Multi-Topology Routing

Truman Boyes
truman@juniper.net
Traffic Engineering

- **Choices Today:**
  - IGP Metric Costing
  - RSVP-TE end-to-end
  - Policy based routing
  - EROs, Offline TE calculations, etc, etc
  - Class Based Forwarding
Overview

- *This is not an alternative to MPLS.*
- *There is much confusion on MTR and where it best fits.*

- MTR divides a network into multiple logical groups. It provides isolation but not VPNs.

- MTR creates multiple RIB / FIB combinations in addition to the base routing context.
- Topologies can be shared or separate.
- N-dimensional vectors of different costs for each link in the network
Use Cases

A FEW REASONS WHY YOU MIGHT WANT MTR

Voice

Video

BE Data
MPLS + IP TE

- You already have an MPLS Network.
- You want to carry IPv4 Internet traffic inside your global default routing instance.
- You don’t want Internet traffic to traverse some links in your network.
IPv6 video traffic MTR

- You carry IPv4 and IPv6 routes in your default routing context.
- You use OSPF to populate the respective RIBs.
- You want IPv6 video traffic to use specific links.
- You want IPv6 data traffic to use any link.
Wholesale / Reseller MTR

- Specific prefixes / Interfaces are mapped into topology contexts. (ie. Wholesale).
- Wholesale context is engineered to stay on specific technologies (Ethernet) and specific nodes.
- Wholesale links use different queuing characteristics and may not have MPLS FRR.
ISP with MPLS LSPs for EF traffic, but BE should use IP Next-Hops

- ISP may not want a lot of MPLS-encap overhead for high volume BE traffic.
- MPLS is used in strategic location to provide FRR and VPNs to Business Traffic.
- Many IP-Only Routers in the Network.
- By default JUNOS will prefer LSPs instead of IP Next-Hops.
- This concept can be achieved with MTR implementation.
Non-Congruent Inter-domain Topologies

- Autonomous System may want to receive VoIP on a given prefix on one link (due to latency) and all other data for the same prefix on another link.
- Could be solved with policy and MTR.
- Alternatives could be multiple LSPs with class-based-forwarding policy, and Inter-AS VPNs. This may be more complex...
WHAT MTR LOOKS LIKE
On a single router: multiple topologies

- The default topology is populated from the IGP and BGP.
- Individual topologies are created, which contain their own RIB / FIB.
- Firewall policies / Classifiers are created to place traffic into the appropriate FIB.
- Routing protocols (OSPF, BGP, static) supported.
Not Quite a Standard…

- **OSPF-MT (RFC 4915)**
  - Copy MT-ID into ToS field (32-127)

- **BGP MTR – No RFC**
  - JUNOS uses community values for each topology ID.
  - Each route is installed into default RIB & topology RIB.
  - Route target Community-based, not a new SAFI.
  - IOS version uses BGP session per topology. This does not require a new protocol extension.

- **ISIS-MT (RFC 5120): new TLV.**
  - Reserved MT-IDs for IPv4 and IPv6 routing topology.
Multi-Topology Example

Base topology
Multicast topology
Voice topology
Multicast Group Redundancy

Server A

Stream A

Server B

Receiver A

Stream B

Receiver B
Constructs

- **Topology RIB/FIB creations.** (ie. Define the topology names you want).
  - Do this on all routers.

- **Enable MT extensions on IGP**
  - OSPF MT-IDs

- **Enable BGP topology / route control**
  - iBGP routes inside a topology context will use the respective OSPF/ISIS context for next-hop route resolution.

