Tutorial: MPLS Applications
Overview

Ina Minei
ina@juniper.net
Disclaimer

- The views presented are of the author and do not necessarily represent Juniper Networks.
Topics

- Traffic engineering
- Fast reroute
- VPNs
- Pseudo-wires
- Protocol comparison
Topics

◆ Traffic engineering
◆ Fast reroute
◆ VPNs
◆ Pseudo-wires
◆ Protocol comparison
Traffic engineering (TE)

- Avoid a situation in which some parts of the network are congested while others are underutilized.

- Goal: most efficient use of available resources, especially bandwidth.
Routing determines the paths taken by traffic, thus controls how much traffic traverses each link.

IP routing characteristics:
1. Forwarding based on the destination address.
2. Routing changes can be made via manipulation of protocol metrics.
An example –
The fish topology – shortest path

All links are 150Mbps.
A and B are each sending 100Mbps
The fish topology - equal cost
The fish topology - controlling forwarding

Traffic from A should take the shorter path.
The fish topology - bandwidth constraints

All links are 150Mbps, except E-F which is 50Mbs
A is sending 120Mbs, B is sending 40Mbs
The problems

- Need path that is optimal with regard to some metric and also needs to take into account other constraints => constraint-based routing.
Constraint-based routing

1. Requires path calculation at the beginning of the path (source of the path).

2. When the path is determined by the source, can’t use destination-based forwarding to forward traffic along that path.
Traffic engineering with constraint-based routing

- Path calculation and setup:
  1. Information distribution
  2. Path selection
  3. Path signaling

- Forwarding
Path calculation

- CSPF - constrained SPF

- Like conventional SPF, computes shortest path (with regard to some administrative metric).

- But takes into account only paths that satisfy one or more user-defined constraints (e.g. available bandwidth)
Path calculation (cont)

- Requires extensions to the I GPs to carry additional information (available bw).

- Can be done online or offline.

- Explained in detail at the previous nanog.
Path signaling and forwarding

- Useful properties of MPLS for this context:

1. Can establish paths based on explicit routes using RSVP (the ERO object).

2. Forwarding is based on a label rather than on a destination address.
Traffic engineering with MPLS

- Information distribution => IGP extensions
- Path selection => CSPF
- Path signaling => RSVP
- Forwarding => MPLS

The result:
Paths that are optimal with regards to both routing protocol metrics and comply to given constraints.
All links are 150Mbps.
A and B are each sending 100Mbps
The fish topology revisited (2)

All links are 150Mbps, except E-F which is 50Mbps

A is sending 120Mbps, B is sending 40Mbps
Split the traffic from B.
Traffic engineering BGP destinations

- When an LSP is available to the BGP nexthop of a particular route, can use the LSP to forward the traffic to that route.
LSP between ASBRs. All transit traffic will use this LSP.
Traffic engineering BGP destinations (cont)

- Control the path that transit traffic takes inside the domain.

- Forwarding is done based on MPLS labels. The routers in the middle of the network don’t need to have knowledge of the destinations.

- It is possible to have a BGP-free core.
What about load sharing?

- Create several tunnels to the same destination.

- Load balance the BGP traffic across these tunnels.
Traffic engineering IGP destinations

- Advantage: allows mixing paths determined by constraint-based routing with paths determined by IP routing -> can apply traffic engineering to only a portion of the network.

- Attractive scaling property.
Traffic engineering IGP destinations

- Two concepts:
  1. Allow IGP on the LSP head-end to use the LSP in the SPF computation. (other routers in the network will not know about the existence of this LSP).
  2. Advertise the LSP in the link-state advertisements, so that other routers can also take it into account in their computations.
Allow SPF at the head-end to use the LSP

- The idea: modify SPF to take advantage of this LSP.

