MPLS OAM Tutorial

Sam K Aldrin

sam.aldrin@huawei.com

Agenda

- **Introduction**
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

What is OAM

- Means different things to different people and organizations.
- Worst, some times it means different things to different people within the same organization
- IETF standardized the meaning of OAM within the IETF
  - June 2011, RFC 6291

IETF definition of OAM

- **O**perations: Operational activities to keep network up and running. *E.g. Monitoring, finding faults*
- **A**dministration: Involves keeping track of network resources. *E.g. Bookkeeping, (available ports, BW)*
- **M**aintenance: Involves repair and upgrades. *E.g. Software upgrades, configurations, corrective and preventive measures.*

Scope of the Tutorial

- Today's presentation mainly focus on IETF defined Operations aspects of MPLS OAM.
- Various OAM operations and techniques are presented for MPLS networks

Agenda

- Introduction
- **Terms and Terminology**
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

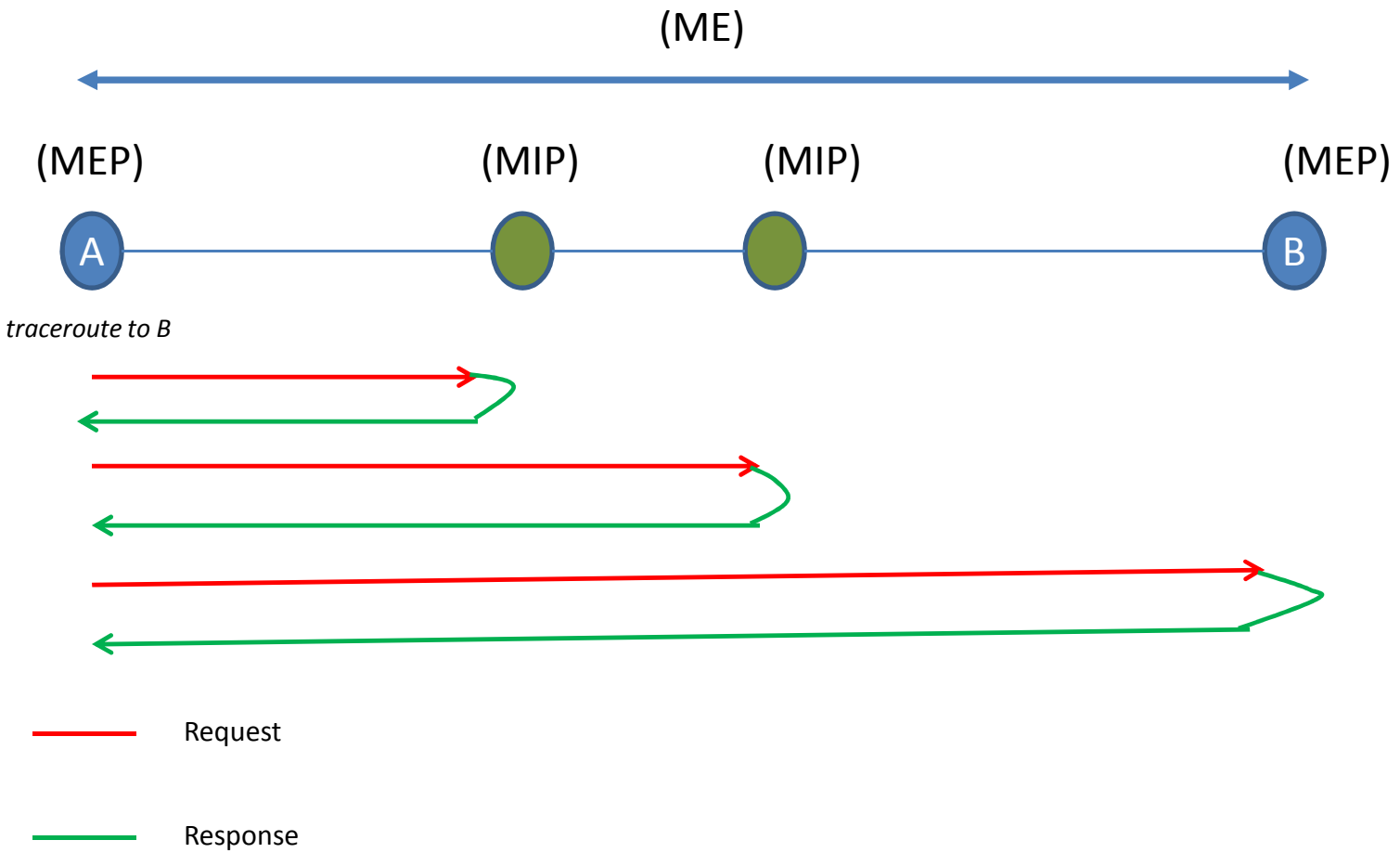
Important Terminologies

- Before we dive deeper, it is important to understand some of the terminologies and their meanings
- What are they ?
  - Various organizations (IEEE, ITUT, IETF) all have their version
  - We will discuss here selected set of definitions from RFC 5860, RFC 6371 and draft-ietf-opsawg-oam-overview-05
- Good understanding of these Terminologies will help us to appreciate modern OAM protocols better.

Important Terminologies

- Maintenance Point (MP)
  - Is a functional entity that is defined within a node that either initiate or react to a OAM message
- Maintenance Entity (ME)
  - Point to Point relationship between two MP
  - In MPLS this is LSP, In BFD this is session
- Maintenance Point can be either MEP or MIP
  - Maintenance End Point (MEP)
    - Can either initiate or react to OAM Messages
    - MEP are the two end points of the ME
  - Maintenance Intermediate Point (MIP)
    - Is an intermediate MP between two MEP
    - It can only respond to OAM messages

Relationship of MP



Important Terminologies (contd..)

- Continuity Check
  - Ability of endpoint to monitor liveliness of a path (BFD)
- Connectivity Verification
  - Ability of an endpoint to verify it is connected to a specific endpoint. (BFD,Ping)
- Route Tracing
  - This is also known as path tracing, allows to identify the path taken from one MEP to another MEP (traceroute)
- Fault Verification
  - Exercised on demand to validate the reported fault. (Ping)
- Fault Isolation
  - Localizing and isolating the failure domain/point (traceroute)
- Performance
  - Includes Packet Loss Measurements and Packet Delay Measurements
  - E.g. IP Performance Metrics (IPPM) (RFC 2330)

Agenda

- Introduction
- Terms and Terminology
- **An Introduction to Tools**
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

Ping

- Ping refers to tools that allows to detect liveliness of a remote host
- Most commonly known Ping is based on ICMP Echo Request and Response
- Security policies and firewalls sometimes prevent forwarding of ICMP messages.
- UDP/TCP version of the Ping has surfaced to circumvent barriers introduced by security policies and Firewalls on ICMP Echo Requests
  - RFC 4379 use UDP port 3503 for LSP Ping
- Different implementations of Ping has different options

Ping – traceroute simulation

- Ping an IP address with increasing the TTL count at each step.
- In the example below TTL increased by 1 at each iteration..

ping -c 1 -t 2 -n www.yahoo.com

PING any-fp3-real.wa1.b.yahoo.com (98.139.127.62) 56(84) bytes of data.

From **10.35.78.17** icmp_seq=0 *Time to live exceeded*

--- any-fp3-real.wa1.b.yahoo.com ping statistics ---

1 packets transmitted, 0 received, +1 errors, 100% packet loss, time 0ms, pipe 2

ping -c 1 -t 3 -n www.yahoo.com

PING any-fp3-real.wa1.b.yahoo.com (98.139.127.62) 56(84) bytes of data.

From **10.34.159.13** icmp_seq=0 *Time to live exceeded*

--- any-fp3-real.wa1.b.yahoo.com ping statistics ---

1 packets transmitted, 0 received, +1 errors, 100% packet loss, time 0ms, pipe 2

Traceroute

- Design to trace the path taken from a node A to a node B.
- Probe packets are generated with monotonically increasing TTL value
  - Forcing ICMP TTL expiry message from each intermediate node.
  - In Linux Echo request packet is UDP (default destination port is UDP:33434)
  - In some other platforms it can be ICMP Echo request.

