MPLS Deployment
Over PERN

By: JAWAD RAZA
Topics

- MPLS Basics
- MPLS TE (Traffic Engineering)
- MPLS Data Transport Services
- MPLS over PERN
- FRR (Fast Reroute)
- BFD (Bi-directional Forwarding)
MPLS Basic Knowledge
MPLS

MPLS—Multi-Protocol Label Switching

- Multi-Protocol

- Support multiple Layer-3 protocols, such as IP, IPv6, IPX, SNA

- Label Switching

- Label packets, and replace IP forwarding with label switching

- Work in Core
Origin: To Integrate IP with ATM

IP
- Connectionless control plane
- Connectionless forwarding plane

MPLS
- Connectionless control plane
- Connection-oriented forwarding plane

ATM
- Connection-oriented control plane
- Connection-oriented forwarding plane
Traditional IP Forwarding

<table>
<thead>
<tr>
<th>Dest</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>47.1</td>
<td>1</td>
</tr>
<tr>
<td>47.2</td>
<td>2</td>
</tr>
<tr>
<td>47.3</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dest</th>
<th>Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>47.1</td>
<td>1</td>
</tr>
<tr>
<td>47.2</td>
<td>2</td>
</tr>
<tr>
<td>47.3</td>
<td>3</td>
</tr>
</tbody>
</table>

47.3

IP 47.1.1.1

147.1

IP 47.1.1.1

47.2

IP 47.1.1.1

47.3

IP 47.1.1.1

1

IP 47.1.1.1

2

IP 47.1.1.1

3
MPLS IP forwarding via Label Switched Path (LSP)
Basic MPLS Concepts

LSR: Label Switch Router
LER: Label Edge Router
LSP: Label Switch Path
Basic Working Process of MPLS

Traditional IP forwarding

Label forwarding

Traditional IP forwarding
How MPLS Works – Basic Terminologies

MPLS router roles may also be expressed as “P” or “PE”:

- **Terms** which come from the description of VPN services.
- **P** – Provider Router
  - A core/backbone router which is doing label switching only.
  - A pure P router can operate without any customer/Internet routes at all.
  - This is common in large service provider networks.
- **PE** – Provider Edge Router
  - A customer facing router which does label popping and imposition.
  - Typically has various edge features for terminating multiple services:
    - Internet
    - L3VPN
    - L2VPN / Pseudowires
    - VPLS
  - **CE** is the “Customer Edge”, the customer device a PE router talks to
MPLS Advantages

Replace IP header with short and fixed-length labels as forwarding basis to improve forwarding speed

Better integrate IP with ATM

Provide value-added service without prejudice to efficiency:
  ◦ VPN
  ◦ Traffic engineering
  ◦ QOS
Key MPLS Capabilities

- IP Multicast
- IP/ATM Integration
- IP CoS
- Traffic Engineering
- VPN’s
Traffic Engineering
Concept of TE

**NE (Network Engineering) and TE (Traffic Engineering)**

Two engineering concepts can accompany with the growth of network size:

- **NE (Network Engineering)**
  - NE is the process of how to plan, design and implement a network to satisfy the real traffic requirements. First real traffic requirements exist, then start network design and implementation.

- **TE (Traffic Engineering)**
  - TE is the process of how to configure existing network equipments to make traffic transmitted efficiently. First network equipments exist, then configure.
What is Traffic Engineering

- What is Traffic Engineering?
  - Classic IGPs use non-TE routing, i.e. a metric (cost) per link, and a shortest path first (SPF) algorithm to find the shortest path.
  - Traffic Engineering takes this, and adds additional constraints.
    - For example, find the shortest path that also has available bandwidth.
    - This is also called constrained routing, using a CSPF algorithm.
  - The principal is simple: It is better to take an uncongested path even though the latency may be higher, than to congest the shortest path on one link while leaving available bandwidth unused on another link.

