Evolving OAM Requirements with IP-Based Tools

Monique J. Morrow,

CTO Consulting Engineer

mmorrow@cisco.com

1

Motivation

- **Align MPLS OAM Requirements with IP-based tool constructs**

- **Comparison of various mechanisms**

Agenda

- **Why IP-Based Tools?**

- **Requirements**

- **LSP Ping**

- **BFD/VCCV**

- **OAM Message Mapping**

- **ITU-T/IETF Mechanism Overview**

- **Future Directions**

- **Summary**

Why IP-Based Tools?

Why IP-Based Tools?

- MPLS is IP-based

- All MPLS control protocols are based on the IP protocol suite

  LDP / BGP / RSVP / PIM

  - Facilitate evolutionary implementation of such mechanisms in deployed networks

ATM OAM Does NOT Equal MPLS OAM

Virtual Circuits	Label Switched Paths
Bi-directional	Usually Uni-directional
Established via ATM Signaling or Management	Establishment tied closely to control planes
Fixed hierarchy VP/VC	Variable Label Stack
Connection oriented	Can be "connectionless"
Single route	May use ECMP
No penultimate popping	Penultimate hop popping

Requirements

Requirements (1)

- **Three categories of requirements gathered from 1st tier PWE/MPLS Service Providers (and others).**
  - ✓**VC/LSP Path Verification and Tracing**
  - ✓**Built-in Protocol Operations**
  - ✓**Standard Management APIs/NMS Applications**
    MIBs, CLI, XML, etc…
  - ➢**Documented in:**
  - ➢**draft-ietf-mpls-oam-requirements-01.txt**
  - ➢**Must be addressed *before* many providers will deploy PWE3 services.**

Requirements (2)

- **Control plane verification of information
  Consistency check
  Authentication**

- **Data Plane Verification**

- **Ability to trace paths from PE to PE – Global routing table as well as VPNs**

- **Ability to trace paths from CE to CE within a VPN**

- **Ability to trace LSPs with ECMP**

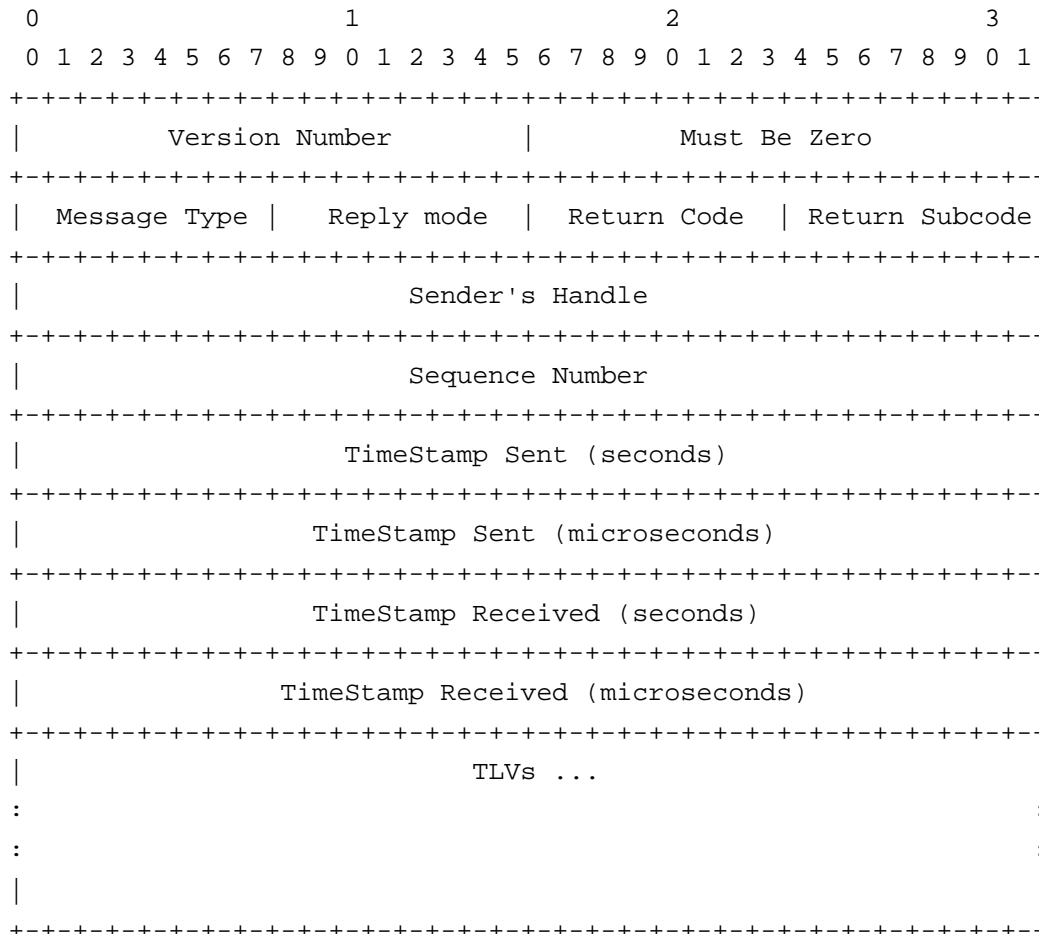- **Ability to Trace TE tunnels**

LSP Ping

LSP Ping

- **Similar to ICMP (IP) Ping**

    Sequence Number

    Timestamps

    Sender Identification

- **Full identification of FEC based the application**

- **Variable length for MTU discovery**

- **Support for tunnel/path tracing**

- **Multiple-reply modes**

- **Handles ECMP**

- **Reference**

    http://www.ietf.org/internet-drafts/draft-ietf-mpls-lsp-ping-03.txt

MPLS Ping: Operation

- **Ping Mode: Connectivity check of an LSP**

    Test if a particular "FEC" ends at the right egress LSR

- **Traceroute Mode: Hop by Hop fault localization**

- **Uses two messages**

        MPLS Echo Request

        MPLS Echo Reply

- **Packet need to follow data path**

MPLS Ping Message Format

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Version Number           |          Must Be Zero         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Message Type |   Reply mode   |  Return Code  | Return Subcode|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sender's Handle                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Sequence Number                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    TimeStamp Sent (seconds)                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 TimeStamp Sent (microseconds)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  TimeStamp Received (seconds)                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               TimeStamp Received (microseconds)               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          TLVs ...                             |
:                                                              :
:                                                              :
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