- The shortest path to the LSP endpoint and to destinations behind it will be through the LSP.
Allow the head-end to use the LSP - example
Allow the head-end to use the LSP - example SPF tree from E
Advertising LSPs in the IGP - why is it needed?
Advertising LSPs in the IGP - example

A-D shortest path 25

LSP E-D metric 15
Advertising LSPs in the IGP

- Advertise the LSP as a unidirectional, point-to-point link in the link-state database.
- All routes can compute paths using the LSP.
Summary 1: Applications of MPLS to traffic engineering

- Load share traffic across paths with unequal cost.
- Enable definition of flexible forwarding policies.
- Route primary paths away from known bottlenecks or points of congestion.
- Control the path of traffic to destinations outside the domain.
- Mix routing determined by MPLS constraint-based routing with routing determined by plain IP. Thus no need for a full-mesh of LSPs everywhere.
Other features useful for traffic engineering

- LSP priorities
- Automatic bandwidth adjustment
- Path protection
LSP priority and preemption

- The idea: some LSPs are more important than others, and can “kick out” the less important ones, when resource contention occurs.

- This happens during computation and setup, not at forwarding time.
Auto-bandwidth

- The idea - monitor the traffic rate on an LSP, and resize the bandwidth on the LSP to match with the traffic rates going down the tunnel.
- May result in an LSP rerouting.
- The traffic flowing down the LSP will not be affected.
- Turned on at the head end of the LSP.
Path protection

◆ The idea - protect an LSP by having an additional (secondary standby) LSP set up in parallel to it.

◆ The secondary standby is signaled ahead of time, from the head end. Can be made to avoid crossing the same links/nodes as the protected LSP.

◆ The secondary standby is used only in case of failure, but is up and ready all the time.
Path protection

Diagram showing network nodes labeled A, B, C, D, and E. Arrows indicate the direction of data flow. The connections are labeled as Primary 50M and Secondary standby 50M.
Path protection (cont)

◆ Advantage:

1. Gives precise control over where the traffic reroutes in case of failure.
2. The secondary standby is taking a diverse path.
Path protection (cont)

- **Disadvantages:**
  1. Reserves resources for the backup, but most of the time they are not used.
  2. Builds one secondary standby for each protected path.
  3. Takes effect when the head end finds out about the failure.
Summary 2: More applications of MPLS to traffic engineering

- Provide a mechanism to prioritize LSPs in the case of resource contention
- Automatic bandwidth adjustment
- Provide precise control over how a path is rerouted in case of a single or multiple failures
Topics

- Traffic engineering
- **Fast reroute**
- VPNs
- Pseudo-wires
- Protocol comparison
Fast reroute

- **Goal** - reduce packet loss during routing transients.

- **Two factors:**
  1. The time it takes to detect the failure (e.g. the time it takes to detect a link-down event - rely on link-layer mechanisms).
  2. The time it takes to distribute the information about the failure and recompute forwarding tables.
Fast reroute with IP?

- IP routing protocols are distributed in nature, and require that all routers converge to a consistent view of the routing information.
- In a large network, may have convergence in the order of a few seconds.
- A link failure can cause congestion in one part of the network, while leaving other parts free of congestion.
Fast reroute - the idea

- The idea: reroute the traffic around the failure along an alternative path.

- Goal: reduce (not eliminate!) packet loss.

- Local protection - link protection, node protection. Will only discuss link protection.
Fast reroute - the idea

Protection path around the link A-B

Will still have traffic loss: the failure needs to be detected, the traffic needs to be switched to the alternate path.
Fast reroute with IP?

Best path C-B this way

Best path A-B this way

Protection path A-B this way

Goal - protect against failure of the link A-B
Fast reroute with IP ? (cont)
Fast reroute with MPLS

The problem: hop-by-hop, destination-based routing.

The idea: construct a “protection” LSP around a point of failure. Nest the LSPs that traverse the point of failure onto the protection LSP.
Fast reroute with MPLS (cont)

LSP X -> Y

Protection LSP for link A-B
Fast reroute with MPLS (cont)

- **Push 101** (X to A)
- **Swap 101->100** (A to B)
- **Pop 100** (B to Y)

Protection LSP for link A-B

- **Swap 201->200** (C to D)
- **Pop 200** (D to Y)
Fast reroute with MPLS (cont)

Push 101, Swap 101->100, Pop 100

Protection LSP for link A-B

Swap 201->200, Pop 200
Fast reroute with MPLS (cont)

- **Push 101**
- **Swap 101->100**
- **Push 201**
- **Pop 100**

- **Protection LSP for link A-B**

- **Swap 201->200**
- **Pop 200**
Fast reroute with MPLS (cont)

