|  |  |
| --- | --- |
|  |  |

**Technical report: deployment of MPLS over MB-NG**

**-25 June 2004-**

|  |  |  |
| --- | --- | --- |
|  | Document identifier: | Technical report: deployment of MPLS over MB-NG |
|  | Relevant Task | Task 7, 9 |
|  | Date: | 18/03/2022 |
|  | Version: | **Version 1.0** |
|  | Document status: | **Draft** |
|  | Description | **Description on how MPLS Traffic Engineering has been configured in the MB-NG network to test bandwidth allocation capabilities** |
|  | Editors: | **N. Pezzi** |
|  | Change Record: |  |
|  |  |  |

**Introduction**

One of the goals of the Managed Bandwidth Next Generation project was the deployment of Multi Protocol Label Switching over the SuperJanet development network in order to test some of its Traffic Engineering (TE) capabilities. The main focus of these tests was to prove what can and what cannot be done in terms of bandwidth allocation and traffic prioritisation in presence of congestion with the above mentioned features. Traffic engineering was originally thought to integrate layer 3 capabilities with MPLS in order to optimize the routing of IP traffic creating the so-called “constraint-based routing” in which the path for a traffic flow is the shortest that meets some resource requirements rather than shortest in terms of hop counts as found in traditional IP networks.

The picture below shows the MB-NG network. The core is composed of Cisco 12000GSRs (Gigabit switch router) and each edge site has two Cisco 7600OSRs (Optical switch router). Because of initial hardware limitation, MPLS-TE was not available on the OSR boxes and it was then restricted to the GSRs in the core (although there are new Supervisor linecards that would allow for the whole network to be MPLS-TE enabled).

**How to enable MPLS-TE**

The main feature of MPLS-TE is a tunnel: a path which is predetermined between two end-points and that is going to be followed by each packet belonging to the same flow. A tunnel can be seen as an “emulated leased line” between two end points. Since the MB-NG network is composed of Cisco routers only, the following description is valid for their boxes only, different manufacturer can implement different protocols with different capabilities and/or proprietary extensions.

In order to create tunnels between end points many different protocols need to be enabled in the network:

* MPLS with TE extensions;
* A link-state Interior Gateway Protocol with TE extensions;
* RSVP (Resource Reservation protocol) with TE extensions.

Starting from the IGP, there are currently two available with TE extensions: ISIS (Intermediate system to intermediate system) and OSPF(Open shortest path first). The reason why a link-state protocol has to be used is because, as previously stated, the path followed by the tunnel is calculated at the head end router so this must have a complete map of all the network resources (i.e. which router/links are MPLS enabled and how much bandwidth is available for RSVP to reserve).

RSVP is used as LDP (label distribution protocol) to carry information about resource availability on each link. This is because the amount of bandwidth that can be allocated on each link does not need to be the total available and in fact a smaller amount can be chosen; for example on an OC48 link, only 600Mbit/sec can be allocated to RSVP rather than 2.5Gbit/sec.

To enable MPLS-TE on Cisco routers, IP CEF (Cisco Express Forwarding) has to be enabled first, a loopback interface with an IP mask of 32 bits need to be created and MPLS-TE has to be enabled globally and then on each router interface involved with the MPLS setup. The loopback interface is used for the setup of the MPLS network and TE by the routing protocol. This loopback address must be reachable via the global routing table and it is mandatory because unlike an IP address assigned to an interface it is supposed to be reachable all the time(if an interface goes down its address disappears from the routing table while a loopback interface cannot go down unless the administrator does it explicitly).

Once all of the above protocols have been properly configured on each router and in turn on each interface that is involved in the MPLS cloud, a tunnel can be created. From the point of view of IOS, a tunnel is seen as an interface but it has to be pointed out that it is a unidirectional one so in some case a tunnel on the reverse path between the same end points might need to be created.

Some of the most important parameters to choose when creating a tunnel are the tunnel destination, the tunnel bandwidth and the type of path desired. The latter can be either automatic (i.e. calculated by the router) or manual (as specified by the user). The tunnel bandwidth is used by RSVP to make sure that enough resources are available in the network before allocating the tunnel.

Finally when the IGP has established the complete topology of the MPLS network, a path is chosen (if automatic) and the tunnel is brought up. If there are not enough resources or a path cannot be found, the tunnel interface will be shown as down.

**Bandwidth allocation and QoS**

Regarding the managed bandwidth service (i.e. a guaranteed bandwidth allocation for a particular flow) it is important to clarify that there is a difference between forwarding plane and control plane. MPLS-TE will “guarantee” bandwidth on the control plane by using RSVP signalling to create the tunnel path. If a new tunnel is created but there is currently not a valid RSVP path with enough bandwidth available to satisfy the tunnel, the tunnel will not become active. This feature can be used to give admission control and to make sure that not too many people are granted access to a prioritized service in the network.

Unfortunately, by default there is no correlation between the MPLS TE control plane and the forwarding plane and without this correlation, there is no guaranteed bandwidth in the network.