traceroute sample output linux

traceroute -n 10.35.78.17

traceroute to 10.35.78.17 (10.35.78.17), 30 hops max, 46 byte packets

1 10.35.75.3 0.292 ms 0.366 ms 0.213 ms TTL=1 ⟸

2 10.35.78.17 0.642 ms 0.429 ms 0.369 ms TTL=2 ⟸

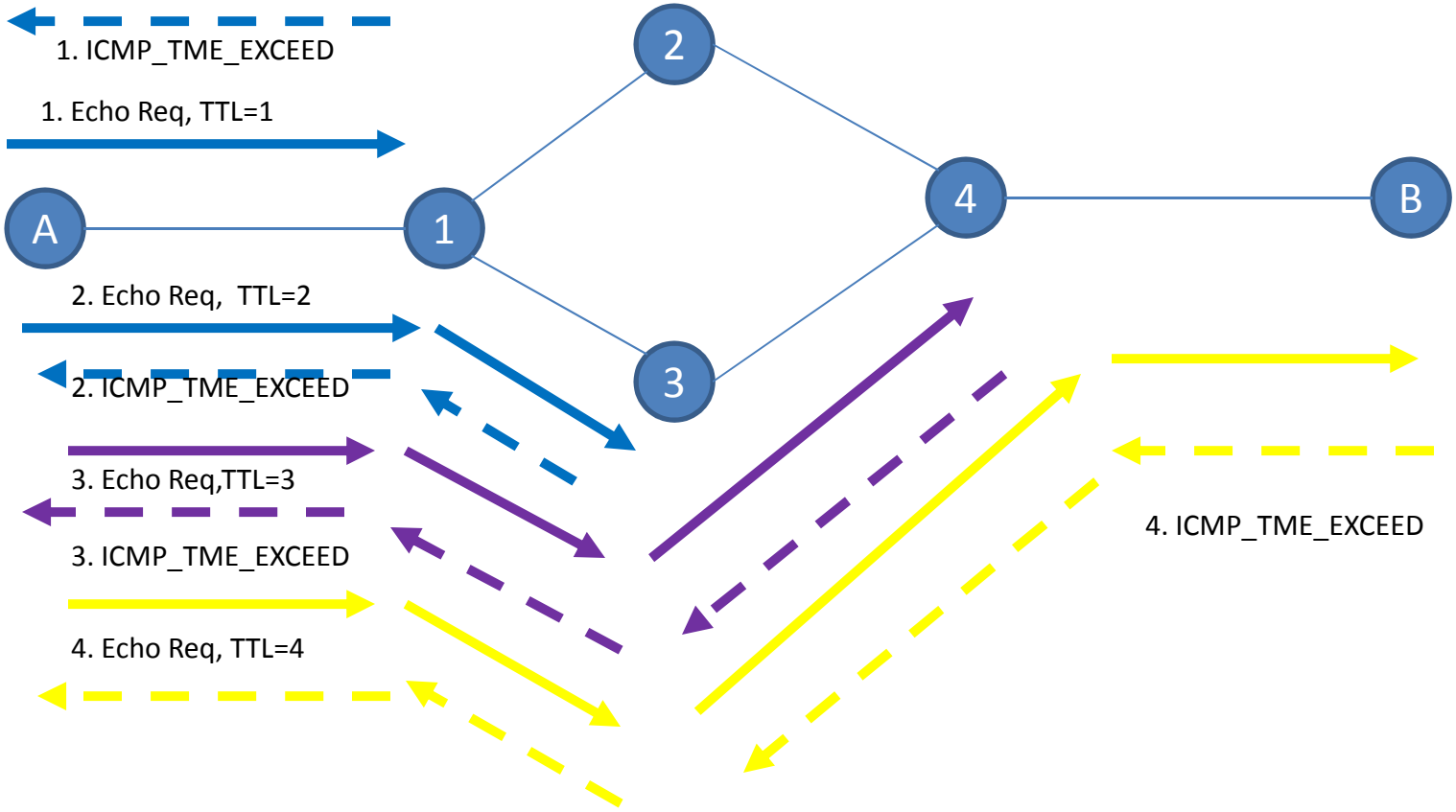
traceroute -n **-I** 10.35.78.17

traceroute to 10.35.78.17 (10.35.78.17), 30 hops max, 46 byte packets

1 10.35.75.3 0.271 ms 0.219 ms 0.213 ms TTL=1 ⟸

2 10.35.78.17 0.442 ms 0.265 ms 0.351 ms TTL=2 ⟸

traceroute


1. ICMP_TME_EXCEED
1. Echo Req, TTL=1
A
1
2
3
4
B
2. Echo Req, TTL=2
2. ICMP_TME_EXCEED
3. Echo Req,TTL=3
3. ICMP_TME_EXCEED
4. Echo Req, TTL=4
4. ICMP_TME_EXCEED

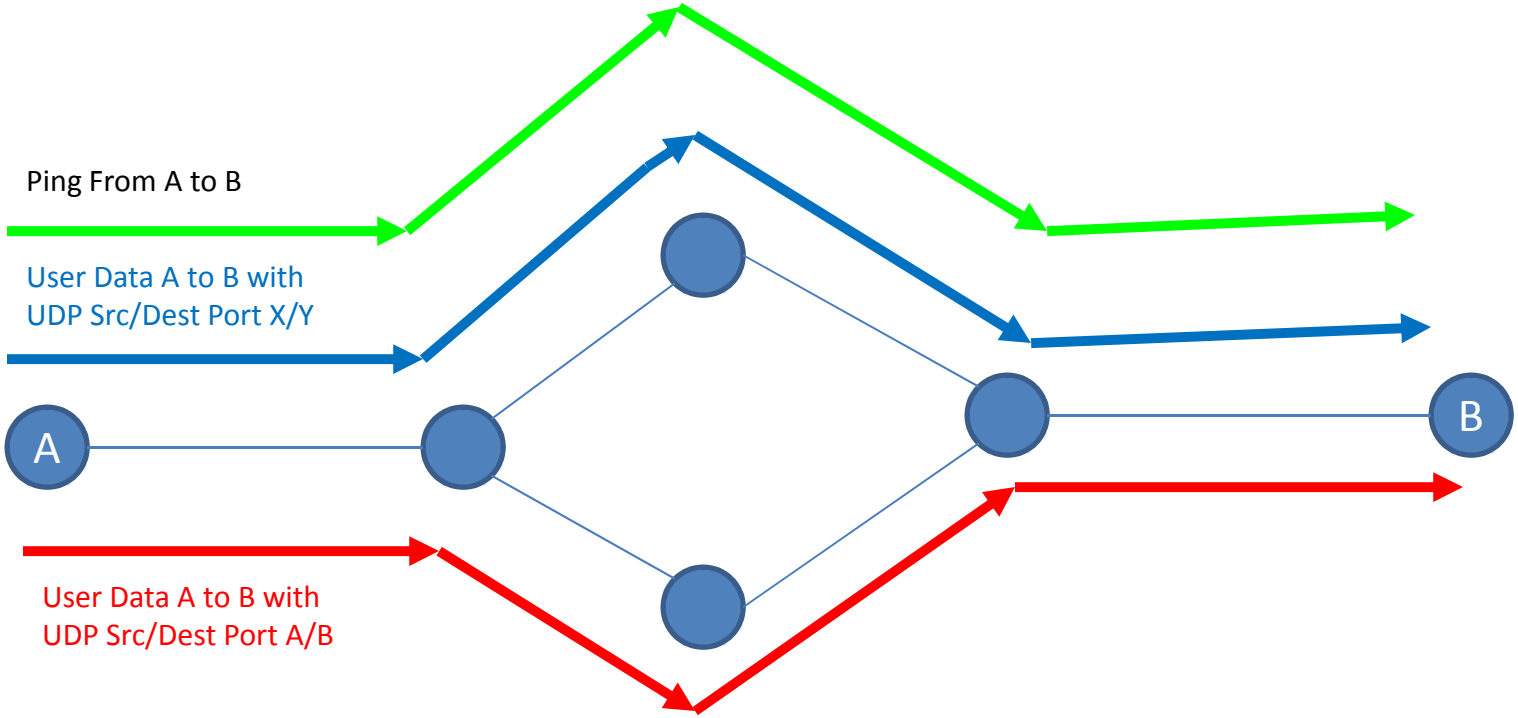
Challenges

- Over the years networking has evolved with that comes OAM challenges
  - ECMP (Equal Cost Multi Path)
  - Multicast
  - Tunneling (MPLS, PW, VPN, TRILL)
  - Firewalls
- ICMP and more traditional OAM are designed for unicast traffic with single path to the destination.

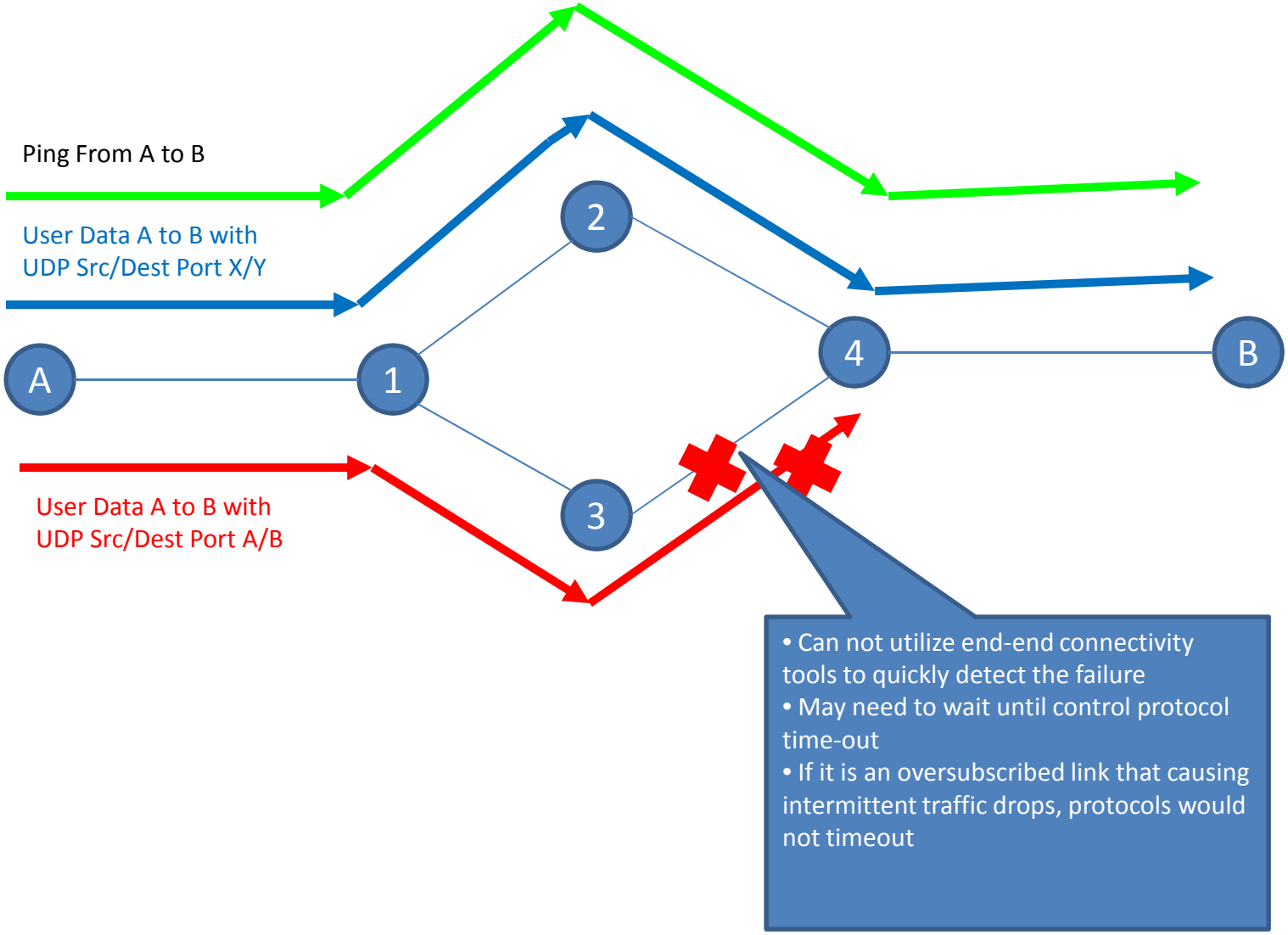
Equal Cost Multipath

- Equal Cost Multi Path (ECMP) allows
  - Protection against failures
  - Increased overall end-end BW
  - ECMP is becoming increasingly popular
- Devices typically use fields in the MAC or IP header to select the forwarding path among multiple equal cost paths
- Connectivity and Continuity verification messages MUST follow the same path as user data.
  - How can we accomplish this ?
  - There is no standard way of doing this in IP world
  - MPLS RFC 4379 has payload discovery approach