- Why can’t I just do this manually with my IGP costs?
  - You can, but this only scales up to a certain point.
  - As networks become more complex, this gets harder to manage.
    - Changing an IGP cost by 1 can easily affect routing dozens of hops away.
How Does MPLS Traffic Engineering Work

- Using RSVP-TE to reserve bandwidth across the network.
  - Remember, an LSP is a “tunnel” between two points in the network.
  - Under RSVP, each LSP has a bandwidth value associated with it.
  - Using constrained routing, RSVP-TE looks for the shortest path with enough available bandwidth to carry a particular LSP.
  - If bandwidth is available, the LSP is signaled across a set of links.
  - The LSP bandwidth is removed from the “available bandwidth pool”.
  - Future LSPs may be denied if there is insufficient bandwidth.
    - They’ll ideally be routed via some other path, even if the latency is higher.
  - Existing LSPs may be “preempted” for new higher priority LSPs.
    - You can create higher and lower priority LSPs, and map certain customers or certain traffic onto each one.
    - This isn’t traditional QoS, no packets are being dropped when bandwidth isn’t available, you’re simply giving certain traffic access to shorter paths.
How RSVP-TE Reserves Bandwidth

PATH message 20Mbps

RSVP PATH: R1 ➞ R2 ➞ R6 ➞ R7 ➞ R4 ➞ R9
RSVP RESV: Returns labels and reserves bandwidth

Bandwidth available on each link
Label value returned via RESV message
How Do You Determine an LSP Bandwidth?

- How do you determine the bandwidth of particular LSP?
  - After all, IP networks are dynamic and packet switched.
  - Bandwidth use can change in an instant, and be unpredictable.

- There are basically two main ways to do it:
  - **Offline Calculation**
    - Calculation which occurs outside of the router, typically based on some bandwidth "modeling", and often using a third party script or tool.
    - This is how MPLS was first implemented, and is still commonly used today by most large networks and early MPLS adopters.
  - **Auto-Bandwidth**
    - The bandwidth value is calculated on the router, by periodically measuring how much traffic is actually forwarding over the LSP.
    - The RSVP reservation is then periodically updated with the new number
MPLS TE level specification

MPLS TE also support some high level specification
- FRR (Fast Reroute)
- Tunnel backup
- Auto Bandwidth Allocation
- Path Re-optimum
MPLS Data Transport Service
MPLS Pseudowires

- An emulated layer-2 point-to-point circuit, delivered over MPLS.
  - Currently standardized by the “PWE3” IETF Working Group.
  - Can be used to interconnect two different types of media:
    - For example, Ethernet to Frame Relay.
  - Useful for migrating legacy transport (e.g. ATM) to an MPLS network.
  - Can be difficult to load balance, due to lack of visibility into the packet.
    - The payload is unknown, so you can’t hash on the IP header inside, etc
MPLS L3VPNs

- L3VPN An IP based VPN.
  - Networks build virtual routing domains (VRFs) on their edge routers.
  - Customers are placed within a VRF, and exchange routes with the provider router in a protected routing-instance, usually BGP or IGP.
  - Can support complex topologies and interconnect many sites.
  - Usually load-balancing hash friendly (has exposed IP headers).
  - But can add a significant load to the service provider infrastructure.
    - Since the PE device must absorb the customer’s routing table, consuming RIB and FIB capacity.
  - Typically seen in more enterprise environments.
  - Signaled via BGP within the provider network.
MPLS VPLS

- VPLS (Virtual Private LAN Service)
  - Creates an Ethernet multipoint switching service over MPLS.
  - Used to link a large number of customer endpoints in a common broadcast domain.
  - Avoids the need to provision a full mesh of L2 circuits.
  - Emulates the basic functions of a layer 2 switch:
    - Unknown unicast flooding
    - Mac learning
    - Broadcasts
  - Typically load-balancing friendly since the L2 Ethernet headers are examined and used, unlike L2 pseudo wires where they are passed transparently.
MPLS OVER PERN
PERN2 MPLS L3 Services

Pre-Deployed Layer 3 VPNS

- Internet L3 VPN:
  - Carrying all Internet traffic of Customer

- Intranet L3 VPN
  - Within PERN

- R&D VPN
  - Carrying R&D Traffic Towards TEIN4 Link
Traffic Categories

A – Internet traffic:
- IP Transit Connectivity

B – Intranet traffic:
- Connectivity among the PERN2 Campuses
- Intranet Bandwidth should be Separate from Internet Bandwidth