- **Push 101**
  - X

- **Swap 101->100**
  - Push 201
  - A

- **Pop 100**
  - B

- **Push 101**
  - Y

- **Swap 201->200**
  - IP 100
  - C

- **Pop 200**
  - IP 200
  - D
Link protection

- The protection LSP is set up ahead of time.
- Traffic loss still occurs, for the period of time until A detects the link-down event and switches the traffic to the protection LSP.
- The intention is to use the protection LSP for a short time until a new path will be recomputed.
Link protection

- The protection LSP protects a link.
- The protection LSP is signaled around the protected link, from the node upstream of the link to the node downstream of the link, ahead of time.
- Traffic from one LSP or from many LSPs can be carried over it. Different scaling properties.
Fast reroute - summary

- Traffic loss is a function of the link failure detection time instead of being a function of the head-end rerouting.
- The protection LSP is intended for short-time use after the failure.
- One can choose to protect just a few of the critical resources.
Topics

- Traffic engineering
- Fast reroute
- VPNs
- Pseudo-wires
- Protocol comparison
VPNs

- Virtual Private Networks - provide a private network over a shared infrastructure.

- Interconnect geographically separate sites, with the same privacy and guarantees as a private network.
VPNs

Site 1

VPN green

Site 2

VPN green

Provider network

Site 3

VPN green

Site 4

VPN green
The Overlay Model for VPNs

- Sites are connected with p2p links - leased lines, FR circuits, ATM circuits, GRE, IPsec.
- Customer routers peer with customer routers.
- The provider needs to design and operate “virtual backbones” for all the customers - scaling issue.
- Problem with VPNs that have a large number of sites.
- Adding a new site requires configuring all the existing sites.
BGP-MPLS VPNs

- **Goal:** solve the scaling issues. Support thousands of VPNs, support VPNs with hundreds of sites per VPN, support overlapping address space.

- **Peer model -** customer routers peer with provider routers.
Properties of the model

- CE router peers with a PE router, but not with other CE routers.

- Adding/deleting a new site requires configuring the PE router connected to the site.

- A PE router only needs to maintain routes for the VPNs whose sites are directly connected.
BGP-MPLS VPNs - areas

- Separation of forwarding
- Distribution of routing information
- New address type
- Forwarding with MPLS
Operation - separation of forwarding

- **Goal**: control connectivity by segregating the forwarding information.

- **PE router connected to CEs from several VPNs.**

- **With a single forwarding table, it is possible to forward packets from one VPN to another.**
Multiple forwarding tables

- Multiple forwarding tables - each table associated with a site.

- Packets from the customer are identified based on the incoming port, which identifies the forwarding table.

- Contents: routes received from the CE, and routes received from remote PEs with constrained routing.
Operation - Distribution of routing information

The idea:

1. CE advertises routes to the local PE via some routing protocol.
2. The local PE marks these routes with a particular (extended) community and advertises them in BGP.
3. The routes are distributed to all remote PE by BGP.
4. Remote PE receives BGP routes, filters them based on the community and advertises them to the CE.
Example

Customer green

Site 2
10.2/16

CE2

Customer green

Site 1
10.1/16

CE1

Customer gray

Site 2
10.1/16

CE4

Customer gray

Site 1
10.2/16

CE3

10.1/16 community green
10.2/16 community green

10.1/16 community gray
10.2/16 community gray
The model so far (1)

- The addresses used in the VPNs need to be unique in the provider’s network.

- The P routers carry all VPN routes.
Operation: overlapping address space and VPN-IP addresses

- Goal: turn non-unique addresses into unique addresses.

- Constructed by concatenating an IP address and an 8 byte unique identifier called the route distinguisher.

- Route Distinguisher - Type (2) AS (2) Assigned Number (4) - doesn’t have to be the same for all routes in the VPN.
VPN-IP addresses (cont)

- Advertised in a special address family by BGP.
- Used only in the provider’s network.
- Used only in the control plane.
- The translation from IP addresses to VPN-IP addresses happens on the PE.
- Not used for route filtering (we use communities for that).
Example using VPN-IP addresses

**Customer green**
- Site 1: 10.1/16
- CE1

**Customer green**
- Site 2: 10.2/16
- CE2

**Customer gray**
- Site 2: 10.1/16
- CE4
- Site 1: 10.2/16
- CE3

RD1: 10.1/16 community green
RD2: 10.1/16 community gray
RD1: 10.2/16 community green
RD2: 10.2/16 community gray
The model so far (2)

- Can use overlapping address space.