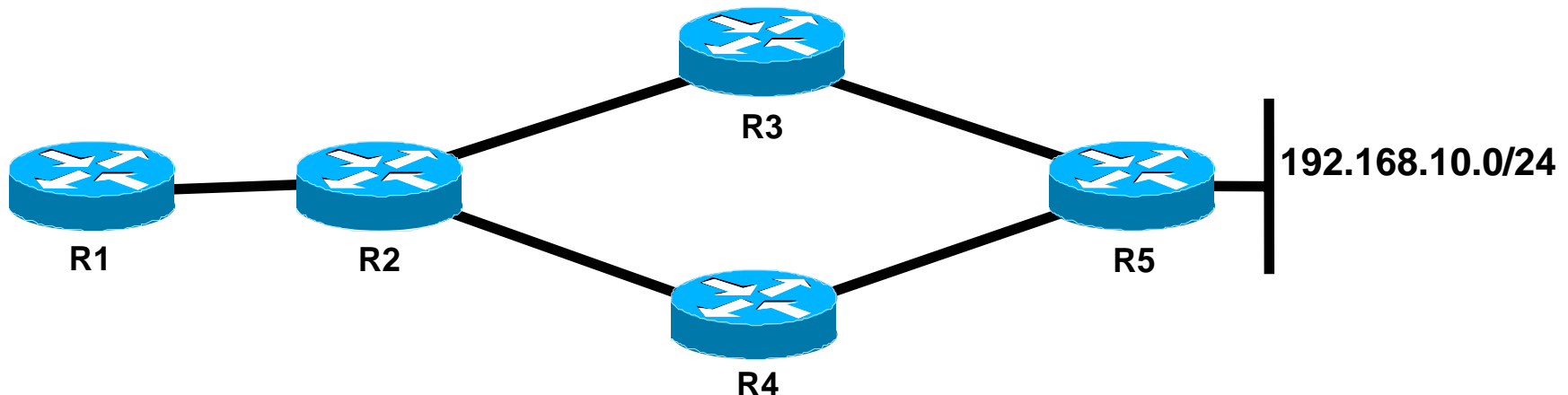**Message Type**
1 Echo Request
2 Echo Reply

Reply Mode
No reply
IPv4 UDP packet
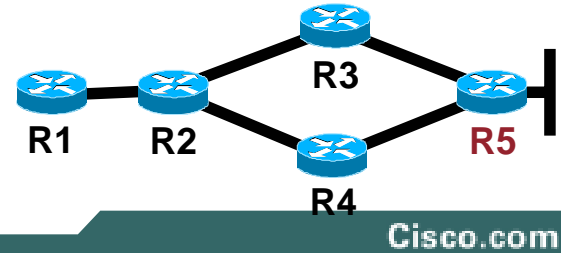IPv4 UDP packet with
  Router alert
Control Plane