C – International NREN (R&D) traffic
- For the R&D traffic Bandwidth must be separate from Intranet and Internet
PERN2 MPLS L3 Services
MPLS L3VPN Service (BGP MPLS VPN)
MPLS L2VPN Service (VPLS)

Legend

- 10G Link
- GE link
- PW Tunnel

Campus

NE80E
NE40E
NE20E

NE20E-PE

VLAN1

VSI 1

VLAN1

C Campus

NE20E-PE

VLAN1

C Campus

VLAN1

MPLS L2VPN Service (VPLS)
Reliability on PERN2

Core:
- TE FRR (FAST Reroute)
- BFD (Bi-direction Forwarding Detection)
What Does Fast Reroute Do
What Does Fast Reroute Do?

- MPLS Fast Reroute improves convergence during a failure.
  - By pre-calculating backup paths for potential link or node failures.
  - In a normal IP network
  - The best path calculation happens on-demand when a failure is detected.
  - It can take several seconds to recalculate best paths and push those changes to the router hardware, particularly on a busy router.
  - A transient routing loop may also occur, as every router in the networks learns about the topology change.

- With MPLS Fast Reroute
  - The next best path calculation happens before the failure actually occurs.
  - The backup paths are pre-programmed into the router FIB awaiting activation, which can happen in milliseconds following failure detection.
  - Because the entire path is set within the LSP, routing loops cannot occur during convergence, even if the path is briefly suboptimal.
LDP over TE sketch map

Legend
- **TE Backup tunnel**
- **LDP LSP**

NE80E
NE40E

QTA-BUITMS-P-PF-NE40E
MPLS TE FRR Tunnels for HEC-ISB Router
## MPLS TE FRR

<table>
<thead>
<tr>
<th>P router</th>
<th>ingress tunnel number</th>
<th>egress tunnel number</th>
<th>transit tunnel number</th>
<th>Max tunnel number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISB-HEC-P-PE-EGRESS-NE80E</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>42</td>
</tr>
<tr>
<td>ISB-AU-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>ISB-FJWU-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>ISB-QAU-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>PSH-HEC-PRR-PE-NE40E</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>LHR-HEC-PRR-PE-NE80E</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>LHR-GCUL-PE-NE40E</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>LHR-PU-NEW-PE-NE40E</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>
## MPLS TE FRR

<table>
<thead>
<tr>
<th>P router</th>
<th>Ingress tunnel number</th>
<th>Egress tunnel number</th>
<th>Transit tunnel number</th>
<th>Max tunnel number</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLT-BZU-P-PE-NE40E</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>42</td>
</tr>
<tr>
<td>FSD-AUF-P-PE-NE40E</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>QTA-BUITMS-P-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>HYD-USINDH-P-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>KHI-HEC-P-PE-EGRESS-NE80E</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>KHI-CPSP-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>KHI-KU-PE-NE40E</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td><strong>Sum (15 P/PE routers)</strong></td>
<td><strong>116</strong></td>
<td><strong>116</strong></td>
<td><strong>110</strong></td>
<td><strong>342</strong></td>
</tr>
</tbody>
</table>
Fast Reroute

Is a mechanism for protecting Multiprotocol Label Switching (MPLS) traffic engineering (TE) label switched paths (LSPs) from link and node failures.

FRR locally repairs the protected LSPs by rerouting them over backup tunnels that bypass failed links or nodes.
MPLS Deployment

- L3 VPN
- L2 VPN

Reliability
- TE FRR
- BFD
BFD (Bidirectional Forwarding Detection)

BFD is a detection protocol designed to provide fast forwarding path failure detection times for all media types Encapsulations, topologies, and routing protocols.

If the primary path Between ISB-AU & ISB-QAU fails, Traffic Will automatically Shift to Alternate path

But IS-IS has its Own link failure Detection convergence time
BFD

BFD is a detection protocol designed to provide fast forwarding path failure detection times for all media types Encapsulations, topologies, and routing protocols.

BFD will Fast forward the Link failure detection time So to increase its convergence Time
BFD for BGP

PERN2 implementation:

- Configure the multi-hop BFD between iBGP peers.
- Fast detect and monitor the links nodes failure between two iBGP peers in milliseconds.
- Inform BGP application to act immediately.
- Similarly in case of IS-IS.
Thanks