- How to forward based on VPN-IP addresses?

- The P routers still carry all the VPN routes.
MPLS-VPNs

- VPN-IP addresses are used by the routing protocols, but do not appear in headers of IP packets.

- Need a way to forward traffic along routes to VPN-IP addresses. MPLS decouples forwarding from the destination information.
Forwarding traffic - so far (1)

10.1/16 interface to CE1
10.2/16 from PE2
Forwarding traffic - so far (2)

Site 1
10.1/16
VPN green
CE1
10.2.1.1
PE1
10.1/16 interface to CE1 10.2/16 from PE2

Site 2
10.2/16
VPN green
CE2
10.2/16 from PE2

Site 1
10.1/16
VPN gray
CE4
10.1/16 interface to CE4

Site 2
10.2/16
VPN gray
CE3

PE2
10.2.1.1 Label

???
VPN labels

- The idea: Use a label to identify the VPN.
- The VPN label is distributed by BGP, along with the VPN-IP address.
- Traffic will carry two labels, the VPN label and the LSP label.
- The remote PE makes the forwarding decision based on the VPN label.
Forwarding traffic - revisited

Site 1
10.1/16
VPN green

CE1

10.2.1.1
VPN green

Site 2
10.2/16
VPN gray

CE2

10.2.1.1
VPN gray

CE3

Site 1
10.1/16
VPN green

PE1

P

P

P

PE2

CE3

Site 2
10.2/16
VPN gray

10.2.1.1
VPN Label
Forwarding traffic - revisited
Summary

- P routers don’t need to maintain VPN routes at all. Only need to maintain routes to other P and PE routers.
- PE routers maintain VPN routes, but only for VPNs that have sites attached to them.
- VPNs can have overlapping address spaces.
Topics

- Traffic engineering
- Fast reroute
- VPNs
- Pseudo-wires
- Protocol comparison
Pseudo wire

- The idea: emulate a point to point layer 2 connection over an MPLS network. Carry PDUs of layer 2 protocols over MPLS.
- Extend the same model from L3VPN. Similarly, use an additional label to improve scalability.
- In L3VPNs, the CEs advertise IP addresses. Here, they advertise “L2 information”.

The idea
Applications to L2

- Need one label per circuit.

- The label can be advertised with BGP or with special extensions to LDP.
Topics

- Traffic engineering
- Fast reroute
- VPNs
- Pseudo-wires
- Protocol comparison
LDP and RSVP

- Why are we talking about them?
- Overview of the protocols
- Comparison and applicability
Protocol LDP – Label Distribution

- Not a routing protocol, relies on other routing protocols for forwarding decisions, loop prevention, etc.
- Label distribution initiated from the endpoint.
- There is no concept of the head-end requesting the establishment of an LSP to an endpoint.
- Labels exchanged between LDP neighbors
LDP - (continued)

- The created LSPs follow the IP shortest path.
- TCP based (reliable), incremental updates.
- Allows creation of multiple paths for the same prefix (load balancing)
- Useful when need to establish a large number of LSPs. Easy to set up.
RSVP - TE

- Soft state, requires periodic refreshes.
- Creates point to point tunnels, initiated from the head end. Labels distributed only along this path.
RSVP - TE (cont)

- Supports explicit paths. Can set up LSPs along paths computed with CSPF, so can take into account bandwidth allocations.
- Supports fast restoration in case of failures.
Comparison

- Ease of configuration - both initially, and when making incremental additions.
- State maintenance
- Tracking the IGP state to determine forwarding state
# LDP & RSVP applicability

<table>
<thead>
<tr>
<th></th>
<th>LDP</th>
<th>RSVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label distribution</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Traffic engineering</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Traffic protection</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>VPN</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
More …

- RSVP soft preemption
- Diffserv-TE
- Node protection
- Graceful restart
- LSP ping
- VPNs
- MPLS management - mibs
- Inter-AS RSVP tunnels
- VPLS
Thank you!

Please send comments to ina@juniper.net