- **Modify forwarding plane with firewall filters to place the respective traffic into each topology FIB.**
OSPF-MT details

- **ToS Field redefined as MT-ID**
- **MT Area Boundary**
  - Each OSPF interface belongs to a single area and all MT’s sharing the link must belong to the same area.
- **Adjacencies**
  - Single adjacency even if there are multiple topologies.
- **LSAs**
  - Router / Summary / Type-5 and Type-7 AS External LSAs are modified to contain MT-ID.
  - No change to Network LSA. DR is shared by all topologies.
More OSPF-MT Details

- Separate SPF calculation for each topology.
- Max metric can be used, but OSPF doesn’t have concept of unreachable link.
  - There is a new concept with MT, where links can be excluded (in the spec) from the default MT-ID #0 topology.
- MT-bit is in Hello packet
Migrating non MT Areas to MT Areas

- If you need to exclude links from the default topology, all routers in the area must support DefaultExclusionCapability.
- MT Area routers will interact with non-MT-Area routers in the default topology.
- Upgrading the backbone area first is desirable.
- Gradually you can upgrade the entire area to MT.
Ideas

- Management Topology RIB
- Video Dual Head Feeds over 2 paths
- Prevent additional overhead on some networks
- Offline TE metric calculations for OSPF-MT
  - Integration with Event/Commit/Op Scripts for Automatic behaviour.
  - XML-RPC/NetCONF reconfigurations of metrics on IP networks.
- Using Overload Bits per Topology!
  - Awesome way to redirect traffic when working on a service.
- Advertising LSPs into OSPF-MT
A bit about BGP Multi Topologies

* 3 approaches
  * Community based
  * Multi-Session
  * New AFI/SAFI

* Community based approach is better on memory, but is more limited.
  * Communities might make sense between border routers, but beyond that they don’t work well between AS if the same prefix is advertised with different topologies.
Summary

- Another tool to add to the tool belt.
- Relatively easy to setup in IP networks.
- Integrates with MPLS networks.
- There are multiple ways to match traffic.
  - BA or Firewall Terms
- Troubleshooting with standard tools
  - Ping / Traceroute / Tcpendump
- Available Now
IMPLEMENTATION
Configurations

```
[edit protocols ospf]
topology (default | ipv4-multicast | name) {
    topology-id number;

    spf-options {
        delay milliseconds;
        holddown milliseconds;
        rapid-runs number;
    }
}
```
Static Routes in MTR

[edit routing-options]

rib routing-table-name {
    static {
        route destination-prefix {
            next-hop;
        }
        static-options;
    }
}

JUNOS MTR with BGP

group ebgp {
    type external;
    local-address 11.19.30.1;
        family inet {
            unicast {
                topology voice {
                    community target:40:40;
                }
                topology video {
                    community target:40:40;
                }
            }
            peer-as 101;
            neighbor 11.19.30.2;
        }
    }
}
MTR and the Forwarding Table

```c
firewall {
  family inet {
    filter ef_path {
      term ef {
        from {
          forwarding-class expedited-forwarding; Filter on DSCP
        }
        then topology voice;
      }
      term video {
        from {
          source-address {
            Filter on address
            11.19.132.0/24;
            11.19.133.0/24;
            11.19.142.0/24;
            11.19.144.0/24;
          }
          then Topology video;
        }
      }
      term default {
        then accept;
      }
    }
  }
}
```
Router(config-router-af)# topology VOICE tid 10

router bgp <autonomous-system-number>
! Global commands
scope {global | vrf <vrf-name>}
! Scoped commands
address-family {<afi>} [<safi>]
! Address family specific commands
topology {<topology-name> | base}
! Topology specific commands
More IOS Snippets

global-address-family ipv4
topology VOIP
  all-interfaces
  exit
topology HDTV
  forward-base
  maximum routes 10000 90
  exit
  exit
class-map match-any VOIP
  match ip dscp 9
  exit
class-map match-any HDTV
  match ip dscp af11
  exit
policy-map type class-routing ipv4 unicast MTR
class VOIP
  select-topology VOIP
  exit
class HDTV
  select-topology HDTV
  exit
  exit
global-address-family ipv4
  service-policy type class-routing MTR
end
Specific Route Resolution

routing-options {
    autonomous-system 65300;
    resolution {
        rib inet.0 {
            resolution-ribs inet.0;
        }
        rib :voice.inet.0 {
            resolution-ribs [ inet.3 :voice.inet.0 ];
        }
    }
    topologies {
        family inet {
            topology voice;
        }
    }
}
Thanks!

- truman@juniper.net