TLVs include
FEC to be checked

MPLS Ping: Packet Flow

- **Ping with label for FEC=192.169.10.0/24**

- **Label Switched at R2, R3**

- **R3 pops label off**

- **R4 processes packet**

R3

192.168.10.0/24

R1   **R2**   **R5**

R4

Packet Flow Ping Mode: Egress node

- **Check Packet integrity**

- **Check if FEC distribution protocol is associated with incoming interface**

- **Check if valid egress node for the FEC**

- **Send echo Reply according to value of Reply Mode**

MPLS Traceroute: Packet Flow

- **MPLS Ping Packets are sent with TTL=1,2,3**

- **Label switched if TTL > 1**

- **Processed where TTL expires**

R3

R1   **R2**   **R5** **192.168.10.0/24**

R4
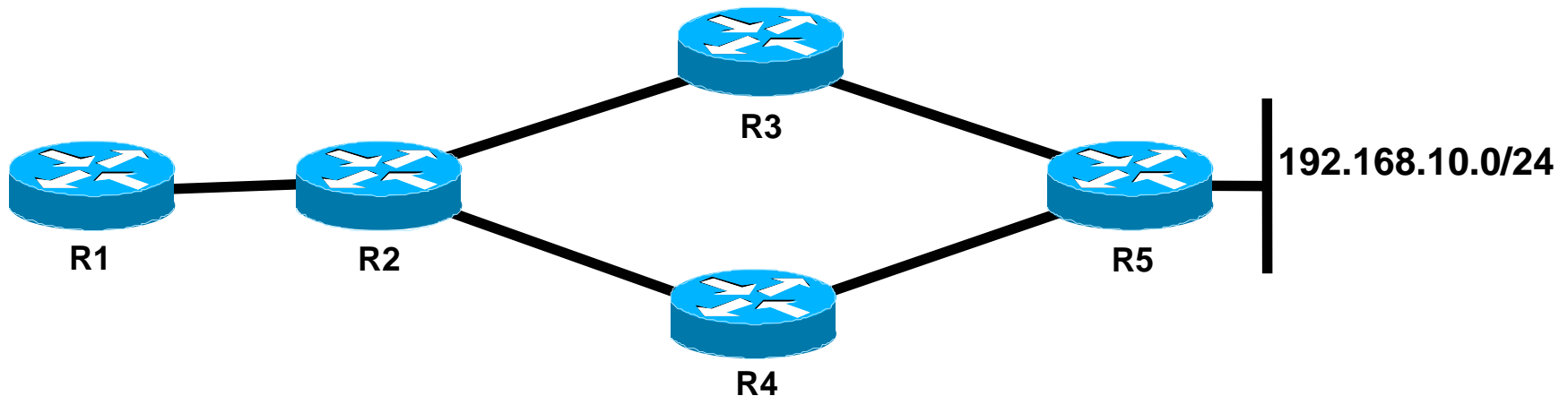
Trace Mode: TTL>1

- **Copy one Downstream Mapping (DM) TLV from Echo Reply**

- **Pick one IP Address from address in DM TLV**

- **Send a new Echo Request with TTL+1**

- **Repeat (if appropriated) for each DM TLV**

- **Reply from Egress stops iteration**

Packet Flow Trace Mode: R1
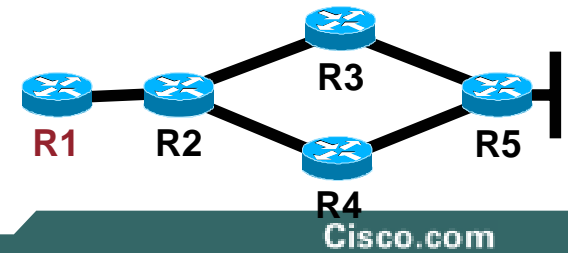
MPLS Header
TTL=1,2,3,4…..

IP
IP Src: IP-R1;
IP Dst: 127.x.y.z
TTL: 1
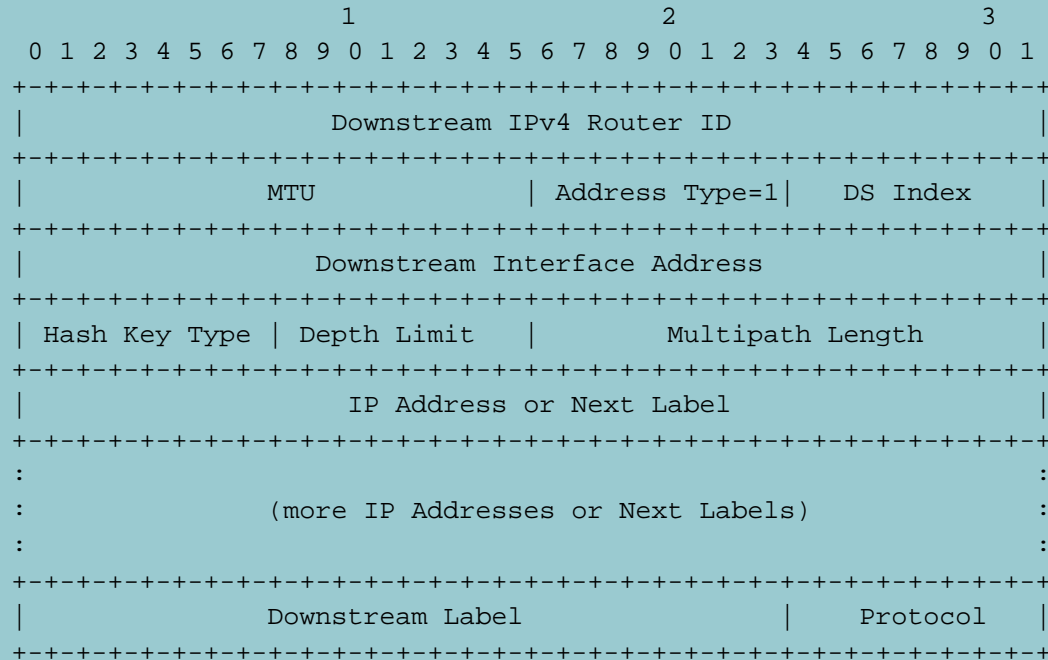Router Alert

UDP
Dest Port: 3503

Payload
Message Type=1
Reply Mode=2,
ReturnCode=0
Sender Handle
Sequence Number
TimeStamp Sent