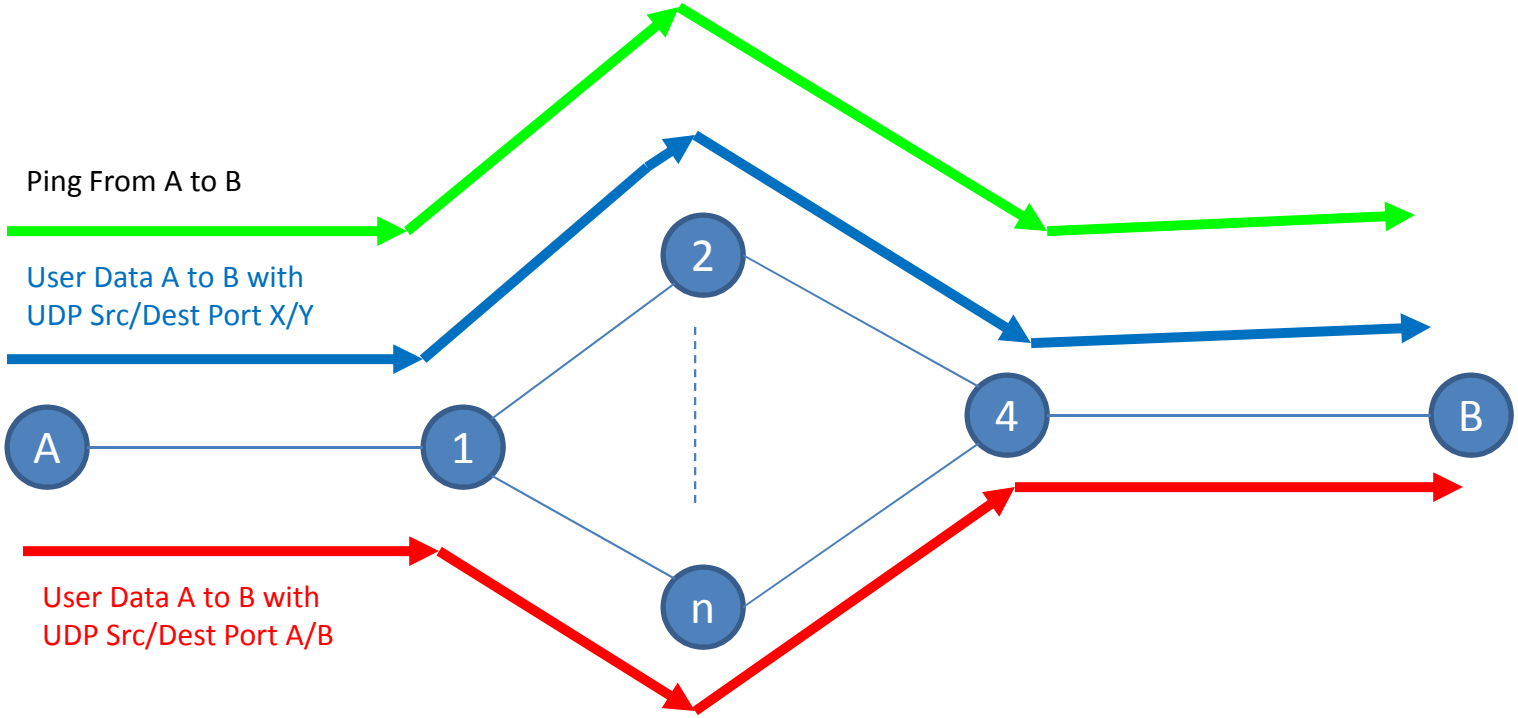
ECMP


Ping From A to B
User Data A to B with
UDP Src/Dest Port X/Y
A
B
User Data A to B with
UDP Src/Dest Port A/B

ECMP Failure Example



ECMP Monitoring Challenges



Challenges:

- Ingress Node (A) may not even know how many ECMP from intermediate node (1)
- Monitoring probes SHOULD take the same path as the normal data
- Different vendors utilize different hash algorithms in selection ECMP paths

ECMP challenges

- Conclusion
  - No standard method to exercise end-end continuity and connectivity verifications that covers all of the ECMP in IP networks

Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- **Introduction to MPLS**
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

What is MPLS

- MPLS is acronym for Multi Protocol Label Switching
- Forwards traffic using labels
- Provides virtual connection (LSP) within the network
- Labels are allocated based on FEC
- Different types of label distribution
- An LSP is usually unidirectional
- Ingress, Transit and Egress router types
- Traditional MPLS networks support PHP processing
- Supports different traffic types like ATM, FR, IP etc
- Private services like VPN for scalable service provider requirements

MPLS LSP signaling protocols

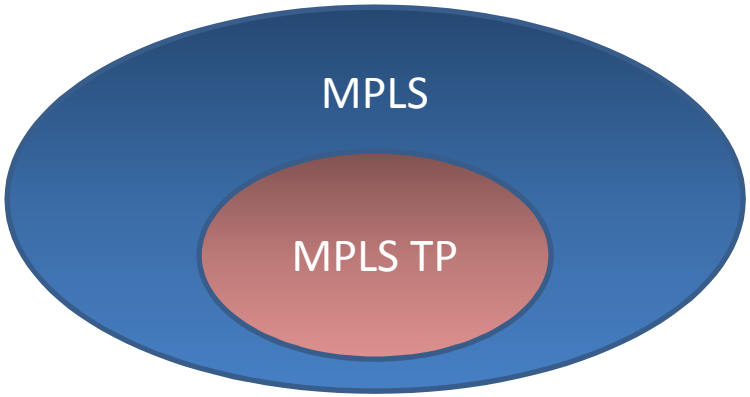
- Resourced Reservation Protocol (RSVP)
  - Extended to support Traffic Engineering
  - Labels are assigned for identified path
  - Explicit bandwidth reservation and paths
- Label Distribution Protocol (LDP)
  - Labels are exchanged between neighbors
  - IGP identifies the shortest path
- Constrained Routing LDP (CR-LDP)
  - Traffic Engineering support using LDP

Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- **MPLS TP 101**
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

What is MPLS-TP

RFC 5654

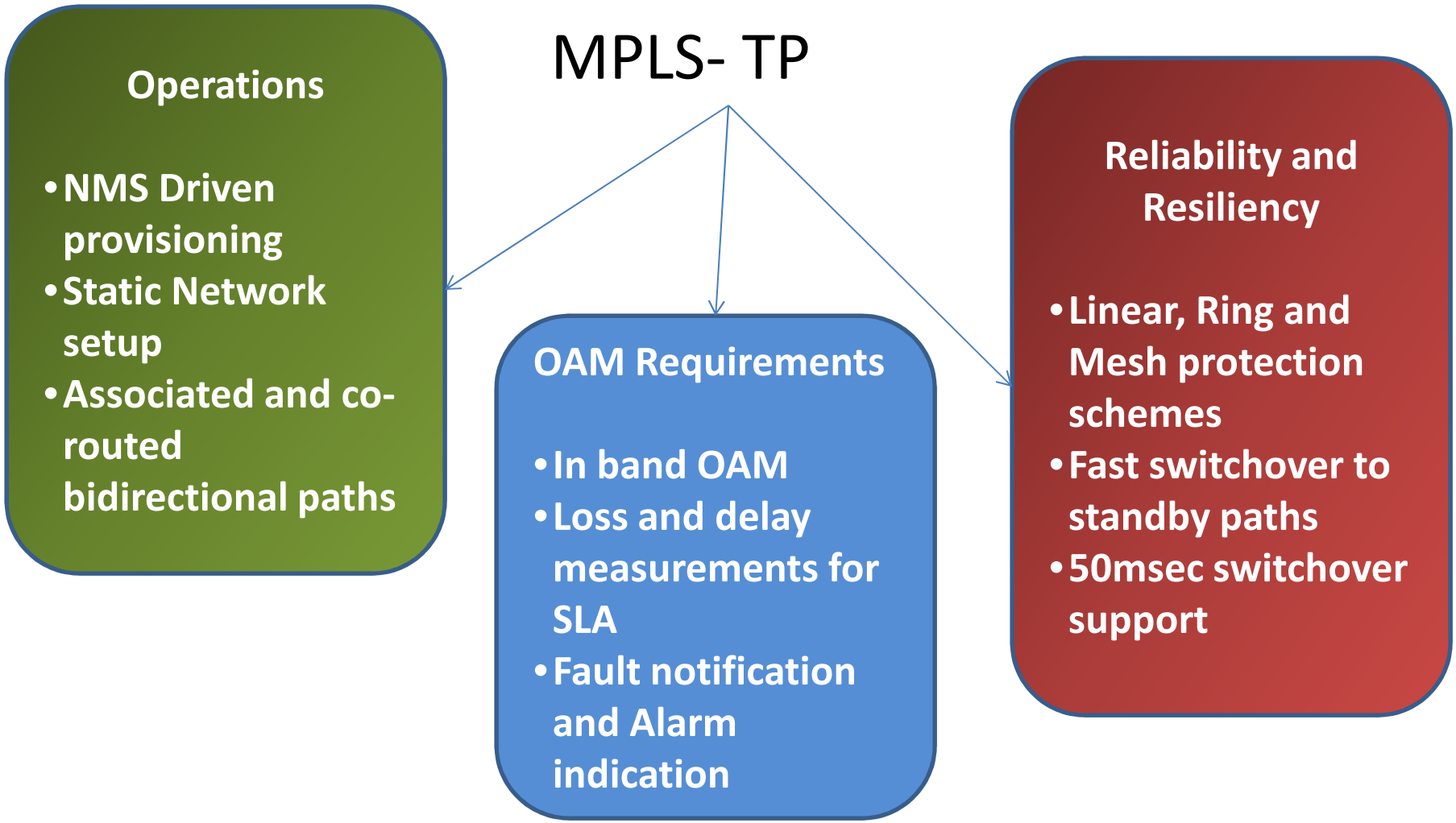


- MPLS TP is a subset of MPLS
- MPLS network enhanced to support Transport requirements
- Bidirectional LSP's with a highly reliable protection schemes
- Inter-op with existing MPLS Technologies
- Transport agnostic protocol extensions

What is being solved by MPLS-TP?

- Next Generation networks are moving
  - SONET/SDH to Packet Switching
  - Bandwidth hungry
  - Lower cost with network resource sharing
- OPEX and CAPEX
  - Provisioning of paths
  - OAM capabilities
  - Fault detection and recovery mechanisms
  - Path computation
  - SLA requirements

MPLS TP – New additions to MPLS



Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- **Troubleshooting MPLS**
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

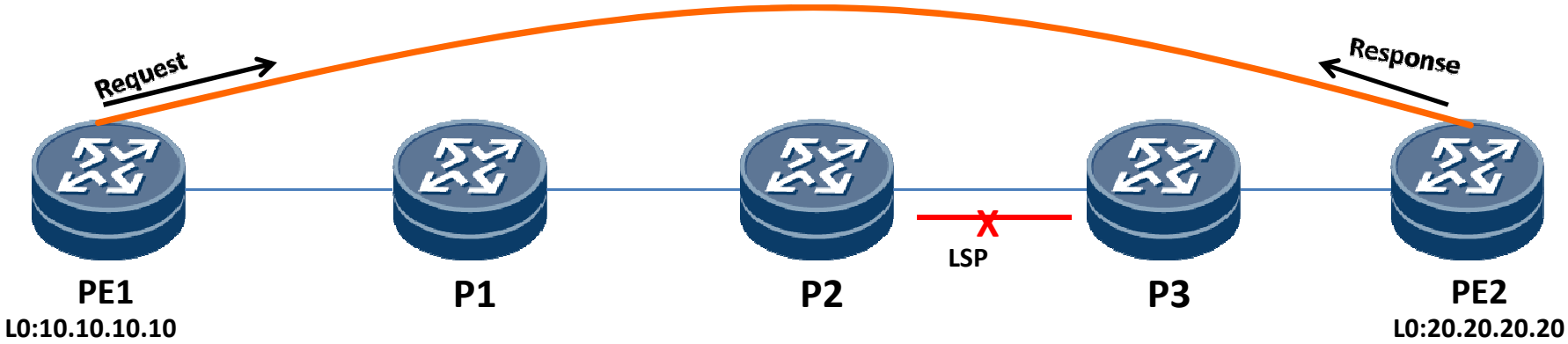
Problems in MPLS Networks

- Control Plane is working, Data Plane is broken
- IGP working but MPLS control protocol is broken
- Proactive monitoring of End-to-End MPLS LSP's
- Identifying the End-to-End packet path
- Unlabelled interface
- MTU issues
- Performance degradation and unable to provide QoS
- Black holes
- ECMP Verification

