MPLS 101

MPLS and VPLS fundamentals

Nick Buraglio

http://www.forwardingplane.net/

6/6/2013

Purpose and Scope

- The purpose of this slide deck is to be a bootstrap and reference for terminology and concepts for the umbrella suite of protocols within MPLS
- It is not meant to be an operational guide or to be a definitive resource but instead a compilation of a superset of references condensed into a portable, convenient format

Overview

- MPLS is broken up into transport and services.
  - L3VPN
  - L2VPN
  - VPLS
- MPLS as a transport is plumbing for the MPLS service to utilize.

Concepts - MPLS

- Label Switching
  - Utilizes labels to switch much faster than layer 3 routing
  - Happens at data link layer
  - Similar to ATM and Frame Relay, which both use label switching at their core
  - Permits finer granularity resource allocation to traffic streams
  - MPLS is an open standard (RFC 3031)

Concepts - MPLS

- Label Switching is similar to:
  - BGP communities (in format)
  - Frame Relay DLCI
  - ATM VPI/VCI

Concepts - MPLS

- MPLS
  - Locates shortest paths based on labels rather than longest network match
  - Avoids lookups in routing table
  - Labels identify virtual links as opposed to endpoints
  - Data Link layer agnostic. MPLS can function over ATM, Ethernet, Frame Relay, SONET, etc.
  - More like switching than routing
  - Sometimes referred to as layer 2.5
  - Packet switching technology

Concepts - MPLS

- MPLS
  - Lookups happen in the switch fabric rather than the CPU
  - Creates a faster lookup time and lower hardware overhead*

- *RIB lookups are significantly faster then they were at the inception of MPLS

Functionality

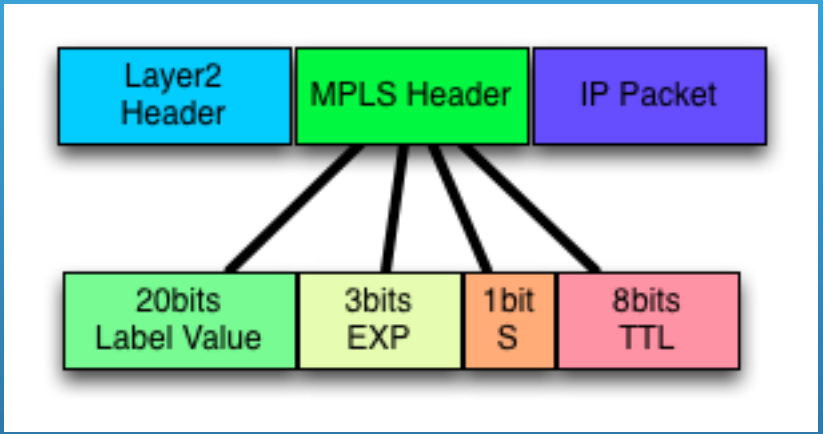
- MPLS
  - Only necessary to do one IP lookup
    - Lookup occurs when the packet first enters the MPLS domain for the duration of the transit of the packet
    - Lookup for the last hop
  - Adds a prefix to packets with an MPLS header, containing one or more labels called a label stack

Functionality

- Each label stack entry contains four fields:
  - A 20-bit label value.
  - A 3-bit *Traffic Class* field for Quality of Service (QoS) priority (experimental) and Explicit Congestion Notification (ECN)
  - A 1-bit *bottom of stack* flag. If this is set, it signifies that the current label is the last in the stack
  - An 8-bit time to live field

Functionality

- MPLS Header



Terminology

- LSP – Label switched path (similar to PVC in ATM but not dependent on L2 technologies)
- PE – Provider Edge Router
  - Sometimes referred to as Label Edge Router (LER)
- P Router – routers that function specifically as transit devices
  - P routers simply transit traffic and can be significantly less complicated and robust*
  - Also known as the Label Switch Router (LSR)
- CE – Customer Edge device.
  - Last hop into customer network
  - Generally not MPLS aware

Terminology

- LIB – Label information base
  - Software table maintained by each router for storage of learned and local port and corresponding label to be popped or pushed on incoming or outgoing packets
  - Used by label distribution protocol (LDP)
  - Used by LDP for mapping next hops
  - Is a function of the MPLS control plane

Terminology

- LFIB - Label Forwarding Information Base
  - FIB for MPLS label forwarding
  - Used for managing forwarding in which destinations and incoming labels are associated with outgoing interfaces and labels

Terminology

- P Router / LSR – Label switch router
  - Routers that perform routing based only on the label
- PE Router / LER – Label Edge Router
  - The entry and exit points of an MPLS network
  - *push* an MPLS label onto an incoming packet and *pop* it off the outgoing packet