To implement a guaranteed bandwidth service, the correct associations between the control plane and the forwarding plane need to be created.

In particular:

1. Traffic entering each tunnel interface should be policed to the tunnels configured bandwidth;
2. Traffic entering each tunnel interface should be marked to a unique IP precedence (or DSCP if DiffServ is being used) to indicate that it is tunnel traffic; it must also be true that all other traffic entering the network (non tunnel traffic) must not be marked with this unique value;
3. Different queuing strategies should be configured on each interface that belongs to the MPLS cloud.

The first point is necessary to make sure that users that are entitled to use tunnels do not send more than they are supposed to which could result in a service with a perceived lower quality. A particular class does not automatically go into the tunnel since there is no way of mapping a particular QoS class to the tunnel itself so a mechanism to force the tunnel traffic to be “directed” inside the tunnel is required (for example Policy Based Routing or a static route).

The second point is required in order to distinguish which traffic has to be treated in a “special” way. It is important to know that the MPLS header has an Experimental field (called EXP) composed of 3 bits which is currently being used for QoS purposes. If the router is marking the QoS field in the IP header, by default IOS copies the three most significant bits of the DSCP or the IP precedence of the IP packet to the EXP field in the MPLS header. This action happens when the MPLS header is initially imposed on the IP packet.

 Another option is to define a mapping between the DSCP or IP precedence and the EXP bits. Either way QoS setting in the rest of the network can look at either the EXP field or the QoS field in the IP header or both.The third point is the one that actually gives the priority to the tunnel traffic at each node; it can be implemented in different ways (priority queuing, MDRR or any scheduling algorithm implemented in the router).

Obviously in the case of an over provisioned network where congestion never occurs, the third point is not necessary. In case the tunnel users are interested in other parameter like jitter then a queuing mechanism needs to be in place everywhere in the network even if the network is overprovisioned.

**Configuration samples**

The following configuration samples are taken from the GSRs in the core of the MB-NG network and show some of the steps described above. When configuring all the protocols needed for MPLS-TE and described, many parameters need to be chosen so for a more detailed description follow the links at the end of the document.

>on all routers in configuration mode

mpls traffic-eng tunnels

>on each core interface (Y is the amount of bandwidth that RSVP is going to reserve on that interface)

mpls traffic-eng tunnels

ip rsvp bandwidth Y

>on all routers there has to be an IGP with TE extensions, below a sample from a router with ISIS

router isis

 metric-style wide

 mpls traffic-eng router-id loopback0

 mpls traffic-eng level-2

>tunnel head end router (i.e. where the tunnel starts). X is the IP address of the destination router.

interface Tunnel0

 ip unnumbered Loopback0

 no ip directed-broadcast

 mpls traffic-eng tunnels

 tunnel destination X

 tunnel mode mpls traffic-eng

 tunnel mpls traffic-eng priority 3 1

 tunnel mpls traffic-eng bandwidth Z

 tunnel mpls traffic-eng path-option 1 dynamic

The most interesting lines in the above configuration are the last three.

The first one specifies two priorities: the former is the setup priority and the latter is the hold priority, by specifying different values it is possible to give different priorities to different tunnels that could start from the same router. When trying to create a tunnel in addition to existing ones, if there is not enough bandwidth the new tunnel is brought up only if it has a bigger setup priority than the hold priority of the existing ones.

The second line from the bottom specifies the bandwidth that has to be reserved for the tunnel (value Z), as previously explained this affects the control plane only and requires a matching QoS configuration to affect the forwarding plane as well.

Finally the last line specifies the kind of path that has to be used, there can be more than one option and the order in which they have to be tried is represented by the number. The last parameter specifies if the tunnel path has to be dynamic (i.e. calculated by the router itself) or specified manually hop by hop.

The following is an example of QoS configuration that gives 600Mbit/sec to the traffic labelled tunnel and the available bandwidth to the rest of the traffic.

policy-map gbw

 class tunnel

 police cir 600000000 bc 4470 be 4470 conform-action transmit exceed-action drop

 bandwidth percent 25

 random-detect

 random-detect precedence 5 2000 packets 20000 packets 1

 class class-default

 bandwidth percent 74

 random-detect

 random-detect precedence 0 2000 packets 20000 packets 1

The class tunnel is defined as follow:

class-map match-any tunnel

 match mpls experimental 5

So the traffic entering the network that is supposed to go in the tunnel is marked with IP precedence 5, this value is copied into the EXP field in the MPLS header and in the rest of the network the QoS settings on each interface can check either of them.

**Useful links**

MPS-TE with ISIS

http://cisco.com/en/US/tech/tk436/tk428/technologies\_configuration\_example09186a0080093fcb.shtml

MPLS-TE with OSPF

http://cisco.com/en/US/tech/tk436/tk428/technologies\_configuration\_example09186a0080093fd0.shtml

MPLS-TE and QoS

http://www.cisco.com/warp/public/cc/pd/iosw/tech/mpotc\_qp.htm

Guaranteed Bandwidth and MPLS

http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/gurtb\_wp.htm