TLVs
Target FEC stack
Downstream mapping

Downstream Mapping TLV

```
                          1                   2                   3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                  Downstream IPv4 Router ID                    |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |          MTU                  | Address Type=1|   DS Index    |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                  Downstream Interface Address                |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     | Hash Key Type | Depth Limit   |        Multipath Length       |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                  IP Address or Next Label                    |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     :                                                              :
     :             (more IP Addresses or Next Labels)               :
     :                                                              :
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                  Downstream Label             |   Protocol   |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Packet Flow Trace Mode: Transit Node
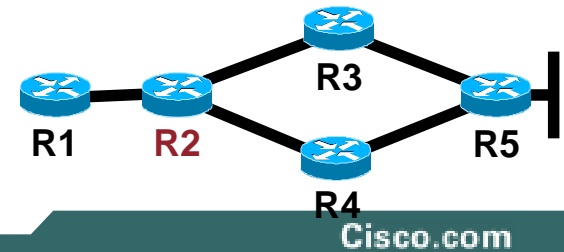
- **Reply processing same as Ping, then**

- **Check for Downstream Mapping TLV**

    Determine nexthop routers

- **Add Downstream Mapping TLVs for each**

    Compute label stacks, address/label ranges

- **Return received Label Stack if requested**

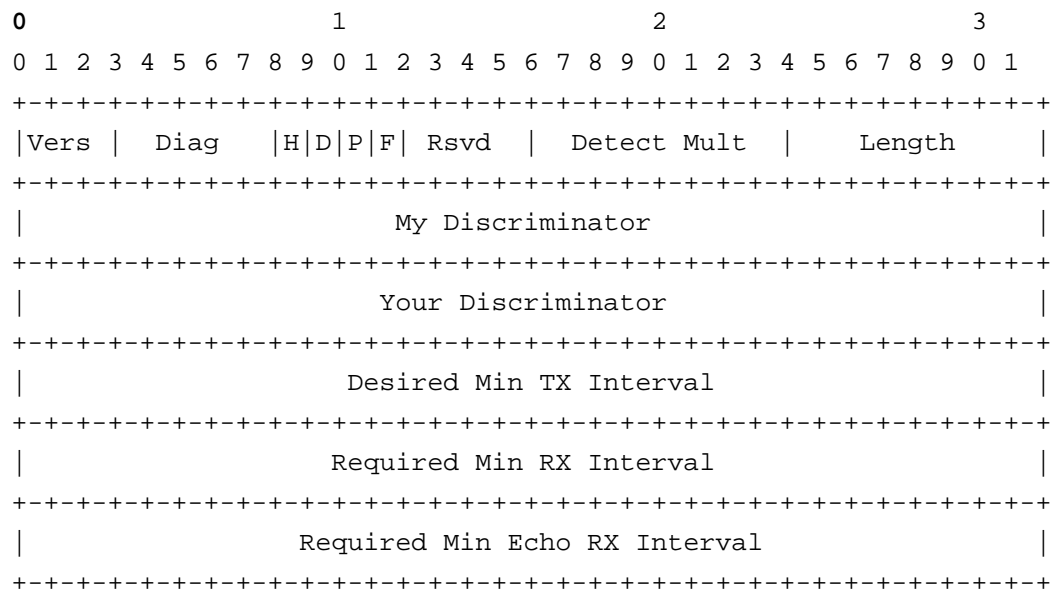# Bidirectional Forwarding Detection/ Virtual Circuit Connection Verification

Bidirectional Forwarding Detection

- Simple, fixed-field, hello protocol

- Nodes transmit BFD packets periodically over respective directions of a path

- If a node stops receiving BFD packets some component of the bidirectional path is assumed to have failed

- Several modes of operaton

  VCCV uses Asynchronous mode

