineering

The key concept of MPLS traffic engineering is the ability to steer traffic over any path in the network and not be limited by the SPF route to a given destination. From a network design and operation perspective, MPLS traffic engineering provides the capability for:

- Defining user based paths through the network
- Automatic LSP setup through protocol signaling
- Ensuring adequate bandwidth for protected paths
- Re-routing traffic in the event of a network failure
- Bandwidth usage monitoring
- Prioritization of bandwidth resources with the ability to pre-empt lower priority LSPs

The concepts and design considerations for each of the above are discussed at a high level in the following sections and in more detail with examples in the following chapters.

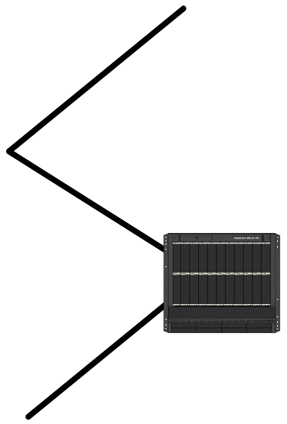
MPLS Traffic Engineering Components

There are multiple components that enable the functions of MPLS traffic engineering. At a high level, these components can be broken down as follows:

- **IGP Traffic Engineering Extensions**

  The tradiational role of an IGP is to find the best route to a given destination, and route selection is based on the information available from the IGP database. This routing database stores information about all links in the network where the routing protocol is enabled. Conceptually, it is similar to the underlying operation of a navigation system used for driving directions. A typical navigation system has knowledge of all roads in the area and provides the best route to use when queried, similarly, an IGP database has knowledge of all paths through a given network and provides the shortest path to a network destination when queried. In the case of traffic engineering, to allow for more options in route selection the IGP databases for OSPF and ISIS have been enhanced with traffic engineering extensions to create a Traffic Engineering Database (TED). Simply stated, these extensions allow more contraints to be applied to the path selection i.e. more options are available rather than always selecting the shortest path. Continuing the navigation system analogy from above, a full featured navigations system allows users the option to avoid highways or pass through a town on the way to their destination. Similarly, with an IGP TED, network operators can also specify route constraints e.g. that traffic should not travel over a given link, and must pass through a given node in the network when traveling to a destination. Path selection with user specified constraints is achieved by running a modified version of the IGP SPF algorithm that takes into account additional parameters that will be discussed in upcoming chapters such as logical bandwidth, link color, strict or loose path definitions etc. This modified version of the SPF algorithm is referred to as Constrained Shortest Path First (CSPF). CSPF calculates the shortest path from a source to a destination that meets all of the user specified constraints.

Referring to the network diagram below, for traffic flow from PE1 to PE3, the shortest path is via P2 with a cost of 20. If all traffic, were to pass over this path the alternate path through P1 would be underutilized and the link from PE1 to P1 could potentially be over utilized.

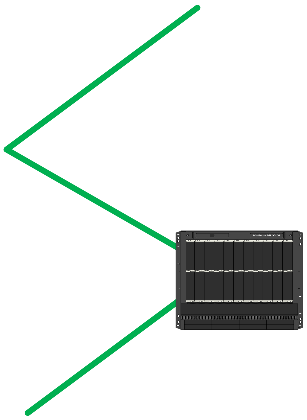


CSPF Constraints: Administrative Groups

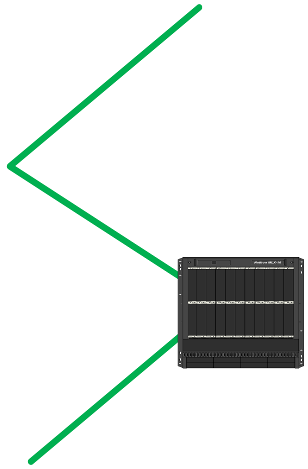
When defining path constraints to reach a destination, there are multiple options that can be used individually or in combination with each other. Administrative groups are one of the constraints that a user can specify for a path.

The concept of an administrative group is a user defined name for referencing a link or group of links in a network by a simple common name. Assigning links to administrative groups is frequently referred to as link coloring, as the group names often include color names e.g. red, green, gold, silver etc. The uniqueness of administrative groups is that their definition and assignment to links is populated throughout the network via IGP traffic extensions. That is, every node in the network has the administrative group link assignments in its TED allowing them to be referenced by name from any node.