Primitive Debugging Methods

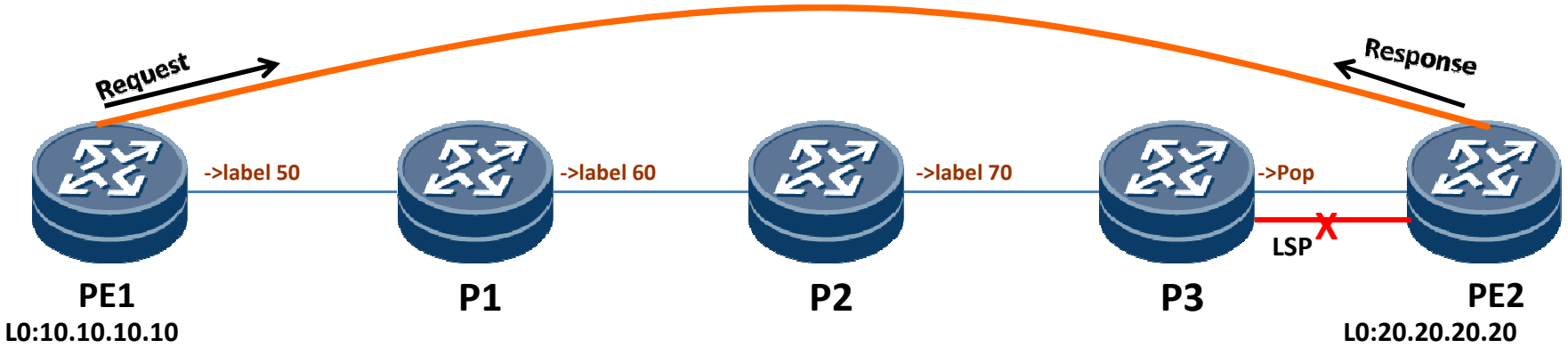
- ICMP provides connectivity verification
- VRF aware ping could test VPN path connectivity
- UDP ping could test the UDP transport
- Route table and Label table provides label entries programmed
- Interface status verification
- MPLS control plane protocols provides control plane information

ICMP ping



- ICMP ping emulates the data but can only verify IP layer
- It cannot verify if MPLS path is broken but IP is working
- It cannot verify ECMP
- It cannot validate control plane to data plane
- It cannot verify various MPLS control plane protocols
- It cannot verify for unlabelled interface, black-holes, control plane to data plane mismatch, etc.

VRF aware ping



- VRF aware could emulate VPN traffic
- Could test VPN connectivity
- Cannot detect LSP breakage
- If IP connectivity is working and MPLS is broken, it cannot detect
- Can detect if there is no label path, but not in all cases
- Cannot detect ECMP failures, CP to DP mismatch, etc.

Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- **MPLS OAM**
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

What is MPLS OAM

- Operations, Administration and Maintenance of MPLS Networks
- Perform proactive and on-demand troubleshooting of MPLS Networks and devices
- Ability to measure MPLS network and aid user in managing the network
- Ability to diagnose defects which cannot be done at other layers or using non-MPLS specific toolset
- Provide carrier class tool set to manage MPLS networks

Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- **LSP Ping**
- ECMP troubleshooting
- BFD for MPLS
- Tools Galore

LSP ping

Requirements

- Detect LSP failures
- Detect label mismatch
- Detect CP to DP mismatch
- Pin point the failure
- Detect MTU failures

Applications

- Verify all MPLS FEC types
- Verify PE, P, MPLS TP devices
- Ability to verify MPLS VPN, TE, LDP, TP, P2MP, etc., LSP's.

Solution

- LSP ping to detect connectivity checks
- LSP ping based traceroute for path verification
- LSP ping based topology tree verification

Standards

- RFC4379 and all other extensions

LSP Ping – What is it?

Function

- Modeled like ICMP ping but based on UDP
- Connectivity between two end points of an LSP

Format

- Encapsulated like data frame for the FEC
- The IP destination of the packet is local host address

Behavior

- Cannot leak out onto non-MPLS interface
- Response packet contains a code indicating the reason
- Destination IP address used as entropy simulate ECMP
- OAM packets are treated the same as data packets
- TTL field is used to test intermediate hops

LSP Ping – What can it verify?

Sub-Type	Length	Value field
1	5	LDP IPv4 Prefix
2	17	LDP IPv6 Prefix
3	20	RSVP IPV4 Prefix
4	56	RSVP IPv6 Prefix
5		Not Assigned
6	13	VPN IPv4 Prefix
7	25	VPN IPv6 Prefix
8	14	L2 VPN endpoint
9	10	FEC 128 PW (Deprecated)
10	14	FEC 128 PW
11	16+	FEC 129 PW
12	5	BGP Labeled IPv4 Prefix
13	17	BGP Labeled IPv6 Prefix
14	5	Generic IPv4 Prefix
15	1	Generic IPv6 Prefix
16	4	Nil FEC

LSP Ping – Constructs

LSP ping packet is encapsulated to simulate data packet in order to test a LSP

- Two types – Echo Request and Echo Response
- The FEC to be verified
- The Label stack for the FEC/LSP
- A UDP/IP packet with LSP ping payload to be send on the LSP
- The interface information on which the packet has to be forwarded
- Forwarding and interface information for the FEC for verification purposes

LSP Ping – Response Codes

Value	Meaning
0	No return code
1	Malformed echo request received
2	One or more TLV's not understood
3	Replying router is egress for the FEC
4	No mapping for the FEC
5	DSMAP mismatch
6	Unknown upstream index
7	Reserved
8	Label switched at stack depth <RSC>
9	Label switched but no MPLS forwarding at stack depth <RSC>
10	Mapping for this FEC is not the given label at stack depth <RSC>
11	No label entry at stack depth <RSC>
12	Protocol not associated with interface at FEC stack depth <RSC>
13	Premature termination of ping due to label stack shrinking to a single label

LSP Ping – Echo Request

Echo Request is sent by the router to test LSP of a given FEC

MPLS encapsulation

- MPLS encapsulated IP/UDP packet
- Label stack is same as data packet for the FEC.
- Default TTL value for the label is 255
- FEC TLV contains the details of the FEC to be verified

IP Encapsulation

- IP/UDP Packet
- Source address: Valid source address
- Destination address: Local host address
- Destination Port: 3503
- RA option : Enable
- TTL : 1

LSP Ping – Echo Reply

Echo Reply is sent by the router to responding to the Echo Request

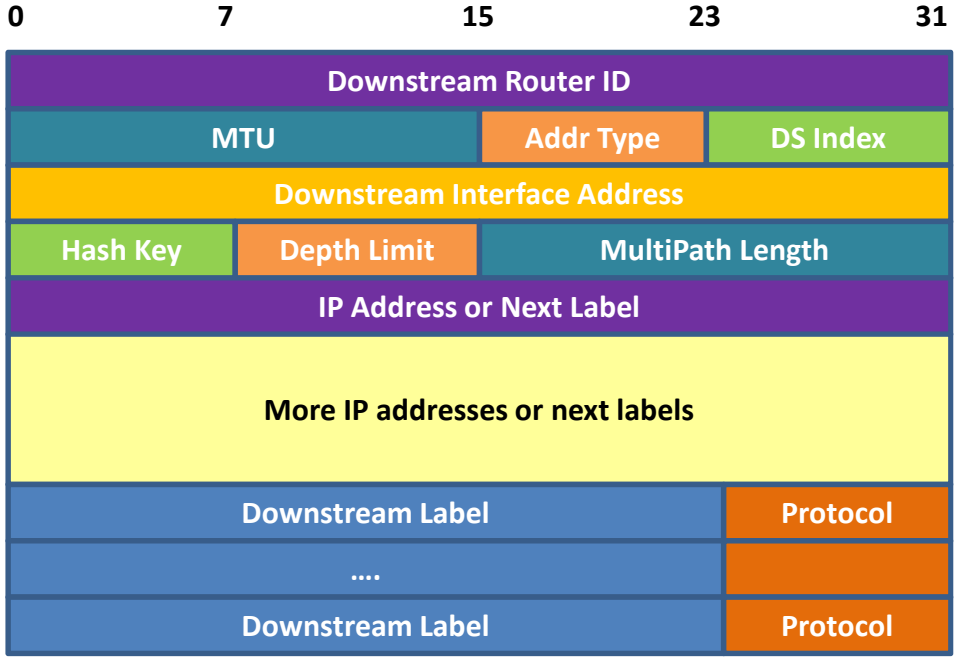
Reply Modes

- IP reply
- No Reply
- IP reply with RA option
- Control Channel

Packet Format

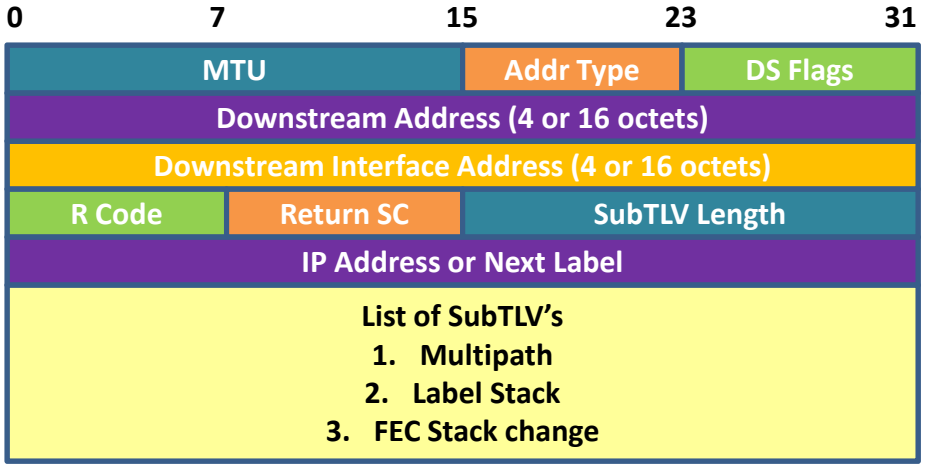
- IP source address : Replying router IP address
- Destination address : Received Source address
- Source port : 3503/other chosen port
- Destination Port : Port number in the echo request
- TTL : 255