  www.ietf.org/internet-drafts **draft-katz-ward-bfd-01.txt**

BFD Control Packet

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Vers | Diag  |H|D|P|F| Rsvd | Detect Mult   |    Length      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      My Discriminator                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Your Discriminator                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Desired Min TX Interval                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Required Min RX Interval                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Required Min Echo RX Interval                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Variable detection intervals

- **Each node estimates how quickly it can send and receive BFD packets**

- **Nodes exchange the follow parameters in every control packet**

    Desired Min TX Interval

    Required Min RX Interval

    Detect Multiplier

- **These estimates can be modified in real time in order to adapt to unusual situations**

Determining Detection Time

TX – Transmission Interval

RX – Receive Interval

> Note that TX(a->b) = RX(b->a)

TX(a->b) = max(Desired Min TX(a), Required Min RX(b))

TX(b->a) = max(Desired Min TX(a), Required Min RX(b))

Detection Time(b) = Detect Mult(a) x T(a->b)

TX is jittered by 25%

Diagnostics

0 -- No Diagnostic

1 -- Control Detection Time Expired (RDI)

2 -- Echo Function Failed (N/A to VCCV)

3 -- Neighbor Signaled Session Down (FDI)

4 -- Forwarding Plane Reset (Indicates local equipment failure)

5 -- Path Down (Alarm Suppression)

6 -- Concatenated Path Down (used to propagate access link alarms)

7 -- Administratively Down

Virtual Circuit Connection Verification (VCCV)
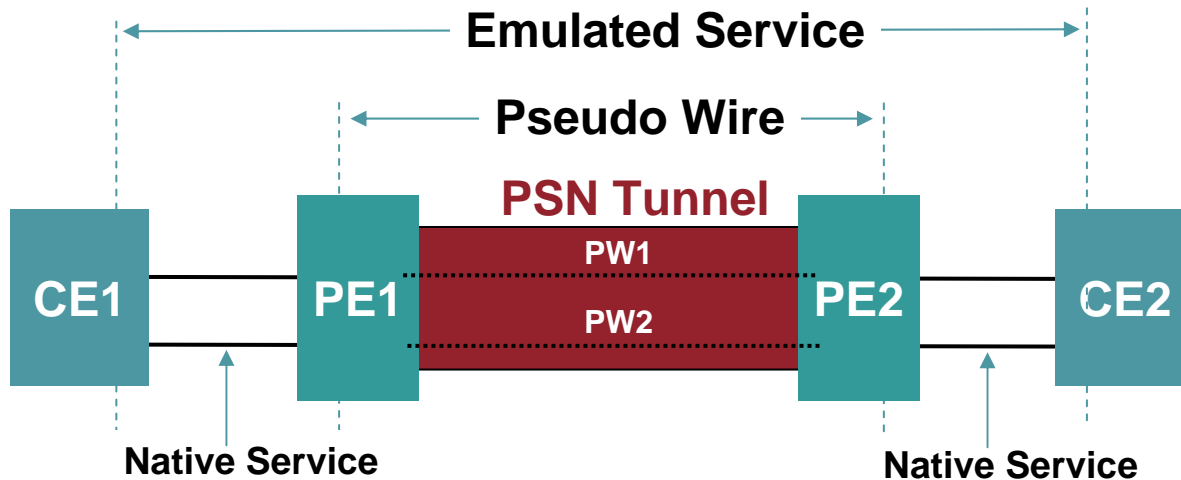
- **Multiple PSN Tunnel Types**
  - **MPLS, IPSEC, L2TP, GRE,…**
- **Motivation**
  - •**One tunnel can serve many pseudo-wires.**
