CCIE Service Provider v4 Advanced Technologies Class

Core MPLS Switching

In This Section

- BGP over MPLS Example
- MPLS Components Overview
- Label Distribution Protocol

The BGP Problem Statement

- Global BGP table is huge, and growing
  - Over 500,000 IPv4 prefixes and growing
  - IPv6 space is growing but currently negligible
  - See http://bgp.potaroo.net/ for table growth stats
- Why is this a problem?
  - IP routing is destination based
  - All devices in the transit path must know the destination
  - E.g. all transit routers must have full BGP feed(s)

Routing Through vs. To the Core

- Transit providers sell transit, not applications
  - E.g. ISP is not the same as an ASP
- Traffic routes through the SP, not to the SP
  - E.g. end client needs to ping end application, not core link
- How does this affect core routing?
  - To SP core, only the ingress point and egress point matter
  - Original source and final destination are arbitrary

Tunnels – The Ultimate Band-Aid

- Simple transit solution is to tunnel traffic over core from ingress to egress
  - Only the ingress and egress devices need full end-to-end information
  - Core only needs info about ingress and egress devices
- How can we tunnel?
  - QinQ, GRE, IPinIP, MPLS, etc.
  - MPLS is de-facto standard

Example Case - BGP over GRE over Core

» Form a GRE tunnel from ingress to egress
  - Tunnel subnet is link-local and arbitrary
» Peer BGP from ingress to egress
» Recurse BGP next-hop to tunnel
  - Either peer through the tunnel, or modify next-hop to the tunnel
» What is the core's data plane result?
  - Core routes ingress PE to egress PE
  - Core does not need end-to-end information

Where MPLS Fits In

- MPLS is the core's tunnel encapsulation
  - Same exact logic as GRE
- MPLS is more flexible
  - Arbitrary transport
  - Arbitrary payload
  - Extensible applications
  - Much more on this later…

Example Case – BGP over MPLS over Core

- **Form an MPLS tunnel from ingress to egress**
  - Typically IGP + LDP is used for this
  - Could be BGP or RSVP (MPLS TE)
  - More on this later…
- **Peer BGP from ingress to egress**
- **Recurse BGP next-hop to MPLS label**
- **What is the core's data plane result?**
  - Core label switches ingress PE to egress PE
  - Core does not need end-to-end information

Q&A

MPLS Components Overview

In This Section

- What is MPLS?
- Why Use MPLS?
- MPLS Device Roles
- MPLS Device Operations
- MPLS Label Distribution Protocols

What is MPLS?

- Multiprotocol Label Switching
- Originally Cisco proprietary
  - Previously called “tag switching”
- Now an open standard
  - RFC 3031 - Multiprotocol Label Switching Architecture

What is MPLS? (cont.)

- Name implies two important parts
  - Multiprotocol
  - Label switching

What is MPLS? (cont.)

- Multiprotocol
  - Can transport different payloads
- Layer 2 payloads
  - Ethernet, Frame Relay, ATM, PPP, HDLC, etc.
- Layer 3 payloads
  - IPv4, IPv6, etc.
- Extensible for new future payloads

What is MPLS? (cont.)

- Label Switching
  - Switches traffic between interfaces based on locally significant label values
- Similar to a legacy virtual circuit switching
  - Frame Relay input/output DLCI
  - ATM input/output VPI/VCI

Why Use MPLS?

- Transparent tunneling over SP network
- BGP free core
  - Saves routing table space on Provider (P) routers
- Offer L2/L3 VPN service to customers
  - No need for overlay VPN model
  - Much more on this later…

Why Use MPLS? (cont.)

- **Traffic engineering**
  - Distribute load over underutilized links
  - Give bandwidth guarantees
  - Route based on service type
  - Detect and repair failures quickly
    - i.e. Fast Reroute (FRR)

MPLS Label Format

- RFC 3032 - MPLS Label Stack Encoding
- 4 byte header used to “switch” packets
  - 20 bit Label = Locally significant to router
  - 3 bit EXP = Class of Service
  - S bit = Defines last label in the label stack
  - 8 bit TTL = Time to Live

How Labels Work

- MPLS Labels are bound to FECs
  - Forwarding Equivalency Class
  - IPv4 prefix for our purposes
- Router uses MPLS LFIB instead of IP routing table to switch traffic
- Switching logic
  - If traffic comes in if1 with label X send it out if2 with label Y

MPLS Device Roles

- MPLS network consists of three types of devices
  - Customer Edge (CE)
  - Provider Edge (PE)
  - Provider (P)

CE Devices

- Customer Edge (CE)
- Last hop device in customer's network
  - Connects to provider's network
- Can be layer 2 only or layer 3 aware
- Typically not MPLS aware

PE Devices

- Provider Edge (PE)
  - Previously called Label Edge Routers (LER)
- Last hop device in provider's network
  - Connects to CE and Provider (P) core devices
- PE performs both IP routing & MPLS lookups

PE Devices (cont.)

- For traffic from customer to core…
  - Receives unlabeled packets (e.g. IPv4)
  - Adds one or more MPLS labels
  - Forwards labeled packet to core

PE Devices (cont.)