Downstream Mapping



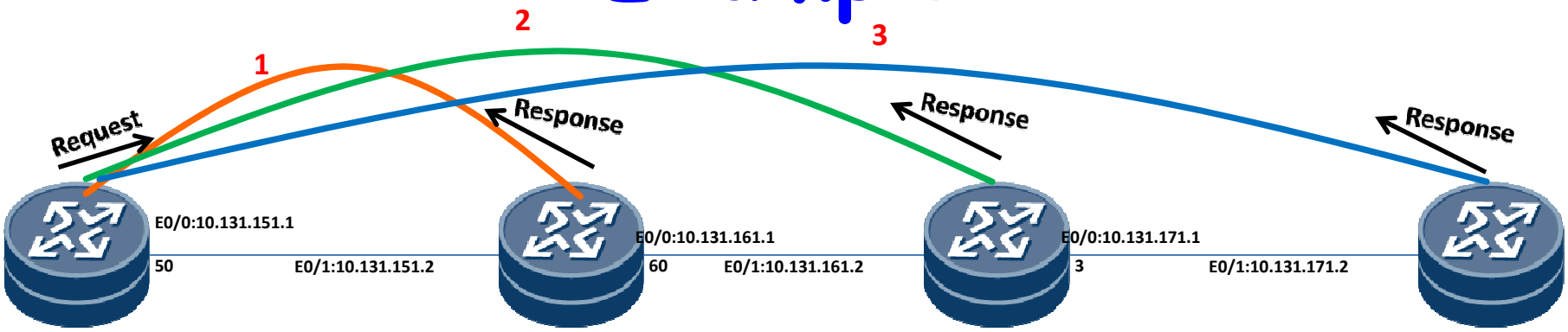
DSMAP TLV

- Downstream interface address is IP address of outgoing interface for the LSP
- Downstream label is the outgoing label for the LSP
- Protocol associated with the label
- DDMAP is enhanced version of the DSMAP TLV (Deprecated)



DDMAP TLV

Downstream Mapping TLV - Example



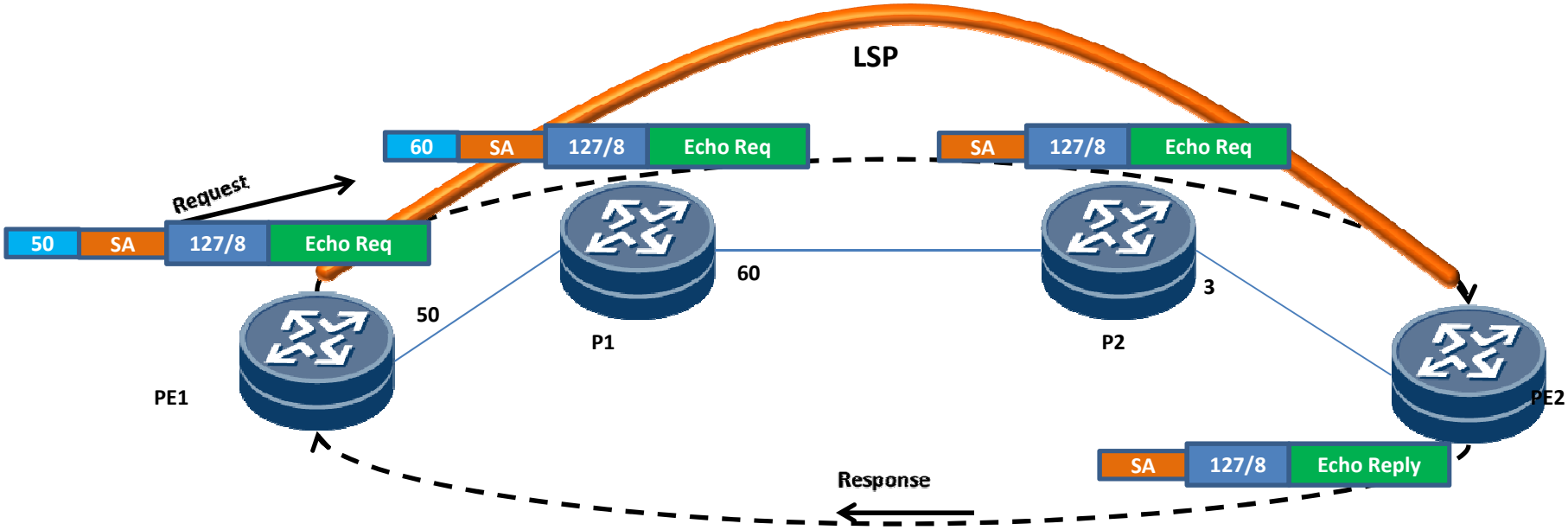
PE1 Downstream for PE2
MTU: MRU of E0/0
Address type: 1
DS Intf Addr: 10.131.151.1
DS Label : 50

P1 Downstream for PE2
MTU: MRU of E0/0
Address type: 1
DS Intf Addr: 10.131.161.1
DS Label : 60

P2 Downstream for PE2
MTU: MRU of E0/0
Address type: 1
DS Intf Addr: 10.131.171.1
DS Label : 3

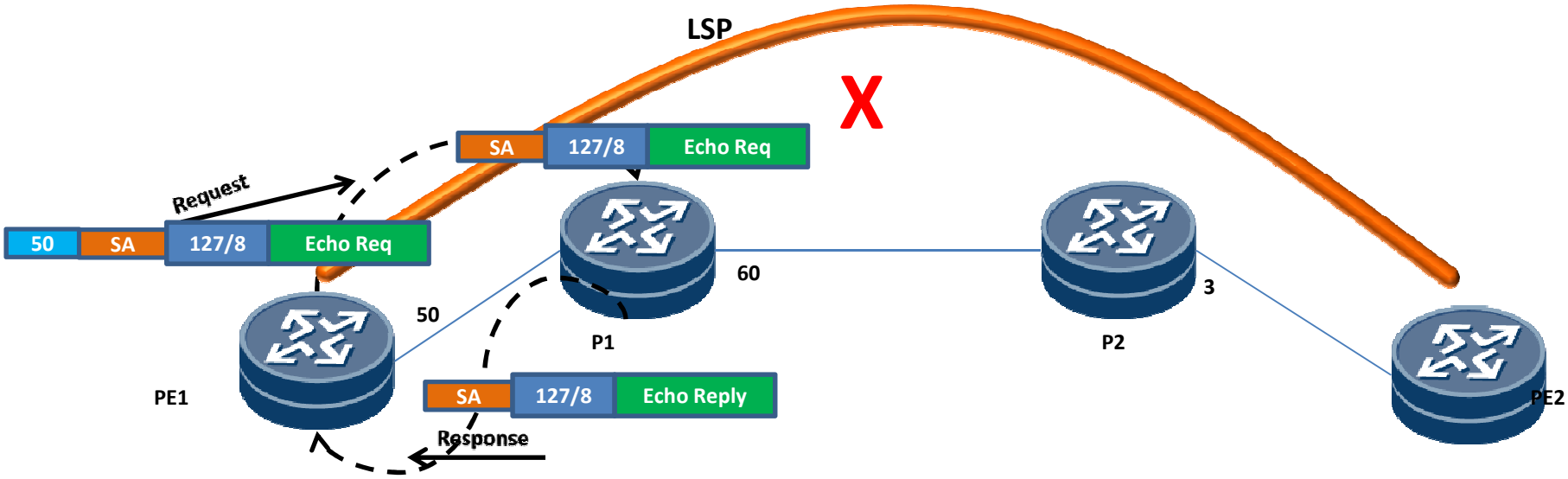
Note: No DSMAP TLV is sent by Egress router

Theory of Operation



- Packet is encodes with the same label stack as data packet
- The destination header of the packet is set as local host address
- The packet is forwarded on Egress interface identified for the FEC
- The packet get labeled switched on transit routers
- No special treatment of OAM packets on transit routers
- The Echo reply is sent as IP as default

LSP ping as diagnostic tool



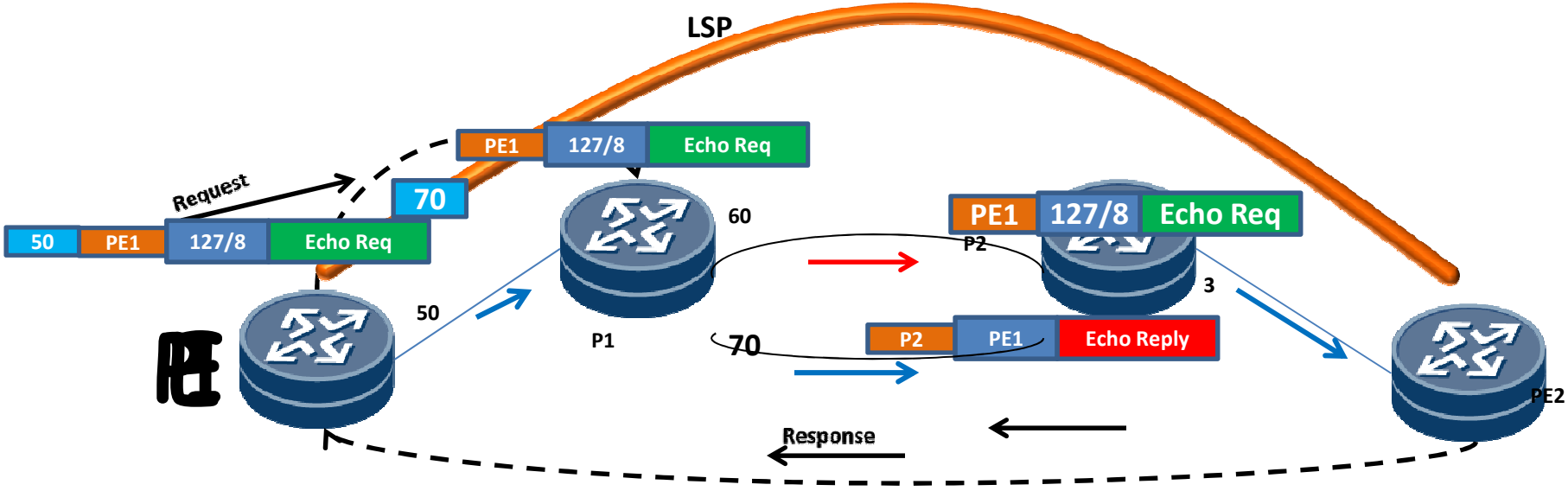
LSP could be broken due to various reasons

- No MPLS interface
- No LDP adjacency
- Label mismatch
- Control Plane and Data Plane mismatch

LSP ping Echo Request cannot get label forwarded due to LSP breakage

- Echo request gets locally processed due to local address
- Reply sent by the processing router with appropriate error code