  - •**MPLS LSP ping is sufficient to monitor the PSN tunnel (PE-PE connectivity), but not VCs inside of tunnel.**

- **www.ietf.org/internet-drafts/draft-ietf-pwe3-vccv-00.txt**

VCCV Overview

- **Mechanism for connectivity verification of PW**

- **Features**

  Works over MPLS or IP networks

  In-band CV via control word flag or out-of-band option by inserting router alert label between tunnel and PW labels
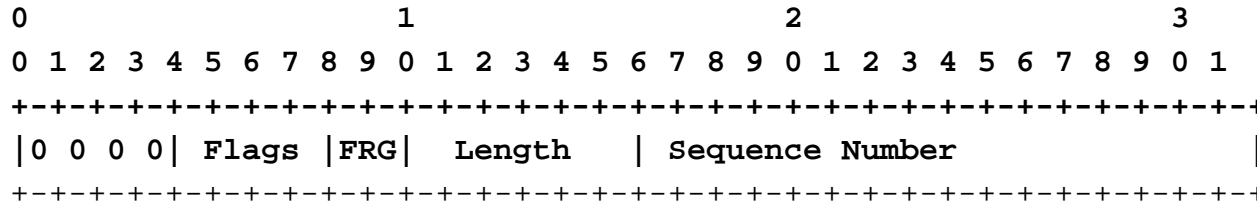
  Works with BFD, ICMP Ping and/or LSP ping

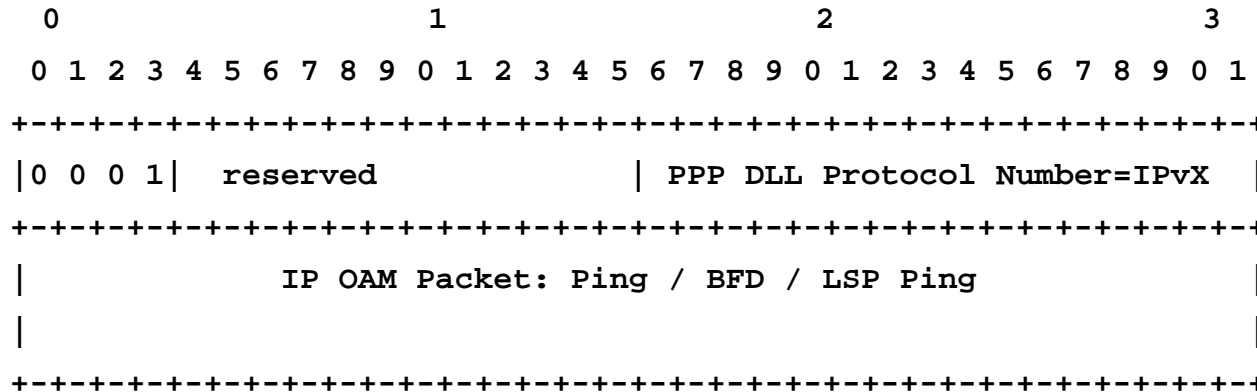- **VCCV results may drive OAM/LMI injection on corresponding AC(s)**

- **http://www.ietf.org/internet-drafts/draft-ietf-pwe3-vccv-00.txt**

In Band VCCV Format

Control word use is signalled in LDP - Standard form:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 0| Flags |FRG|  Length   | Sequence Number               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

OAM uses a different 1st nibble

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 1|  reserved             | PPP DLL Protocol Number=IPvX  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            IP OAM Packet: Ping / BFD / LSP Ping               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

PWE3 OAM Example: Continuity Verification

Attachment VCs

BFD Packet over VCCV channel

LSP Tunnel
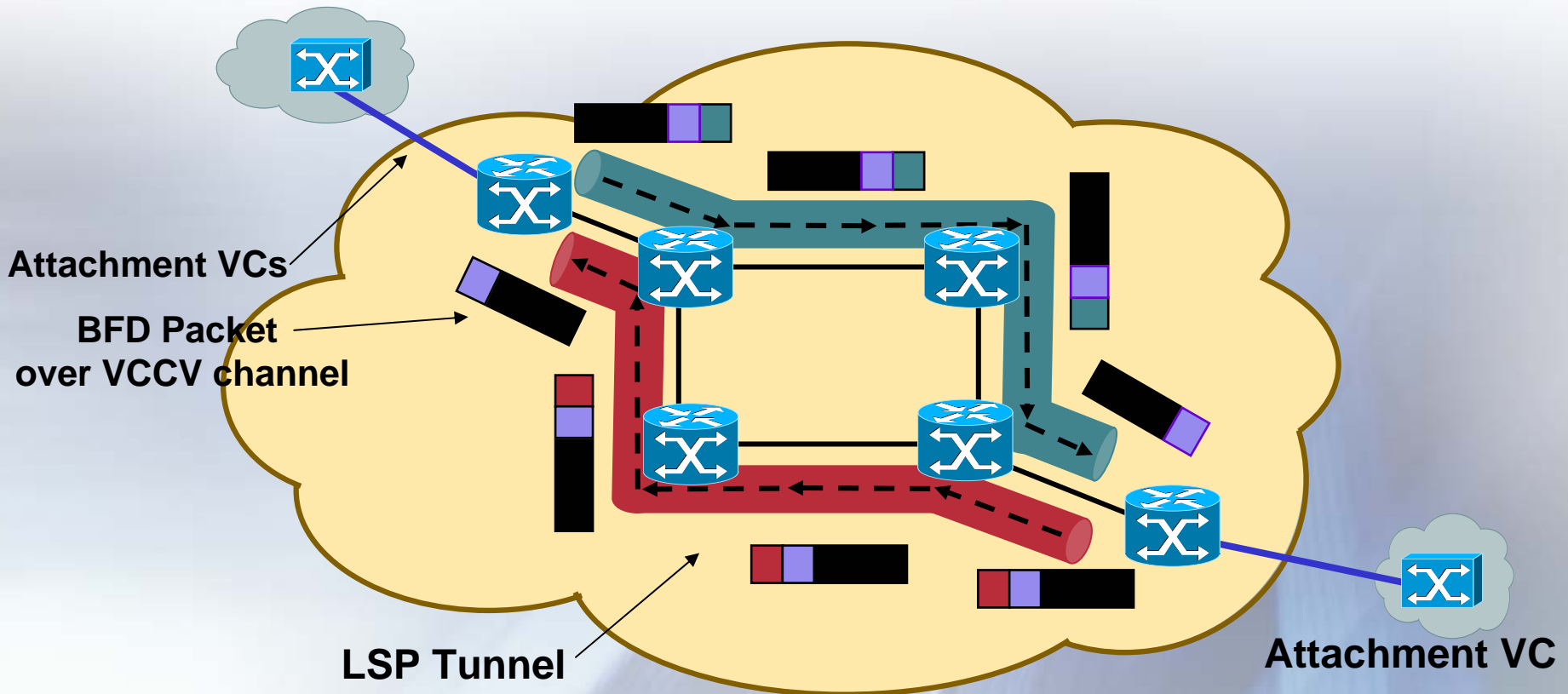
Attachment VC

- **BFD provides a lightweight means of regular periodic CV**

PWE3 OAM Example: Connection Verification

Attachment VC

VCCV Packet

Two Labels
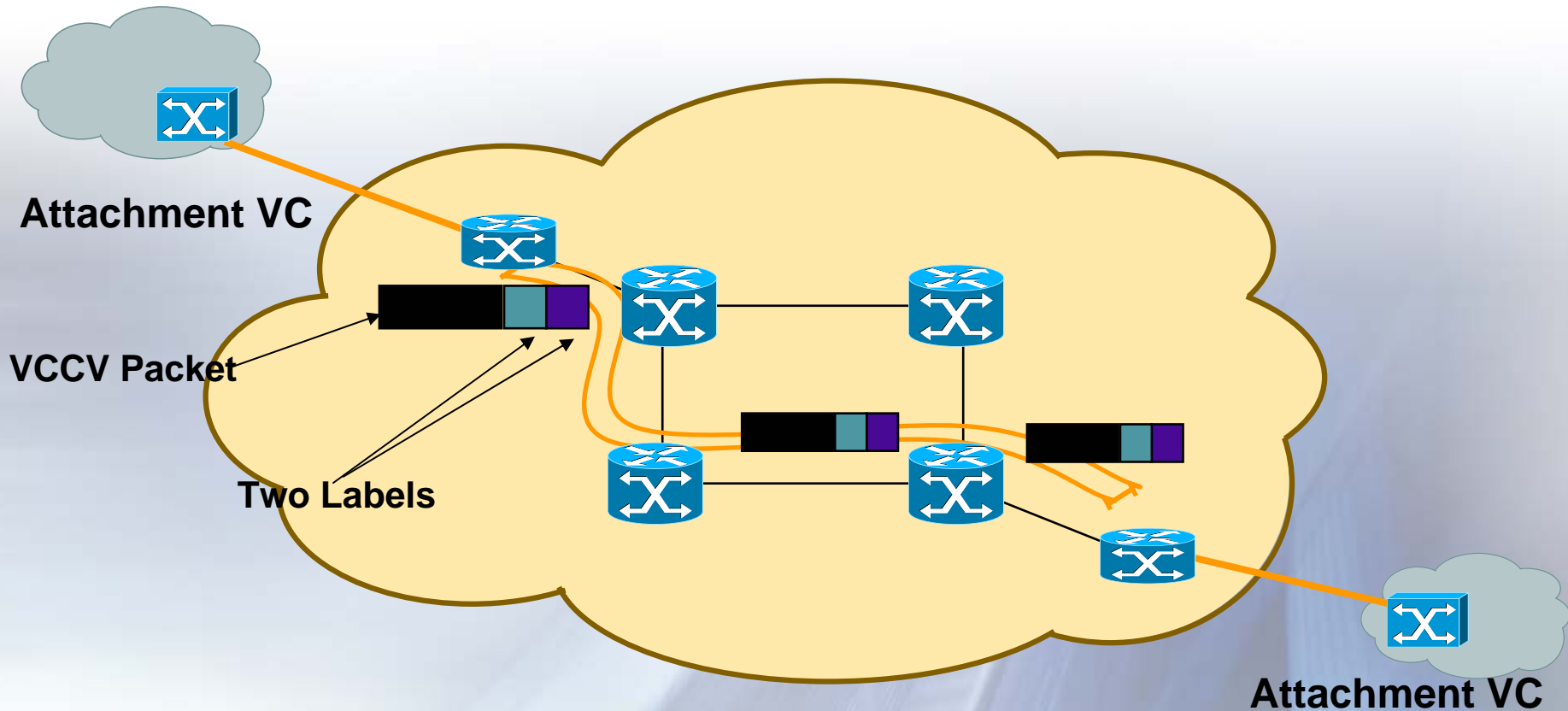
Attachment VC

- •Verify/Trace Path of LSP Tunnels between PEs.
- •Verify/Trace Emulated services (e.g. ATM, FR) mapped to Attachment VCs
- •Trace/Verify packets must take same path as data packets.

Example of Operation
CV/Trace Using VCCV and LSP Ping

NMS/mgr Triggers LSP ping
trace when failure detected

NMS/mgr Triggers VCCV