Referring to the network diagram below, the following links are assigned to administrative group "green" PE1-P1, P1-P2, and P2-PE3 while the link from PE1 to P2 was assigned to administrative group red. After assignment of these administrative groups, a user can specify that the path from PE1 to PE3 should only contain links from administrative group green. In this case, the CSPF algorithm would calculate the shortest path to PE3 from PE1 using only links belonging to admin-group green and would select the best path to be: PE1 – P1 – P2 – PE3.



is also able to allocate the requested resources then it replies to the path message with a reservation message containing the MPLS label. Reservation messages travel upstream, and in the example above the reservation message would be initiated from PE3 and would travel along the reverse path PE3-P2-P1-PE1. Each node that receives the reservation message extracts the label information, and inserts its own label to be used by the upstream router and forwards the modified reservation message to the next upstream router. Reservation messages provide two key functions: a) MPLS label allocation along the path and b) confirmation that the requested resources have been allocated. Once the reservation message is received at the source (PE1), the LSP is brought up. The diagram below illustrates the sequential high level signaling that takes place for LSP setup.


NetIron MLX-16

1 IGP Traffic Engineering Extensions

1.1 Introduction

Standard IGP routing protocols do not have the capability to adverstise and maintain key information (e.g. reserverable bandwidth, unreserved bandwidth, administrative groups etc.) needed make traffic engineering path decisions. Both OSPF and ISIS have protocol extensions to overcome these shortcomings and enable creation of a Traffic Engineering Database (TED) that stores the required information for making path decisions based on user constraints and bandwidth.

1.2 Technology Background

Traffic Engineering relies on a Traffic Engineering Database (TED) which contains topology information about nodes in an MPLS domain and the links that connect them. This topology information is obtained from either the OSPF or IS-IS with traffic engineering extensions. An LSR, when configured to do so, floods OSPF-TE LSAs or IS-IS LSPs with TE extensions for its MPLS-enabled interfaces to its neighboring routers in the OSPF or IS-IS area. Other LSRs store the information from the OSPF-TE LSAs or IS-IS LSPs with TE extensions in their own Traffic Engineering Databases, allowing each LSR in the area to maintain an identical TED describing the MPLS topology. The topology information in the TED is used by the CSPF process when it calculates traffic-engineered paths for signaled LSPs. The specific IGP extensions are as follows:

- OSPF opaque LSA 10 (Area-scope flooding) two top-level TLVs defined:
  1. Router Address TLV: Router Address TLV specifies a stable IP address of the advertising router
  2. Link TLV: multiple sub-tlv’s summarized below

- ISIS TLV 22 (Extended IS reachability): multiple sub-tlv’s summarized below

- ISIS TLV 135 (Extended IP reachability – primarily for QoS implementation outside the scope of this discussion)

The information exchanged by both OSPF and ISIS for building the TED is very similar and includes the following key items:

- Interface and neighbor interface addresses
- Maximum reservable bandwidth per network link
- Current reserable bandwidth
- Traffic engineering metric
- Administrative group information (affinity classes / colors)

For additional information and comparison, the sub-tlv types for ISIS and OSPF TE extensions are summarized below in tables 1-1 and 1-2 respectively.

Table 1.1-1: ISIS TLV 22 Extended IS Reachability sub-TLV types

Sub-TLV type	Length (octets)	Name	Description
3	4	Administrative group (color)	Made up of a 4-octet bit map (user assigned) each bit corresponds to one administrative group note: interfaces can be part of more than one admin group therefore multiple bits can be set at the same time
6	4	IPv4 interface address	Made up of a 4-octet IPv4 address for the interface.
8	4	IPv4 neighbor address	Single IPv4 address for the neighbor router connected via the link the TLV is sent on
9	4	Maximum link bandwidth	Maximum link bandwidth that can be used on this link (in the direction the TLV was received in). 32 bits in IEEE floating point format in bytes/s.
10	4	Maximum reservable link bandwidth	Total maximum amount of bandwidth that can be reserved on the link in the direction the TLV was received in. TE allows this value to be greater than