LSP ping for Control Plane Data Plane Mismatch



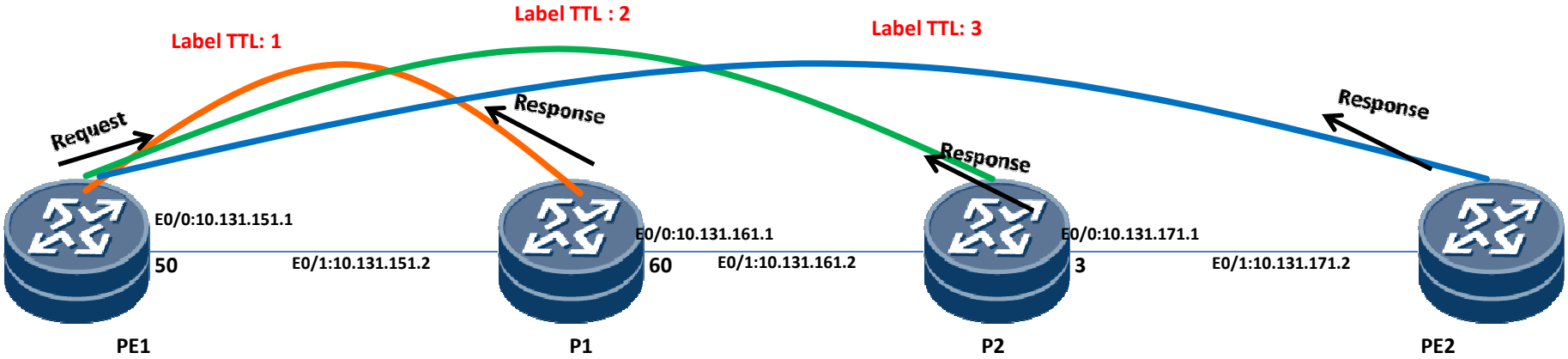
LSP control plane and data plane mismatch

- Control plane advertises label 60 to PE2 FEC
- Data Plane takes different path with label 70
- Though packets reach PE2, they traverse different path

LSP ping with DSMAP or Trace validation

- When LSP ping with DSMAP is set hop by hop, it can identify the fault
- DSMAP mismatch error will be return upon this error

Trace with LSP Ping



PE1 Downstream for PE2
MTU: MRU of E0/0
Address type: 1
DS Intf Addr: 10.131.151.1
DS Label : 50

P1 Downstream for PE2
MTU: MRU of E0/0
Address type: 1
DS Intf Addr: 10.131.161.1
DS Label : 60

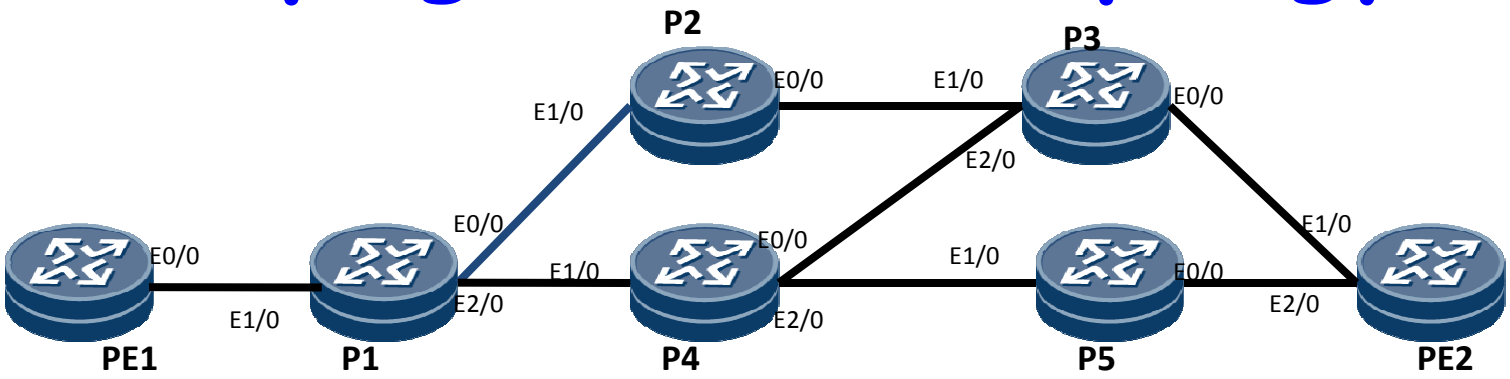
P2 Downstream for PE2
MTU: MRU of E0/0
Address type: 1
DS Intf Addr: 10.131.171.1
DS Label : 3

- LSP Ping with TTL is used to validate every hop of the LSP
- Downstream TLV is used to validate and request downstream info
- If the responding router is Egress of the FEC, a return code of 3 is returned.
- No DSMAP TLV is sent in the response by Egress router for the FEC

Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- **ECMP troubleshooting**
- BFD for MPLS
- Tools Galore

LSP ping in ECMP topology



PE1

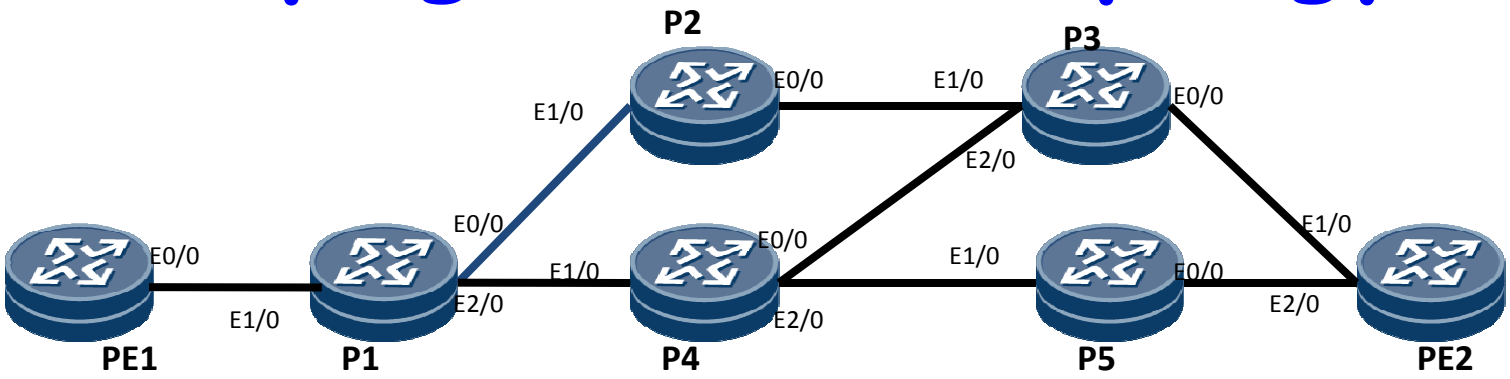
TTL = 1

DA: 127.0.0.0

MapSize/hash: 32/8

Bitmap:0xFFFF

LSP ping in ECMP topology



PE1

TTL = 1

DA: 127.0.0.0

MapSize/hash: 32/8

Bitmap:0xFFFF

1

P1

2

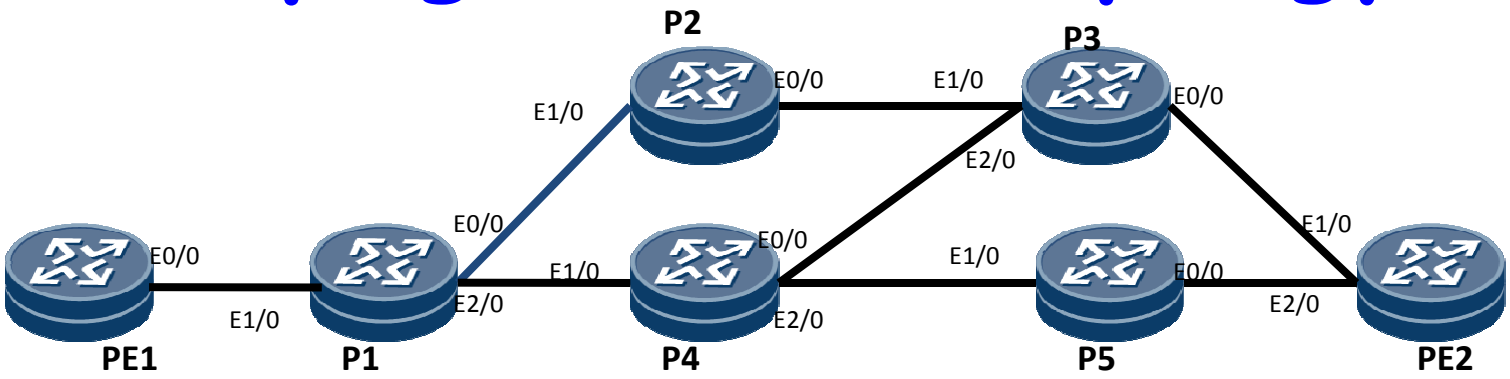
MultiPath1 [E0/0]

•Bitmap: 0x00FF

Multipath2[E2/0]

•Bitmap: 0xFF00

LSP ping in ECMP topology



PE1

3

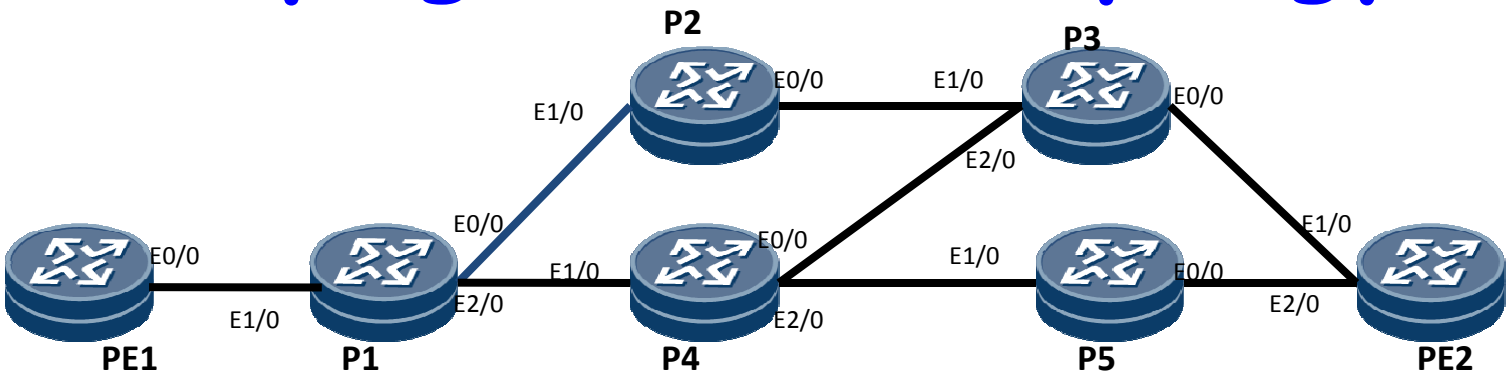
TTL = 2

DA: 127.0.0.24

Mapsize/Hash: 32/8

Bitmap:0x00FF

LSP ping in ECMP topology



PE1

TTL = 2

DA: 127.0.0.24

Mapsize/Hash: 32/8

Bitmap:0x00FF

3

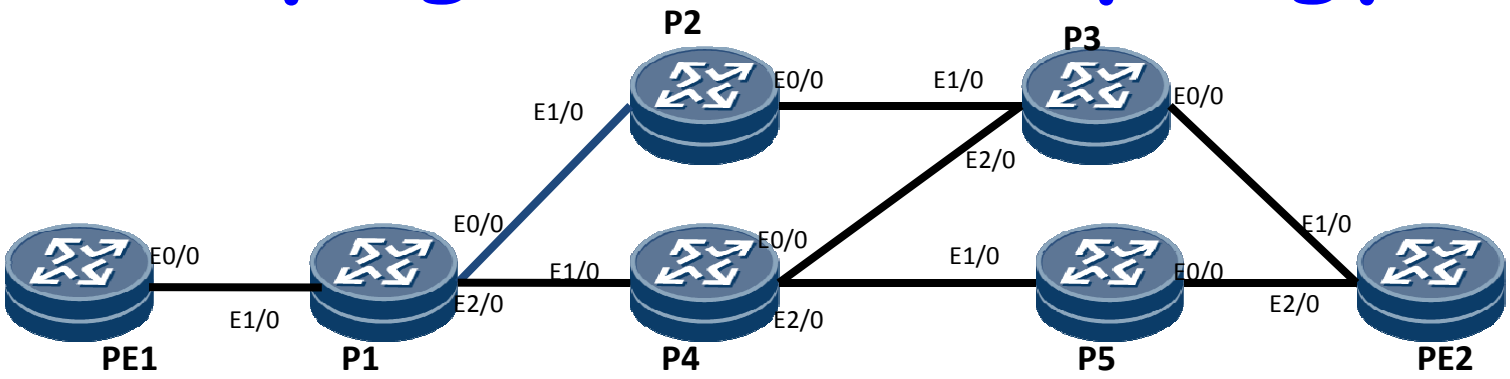
P2

MultiPath1 [E0/0]