Terminology

- LDP – Label Distribution Protocol
  - Protocol for distribution of labels between LERs and LSRs
- Label *swap*
- *Label push* (*impose*)
- *Label pop* (*dispose*)
- pseudo-wire (PW, VLL)
  - An emulation of a point-to-point connection over a packet switched network

Terminology

- LDP
  - Uses UDP multicast 224.0.0.2:646 to discover neighbors
  - Uses TCP port 646 once neighbors are discovered
  - Requires IGP (ISIS, OSPFv2, OSPFv3) to function
  - Advertises labels for IGP learned routes
  - Routers running MPLS distribute their labels for each prefix to their neighbors
  - Each router only knows about its neighbor's labels
    - Each router does not have a full view of the network

Terminology

- RSVP - Resource Reservation Protocol
  - Transport layer protocol for reservation of resources across a network
  - Operates over IPv4 and IPv6
  - Does not transport application data
  - Similar to a control protocol like ICMP or IGMP
  - Can be used to request or or deliver QoS for data streams or flows
  - RSVP-TE is more common today and is an extension of RSVP

Terminology

- RSVP-TE - Resource Reservation Protocol - Traffic Engineering
  - Extension of the RSVP protocol for the purpose of traffic engineering
  - Supports reservation of resources over an IP based network
  - Allows for the establishment of LSPs taking into consideration network constraint parameters such as available bandwidth and explicit hops
  - Operationally RSVP-TE is more more robust and scalable

Terminology

- RSVP-TE
  - Operationally RSVP-TE is more more robust and scalable than LDP
  - Operational complexity can be higher with RSVP-TE due to it’s more feature rich nature

Terminology

- FEC – Forwarding equivalent class
  - Characteristics for determining the FEC of a higher-layer packet
  - Depend on the configuration of the router
  - Typically this is the destination IP address at a minimum
  - Quality of service class is also often used.
  - A Forward Equivalence Class generally corresponds to a label switched path (LSP)
  - The reverse is not normally true
  - LSPs may be (and usually are) used for multiple FECs

Functionality

- Topmost label is examined when a packet enters an LSR
- Actions based on topmost label and pre-built tables allow for very fast operations
- Labels may have pre-built label tables
- Label is inserted between the Layer 2 header and the Layer 3 header

Operation

- Label Operation
  - Label Push – adds a label to incoming packet; also called label imposition
  - Label Swap – replaces the label on an incoming packet
  - Label Pop – removes the label on an outgoing packet; also called label disposition
  - Penultimate Hop Popping (PHP)
    - Without PHP, the egress PE router must perform two lookups. First it has lookup the destination prefix associated with the label, and then once it has the prefix; the router must look up the next-hop for the destination prefix. With PHP, the next-to-last hop router (the penultimate router) will perform the label pop operation. This means the egress PE router will not have to do the label lookup because it is receiving an IP packet (the penultimate router has popped the label off before forwarding it to the egress router).

VPLS

Concepts - VPLS

- VPLS - Virtual Private LAN Service
  - Also called L2VPN
  - Is a mechanism for providing Ethernet based multipoint to multipoint communication over IP/MPLS networks
  - Allows for any-to-any connectivity

Concepts - VPLS

- In a VPLS environment the LAN in question at each site is extended to the edge of the provider network
- Service provider network (SP) emulates a bridge / switch to create a single bridged network
- Designed for applications that require multipoint or broadcast access

Functionality - VPLS

- VPLS emulates a LAN, full mesh connectivity is required between all devices
- Two methods for full mesh establishment for VPLS
  - BGP
    - Provides auto discovery as well as signaling
  - LDP

Functionality – VPLS

- VPLS MPLS packets have a two-label stack
  - Outer label used for normal MPLS forwarding within the service provider's network
  - If BGP signaled, the inner label is allocated by a PE as part of a label block
  - If LDP signaled, the inner label is a virtual circuit ID assigned by LDP when mesh is first established between participating PEs.
  - Every PE keeps track of assigned inner label, and associates these labels with the VPLS instance.

Functionality – VPLS

- VPLS requires a full mesh in both the control and data planes
  - BGP better for large scale due to better scaling utilizing route reflectors
  - Scaling LDP over a large amounts of VPLS sites generally uses a hierarchical design (HVPLS)

Use Cases

- Transparent tunneling over existing networks
- BGP free core
- Layer2 VPN services
- Layer3 VPN services

Use Cases

- Traffic Engineering
  - Redistribution of traffic load over underutilized links
  - Bandwidth guarantees / CIR
  - Fast failure recovery

References

- http://networkstatic.net/juniper-and-cisco-comparisons-of-rib-lib-fib-and-lfib-tables/
- http://convergingontheedge.com/post/16916353412/mpls-primer-for-the-interested
- http://www.wikipedia.org/
- https://www.juniper.net/us/en/
- http://www.cisco.com/
- http://mellowd.co.uk/ccie/
- http://blog.ioshints.info/