- For traffic from core to customer…
  - Receives MPLS labeled packets
  - Removes one or more MPLS labels
  - Forwards packet to customer

P Devices

- Provider (P)
  - Previously called Label Switch Router (LSR)
- Core devices in provider's network
  - Connects to PEs and/or other P routers
- Switches traffic based only on MPLS labels

MPLS Device Operations

- PE & P routers perform three major MPLS operations
- Label push
  - Add a label to an incoming packet
  - AKA label imposition
- Label swap
  - Replace the label on an incoming packet
- Label pop
  - Remove the label from an outgoing packet
  - AKA label disposition

Label Distribution

- Labels are advertised via a Label Distribution Protocol
- Label Distribution Protocol (LDP)
  - Advertises labels for IGP learned routes
  - RFC 5036 - LDP Specification
- MP-BGP
  - Advertises labels for BGP learned routes
  - RFC 3107 - Carrying Label Information in BGP-4
- RSVP
  - Used for MPLS Traffic Engineering (MPLS TE)
  - RFC 3209 - RSVP-TE: Extensions to RSVP for LSP Tunnels

Q&A

Label Distribution Protocol (LDP)

In This Section

- Label Distribution Protocol (LDP) Operations
  - Discovering LDP Neighbors
  - Forming LDP Adjacencies
  - Advertising Labels
- LDP Configuration
- LDP Verification

Discovering LDP Neighbors

- Like IGPs, LDP automatically discovers neighbors with a Hello protocol
- Hellos use multicast UDP
  - 224.0.0.2 – "All Routers Multicast"
  - UDP src & dst port 646
- Hello include "IPv4 Transport Address"
  - Address to use for the TCP session
  - Defaults to the LDP Router-ID

Forming LDP Adjacencies

- LDP sessions are formed reliably over TCP
  - Unicast between transport addresses
  - TCP port 646
- Implies peers must have routes to each other's transport addresses
  - E.g. their Loopbacks
- Transport addresses could be modified, but not normally needed
  - Like IGP, hardcoding Router-ID is recommended

Advertising Labels

- Once LDP session is established, Label is advertised for FEC
  - I.e. Label to IPv4 prefix mapping
- Label distribution can be implicit or explicit
  - Unsolicited Downstream vs. Downstream on Demand
  - Depends on implementation & config options
- Labels could be advertised for some or all routes
  - Cisco default is all IGP routes
  - Really only /32 Loopback matters
    - More on this later…

IOS LDP Configuration

- Enable CEF
  - Should already be on by default
- Agree on label protocol
  - mpls label protocol
  - Should already be LDP by default
- Recommended to define Router-ID
  - mpls ldp router-id
- Enable LDP
  - Interface mpls ip
  - IGP process mpls ldp autoconfig

IOS XR LDP Configuration

- Enable global LDP process
  - mpls ldp
- Recommended to define Router-ID
  - router-id
- Enable LDP
  - Interface under **mpls ldp**
  - IGP process **mpls ldp auto-config**

IOS LDP Verification

- Verify LDP is enabled
  - show mpls interfaces
- Verify LDP sessions
  - show mpls ldp neighbor
- Verify LFIB
  - show mpls forwarding-table
- Troubleshooting LDP Adjacencies
  - debug mpls ldp transport events

LDP Verifications

- Verify LDP is enabled
  - show mpls interfaces
- Verify LDP sessions
  - show mpls ldp neighbor
- Verify LFIB
  - show mpls forwarding
- Verify LIB
  - show mpls ldp bindings
- Verify LDP Transport
  - show mpls ldp discovery
- Troubleshooting LDP Adjacencies
  - debug mpls ldp transport events

Q&A

LDP Features

LDP Features

- LDP Authentication
- LDP IGP Synchronization
- LDP Session Protection
- Label Ranges
- Label Allocation Filtering
- Label Advertisement Filtering

LDP Authentication

- LDP uses TCP for transport
  - Implies that TCP Option 19 (MD5) can be used for authentication
  - Same logic as BGP
- Authentication is in the TCP header
  - Implies that neighbor address is the TCP transport address

LDP IGP Synchronization

- IGP & LDP should be on a 1:1 basis
  - If not, MPLS traffic can be black holed
- LDP IGP Sync attempts to prevent this
  - If LDP is broken, avoid IGP routing on the link
  - Only works if there is an alternate IGP path

LDP Session Protection

- When LDP neighbors go down, labels are flushed
  - Flushed from both LFIB and LIB
- When neighbors come back, labels are re-learned
  - Added to the LIB, then possible the LFIB
- Short flaps can cause add/withdraw/add/withdraw/etc.
- LDP Session Protection prevents this by…
  - Caching information in the LIB when neighbors go down
  - Flushing labels after protection timer expires

Other Misc. LDP Features

LDP Label Ranges

- Global label range can be statically defined
- Needed for static label assignments

Label Allocation Filtering

- Which labels do I want to originate? All? Loopbacks?

Label Advertisement Filtering

- Which labels do I want to advertise?

Q&A

Q&A