•Bitmap: 0x00FF

4

LSP ping in ECMP topology



PE1

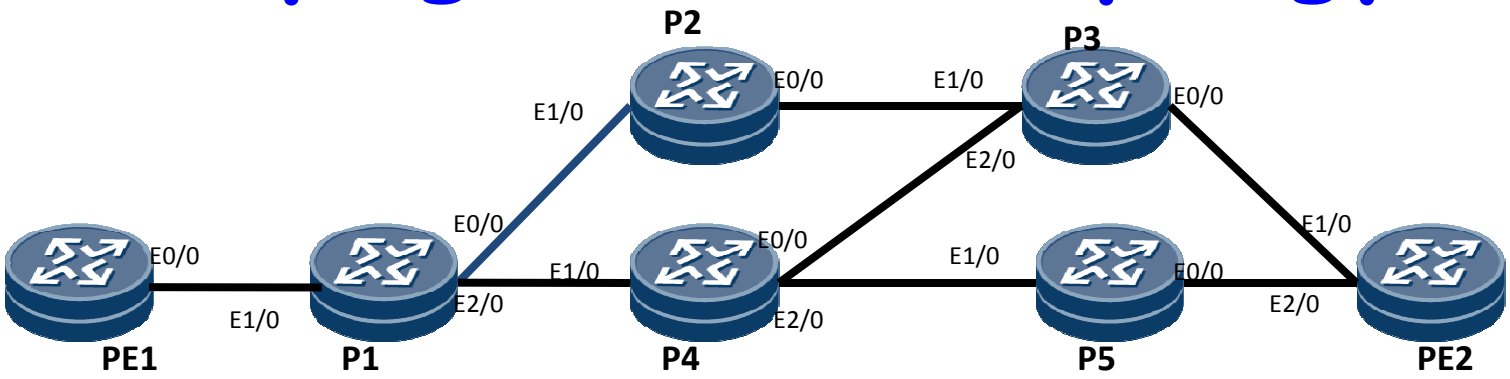
TTL = 3 5

DA: 127.0.0.24

Mapsize/hash: 32/8

Bitmap:0x00FF

LSP ping in ECMP topology



PE1

5

TTL = 3

DA: 127.0.0.24

Mapsize/hash: 32/8

Bitmap:0x00FF

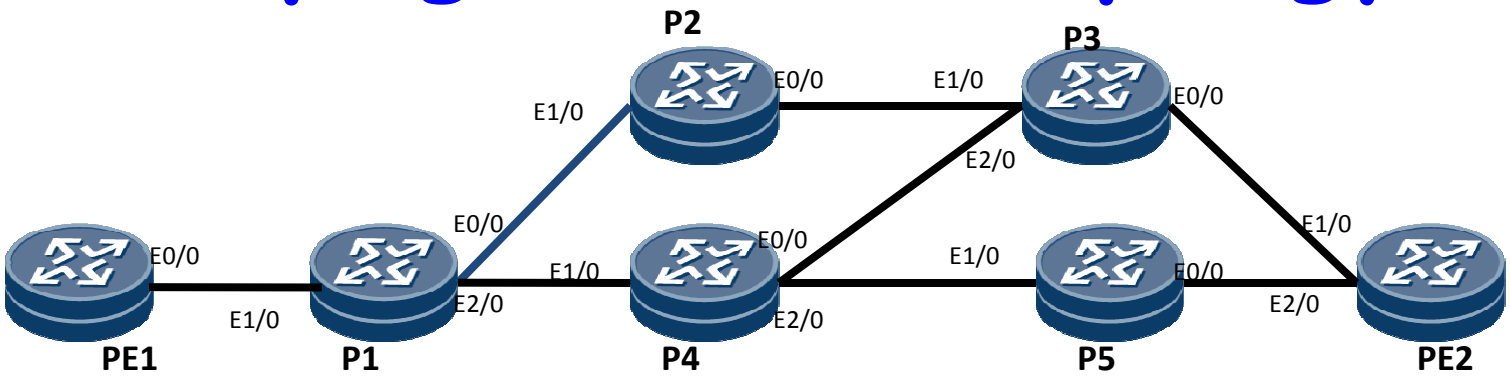
P3

MultiPath1 [E0/0]