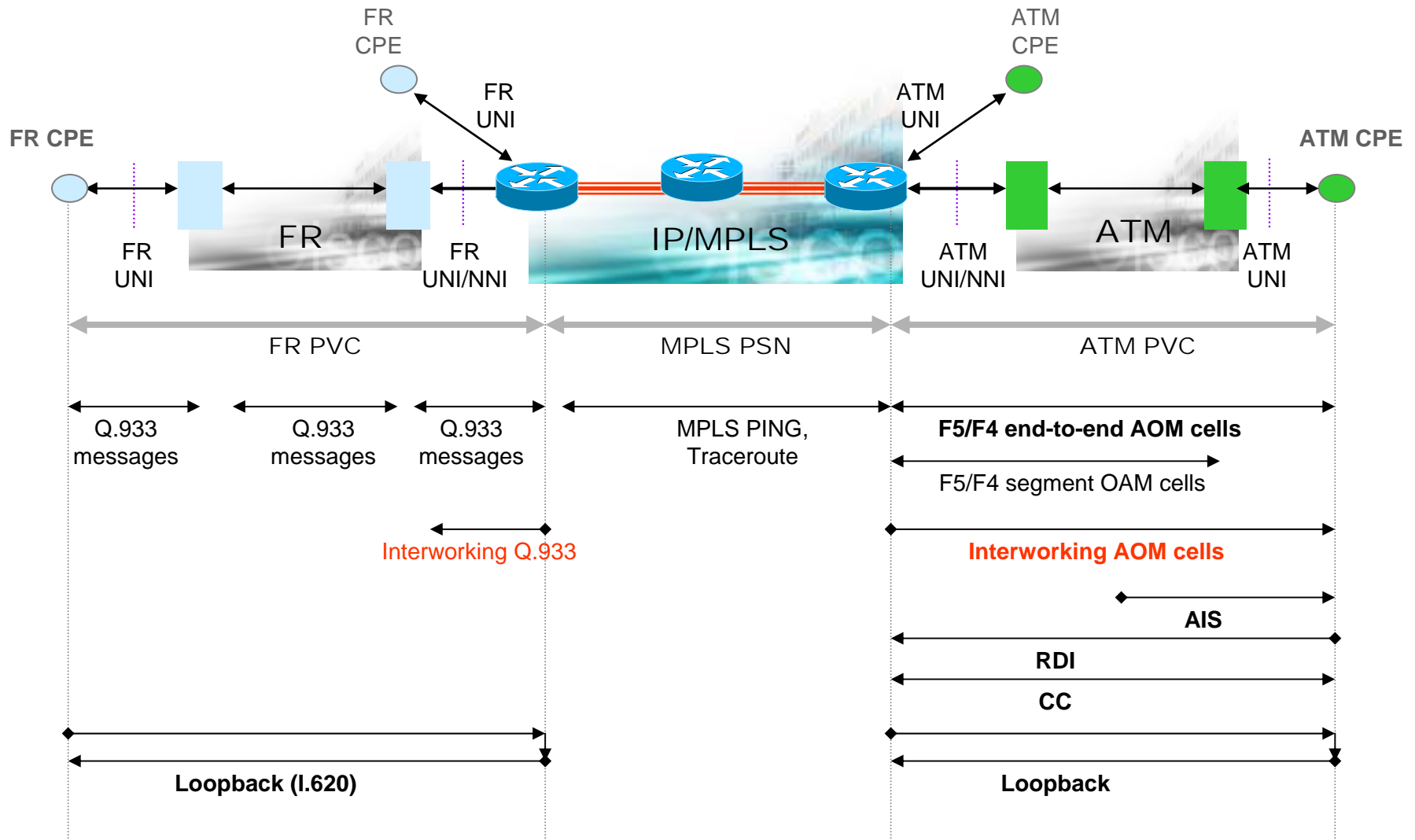**VCCV Packet
Is lost**

Attachment VC
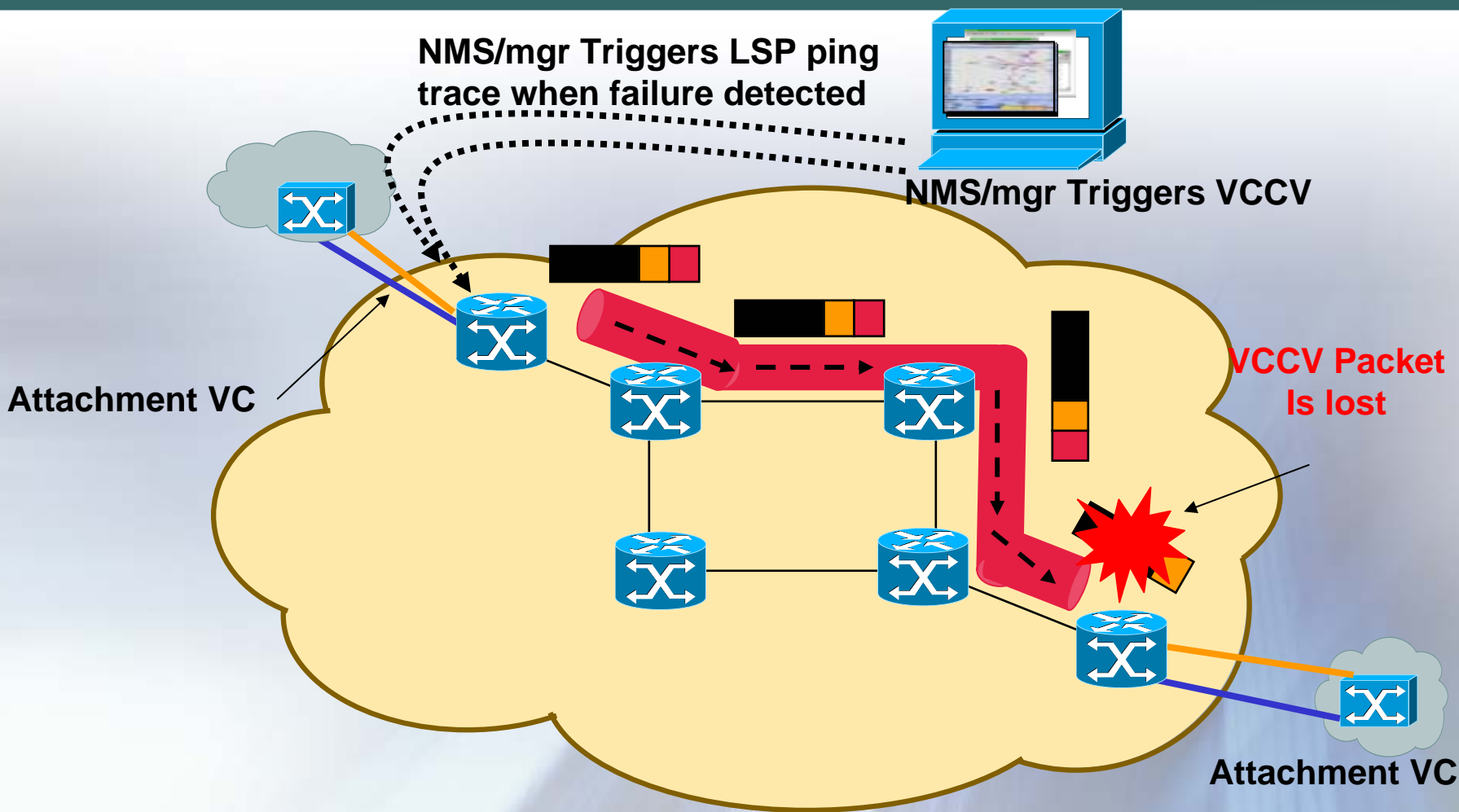
Attachment VC

OAM Message Mapping

OAM Message Mapping

- **Provides details of how LSP ping/VCCV failures should translate into native ATM/FR OAM messages that PEs return to the native attachment interfaces.**

- **OAM Emulation**

  AIS and RDI Generation for ATM AAL5 over MPLS

  AIS/RDI Generation for ATM upon reception of label withdrawal and vice versa

- **LMI/ILMI based status notification upon label withdrawal reception for pseudowire virtual circuits**

- **draft-nadeau-pwe-msg-mapping-01.txt**

OAM Message Mapping Example

FR CPE

ATM CPE

FR UNI

ATM UNI

FR CPE

ATM CPE

FR

IP/MPLS

ATM

FR UNI

FR UNI/NNI

ATM UNI/NNI

ATM UNI

FR PVC

MPLS PSN

ATM PVC

Q.933 messages

Q.933 messages

Q.933 messages

MPLS PING, Traceroute

F5/F4 end-to-end AOM cells

F5/F4 segment OAM cells

Interworking Q.933

Interworking AOM cells

AIS

RDI

CC

Loopback (I.620)

Loopback

ITU-T/IETF
Mechanism Overview

Addressing OAM Requirements
Two Basic Approaches

	ITU	IETF	
Requirements	**Y.1710**	**draft-ietf-mpls-oam-requirements-01.txt**	

Detection
     MPLS
       L3
       L2
	Y.1711 (a.k.a. CV)	**BFD, LSR Self Test**	

Diagnostic
     MPLS
 	
       L3
       L2
	FEC-CV (Y.17feccv), (Y.17fw)
	
	 FEC-CV (Y.17feccv),
	 FEC-CV (Y.17feccv), Y.17fw
		MPLS Ping/Trace
		
		 MPLS Ping/Trace
		 VCCV, OAM state mapping	

Instrumentation	**(Y.17fw)**	**MIB, Syslog**	
Recovery	**Y.1720, (Y.17fw)**	**FRR, MPLS HA, Graceful restart**	
Performance	**Y.MPLSperf**		
Security			

Drawbacks of Y.1711

- **Does not work for Penultimate hop popping, which is very commonly used.**

- **Assumes a fixed path/connection-oriented view of the world as in G.709 which is more the exception than the norm.**

- **Requirement for a reverse path**

   This is the exception rather than the rule for most MPLS applications