•Bitmap: 0x00FF

6

LSP ping in ECMP topology



PE1 7

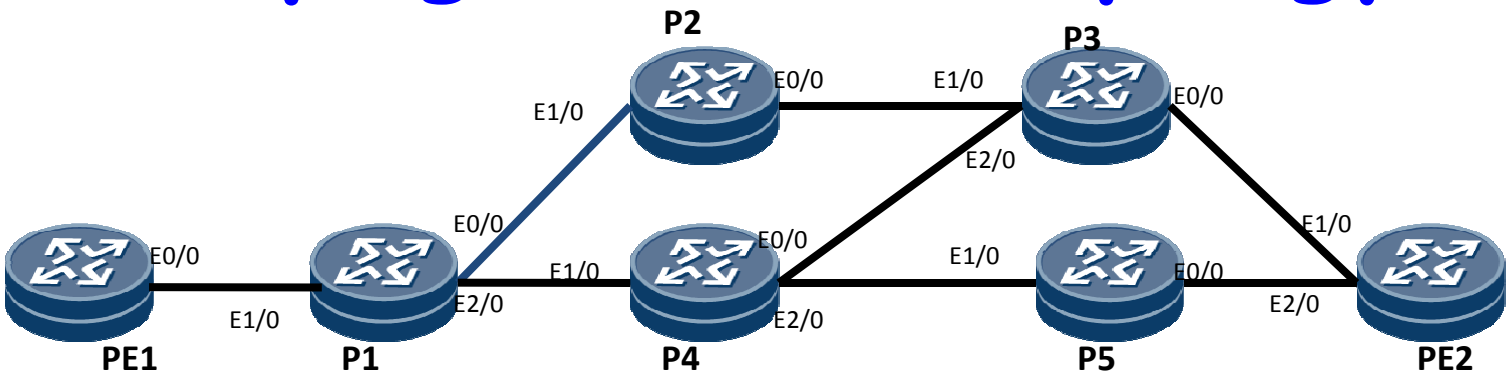
TTL = 4

DA: 127.0.0.24

Mapsize/Hash: 32/8

Bitmap:0x00FF

LSP ping in ECMP topology



PE1

7

TTL = 4

DA: 127.0.0.24

Mapsize/Hash: 32/8

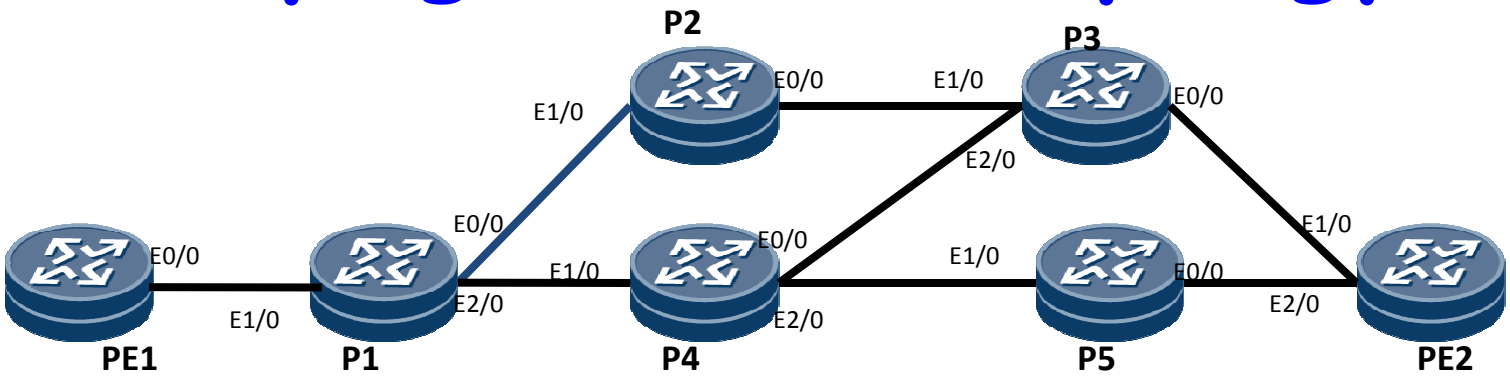
Bitmap:0x00FF

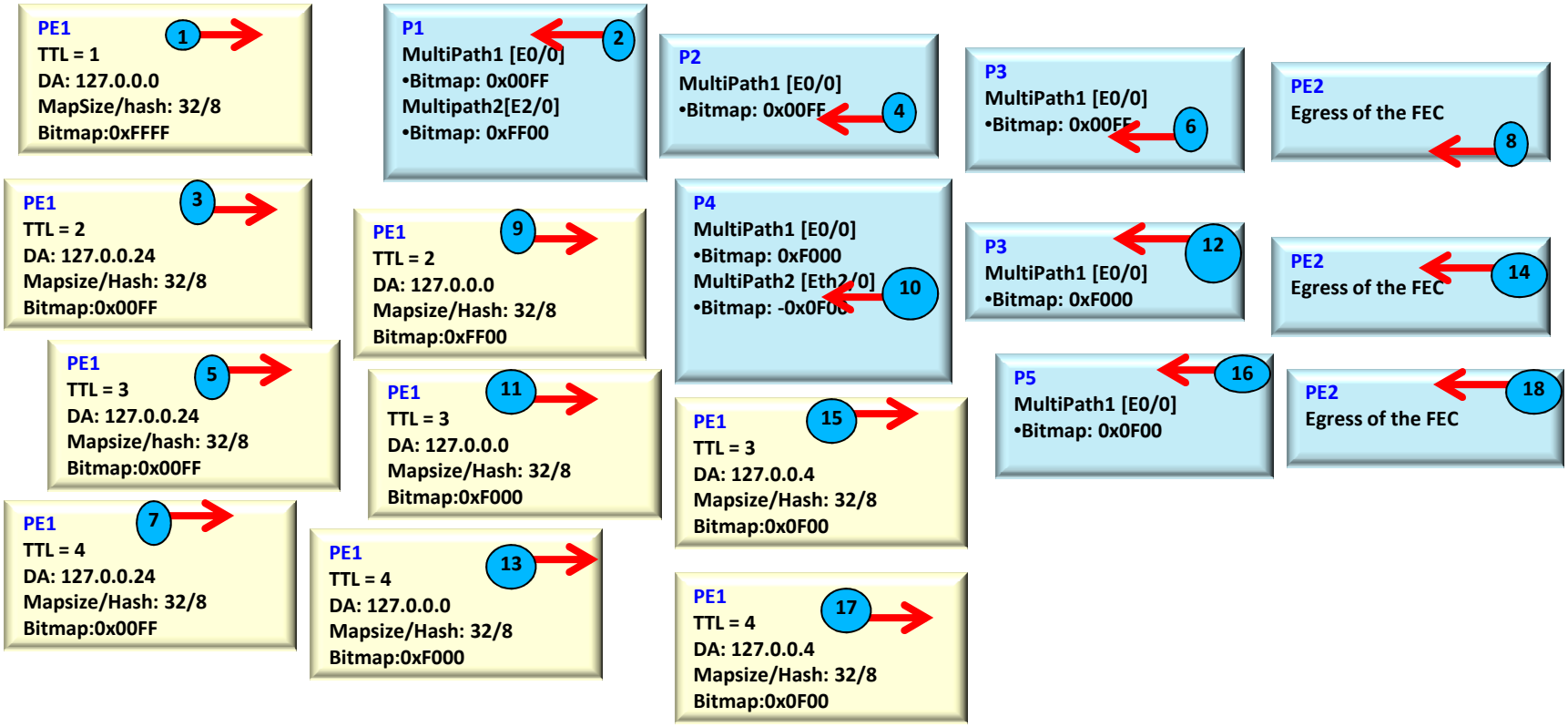
PE2

Egress of the FEC

8

LSP ping in ECMP topology





FEC types support

LSP ping supports various FEC types

FEC Type	LSP Ping	LSP Trace	ECMP Trace
LDP IPv4 and IPv6	Yes	Yes	Yes
RSVP TE v4 and v6	Yes	Yes	N/A
PW v4 and v6	Yes	MSPW(Yes)	Entropy Label
VPN v4 and v6	Yes	Yes	N/A
BGP v4 and v6	Yes	Yes	N/A
P2MP TE and mLDP	Yes	Yes	N/A
MPLS-TP	Yes	Yes	N/A

LSP ping for Pseudowire FEC

Requirement	**Provide end-to-end fault detection and diagnostic features for emulated Pseudowire service** **•P2P PWE3** **•MS-PW end-to-end Ping and Trace** **•Static and Dynamic Pseudowires**
Solution	**VCCV provides control channel to allow control packets over Pseudowires** **•VCCV capabilities are signalled using control protocols** **•Ability to support Control Word encapsulation** **•Router Alert labeled packets are to be punted** **•TTL exhaustion causes the packet to be processed**
Applications	**Layer 2 transport over MPLS** **•EoMPLS** **•FRoMPLS** **•ATMoMPLS**
Solution	**RFC5085**

Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- **BFD for MPLS**
- Tools Galore

Bidirectional Forward Detection (BFD)

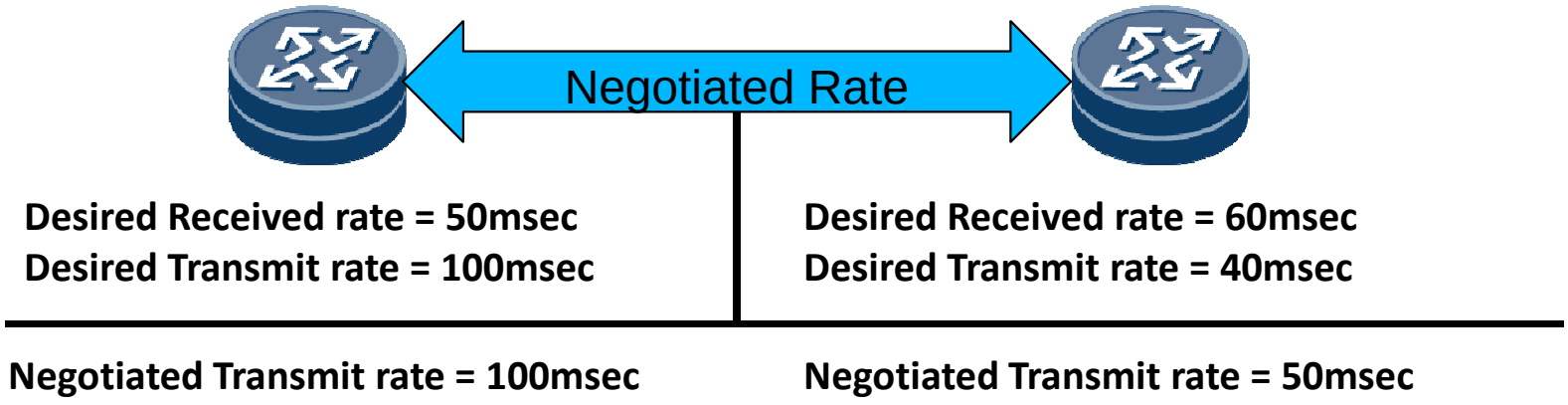
- Simple fixed-field, hello protocol
- Packets are periodically transmitted over respective directions of the path
- If a node stops receiving BFD packets, some component of the bidirectional path is assumed to have failed.
- Several modes of operation

BFD protocol Overview

- Typical hello protocol
- Neighbors continuously negotiate transmit and receive rates in micro seconds
- Dynamic rate adaption
- Neighbor is declared down when hello packets don't show up
- Uses UDP/IP or Non IP packets as BFD packets
- Ability to support single-hop and multi-hop

BFD Timer negotiation

- Neighbors continuously negotiate transmit and receive rates
- Designated UDP ports 3784 and 3785 are assigned to BFD
- Ability to support single-hop and multi-hop

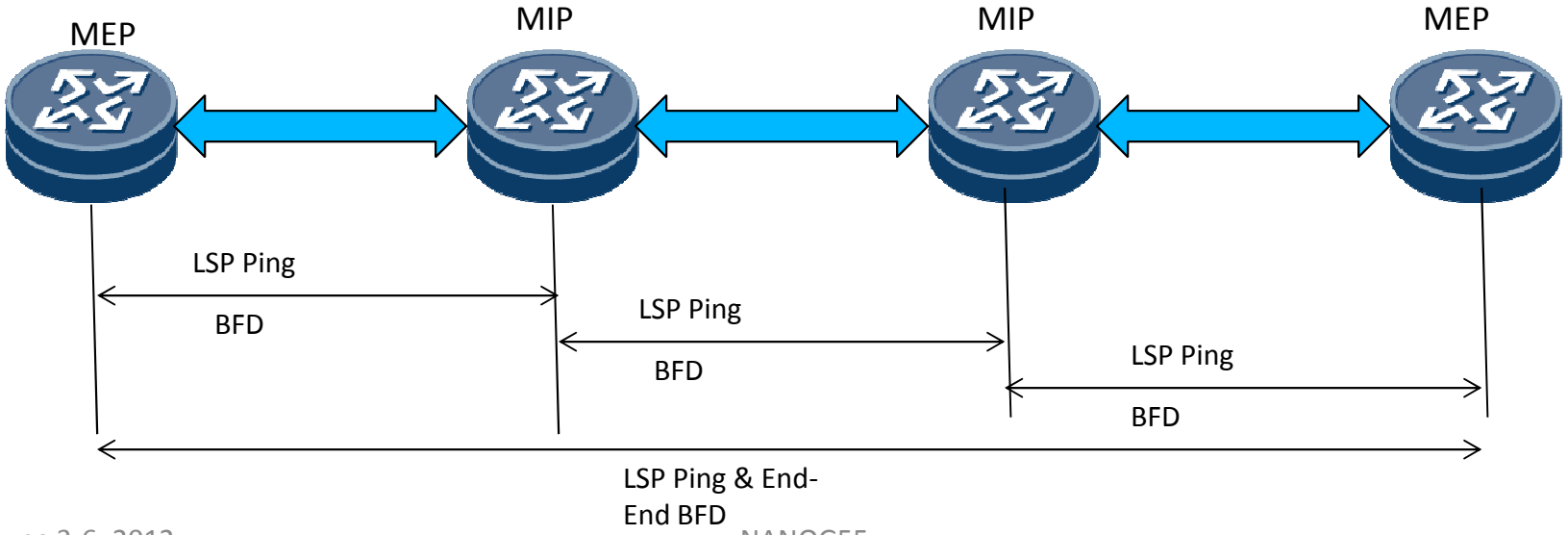


BFD for MPLS

- Ability to verify LSP
- BFD to verify TE tunnels, TP tunnels, PW LSP's etc
- VCCV to be used to verify PW LSP's
- BFD could be used to complement or replace use of RSVP hellos for MPLS FRR Link/Node protection
- BFD to carry AIS, RDI errors to end points of TP tunnels
- BFD the primary mechanism to make fast switchover and meet transport requirements
- BFD to play complimentary role to provide OAM within MPLS

LSP ping & BFD for MPLS-TP

- LSP ping got enhanced to support TP LSP's
- LSP ping plays crucial role in static TP LSP's.
- Ability to perform MEP-MEP, MIP-MEP and MIP-MIP OAM functions
- BFD is used to fast detect failures
- GAL label(13) to identify OAM and BFD packets



Agenda

- Introduction
- Terms and Terminology
- An Introduction to Tools
- Introduction to MPLS
- MPLS TP 101
- Troubleshooting MPLS
- MPLS OAM
- LSP Ping
- ECMP troubleshooting
- BFD for MPLS
- **Tools Galore**

Tools

- CC and CV for MPLS networks using LSP Ping
- Fault Isolation using traceroute with LSP Ping
- Performance monitoring based on Y.1731 model
- 1:1, 1+1, 1:n and m:n protection supported using BFD
- All FEC types supported using LSP ping
- Provides support for IPv4 and IPv6
- Automated tools built around LSP ping and other OAM tools
- No CCIE expertise required to use these tools

Summary of OAM tools

	Continuity Check	Connectivity Verification	Path Discovery	Defect Indications	Performance Monitoring
ICMP		Echo (Ping)	Traceroute		
BFD	BFD control	BFD Echo			
LSP Ping		Ping	Traceroute		
IPPM					-Delay - Packet loss
MPLS-TP OAM	CC	CV	Traceroute	-Alarm Reporting - Client failure Ind - Remote Defect	-Delay - Packet loss

Ref: draft-ietf-opsawg-oam-overview-05

Summary

Summary

- MPLS OAM covers all types of MPLS networks
- No CCIE's required to manage MPLS networks
- Already built into major vendors MPLS devices
- Deployed and being used in major carrier networks
- Inter-op tests carried out at various labs prove the OAM technologies WORK
- MPLS-TP brought forth the usefulness of OAM in transport networks
- "MPLS OAM" a proven technology

Questions