- **Fixed interval between CV Packets of 1 second.**

   Will not scale for typical networks comprised of large numbers of LDP signaled LSPs.

Y.1711 in a Nutshell

- **Y.1711 will not scale for auto-routed LDP networks which comprise about %90 of deployed MPLS networks!**

- **Revolutionary approach requiring re-spin of many pieces of hardware.**

Future Directions for ITU OAM

- *IP-based Tools approach added to y.17fw.*

- *Y.17fw to be consented, Feb 2004*

  Will leverage existing IP-based tools to overcome many aforementioned shortcomings.

  Will integrate new tools (LSP ping/trace, VCCV) into tool box of existing operator management tools.

  Evolutionary not revolutionary approach!

Future Directions

Future Directions

- **Bidirectional Forwarding Detection (BFD)**

  "Bidirectional Forwarding Detection",   IETF draft-katz-ward-bfd-01.txt

  "BFD for IPv4 and IPv6 (Single Hop)",   IETF draft-katz-ward-bfd-v4v6-1hop-00.txt

- **Working on Ethernet OAM mechanisms (IEEE)**

- **Interworking OAM is pivotal for converged services**

Summary

Summary

- **Most current applications offer an IP related service**

- **IP is the basis of MPLS control planes**

- **Most important, customer requirements highlight need for IP-based tool mechanisms**

- **Therefore, IP-based tools makes sense and provides natural evolution of MPLS OAM to current and future services!**

References

References (2)

- ➤ **draft-ietf-mpls-oam-requirements-01.txt**

- ➤ **draft-ietf-pwe3-vccv-00.txt**

- ➤ **draft-nadeau-pwe-msg-mapping-01.txt**

- ➤ **draft-swallow-mpls-lsr-self-test-00.txt**

- ➤ **draft-katz-ward-bfd-01.txt**

- ➤ **draft-katz-ward-bfd-v4v6-1hop-00.txt**

- **Guest Editor Special Edition IEEE Communications Magazine on topic of OAM for MPLS-Based Networks**

- **Scheduled Publication, October 2004**

- **Call For Papers:**
  http://www.comsoc.org/pubs/commag/cfpcommag1004.htm