BFD and MPLS: Marriage Made in Heaven

Zafar Ali and Mallik Tatipamula
Cisco Systems, Inc.

MPLS 2004

CISCO SYSTEMS

Outline

- History of BFD
- Brief overview of BFD Protocol
- Applicability of BFD in MPLS
- BFD updates from IETF
- Engineering Considerations
- Summary

Failure Detection Mechanisms

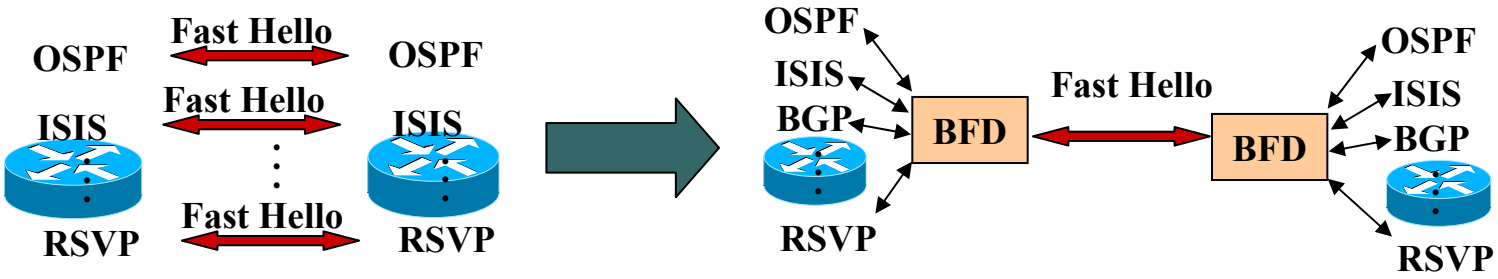
- Most routing and signalling protocols have some form of liveliness detection built in, e.g.,
  - OSPF, ISIS, RSVP-TE, LDP.
- Some protocols use the underlying transport to determine the liveliness of their neighbors, e.g.,
  - BGP uses TCP keepalives.
- Some protocols have intrinsic liveliness mechanisms, e.g.,
  - RIP

Fast Failure Detection Requirements

- For faster convergence of the network, requirement is to provide fast failure detection for these protocols
  - In a few tenths or even hundredths of a second.
- Issues
  - Time to detect for these mechanisms that one's neighbor is down ranges from seconds to tens or even hundreds of seconds.
  - Not all link layer technologies that inherently supply link outage detection.

Obvious Solution:
Allow sub-second protocol hello timers

Issues with Fast Protocol Hellos



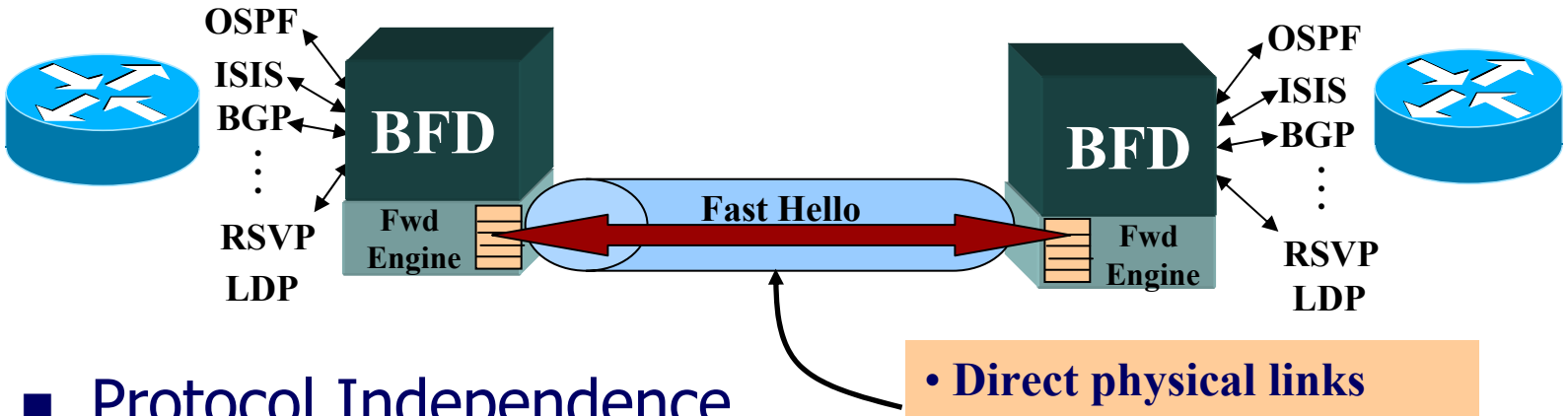
- Only feasible for protocols that are equipped with the Hellos
  - Does not cover all protocol space.
- Hello messages often carry more than just liveness information
  - Protocol Specific Parsing requirement
  - Computation Burden
  - High chances of False Positives
- "n" Fast Hellos for "n" protocols.

Several solution were proposed to avoid a proliferation of different failure detection methods, BFD is adapted by the IETF

Outline

- History of BFD
- Brief overview of BFD Protocol
- Applicability of BFD in MPLS
- BFD updates from IETF
- Engineering Considerations
- Summary

Attributes of BFD



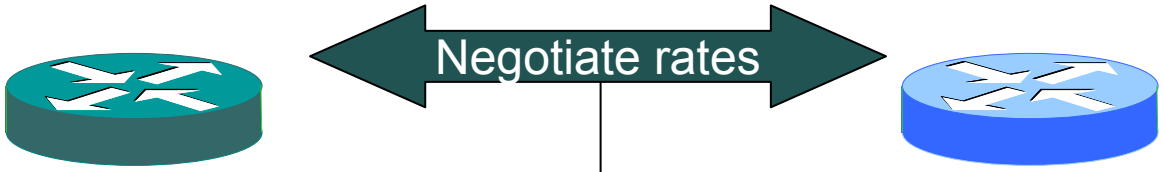
- Protocol Independence
- Media Independence
- Fast failure detection
  - Light Weight, Fixed Length; simple to parse
- Forwarding plane liveliness
  - E.g., Link may be up but forwarding engine may be down or an entry may be incorrectly programmed.
- **No discovery mechanism in BFD**
  - Applications bootstrap a BFD session

BFD Protocol overview

- It's a typical Hello Protocol.
- Neighbors exchange hello packets at negotiated regular intervals. A neighbor is declared down when expected hello packets don't show up.
- BFD control packets are encapsulated in UDP datagram.
- Destination port 3784 and source port between 49252 to 65535.
- Echo packets uses source and dest UDP port 3785.
- Various Modes of Operations.

Timer negotiation

- Neighbors continuously negotiate their desired transmit and receive rates in terms of microseconds.
- The system reporting the slower rate determines the transmission rate.
- Dynamic Rate Adaptation



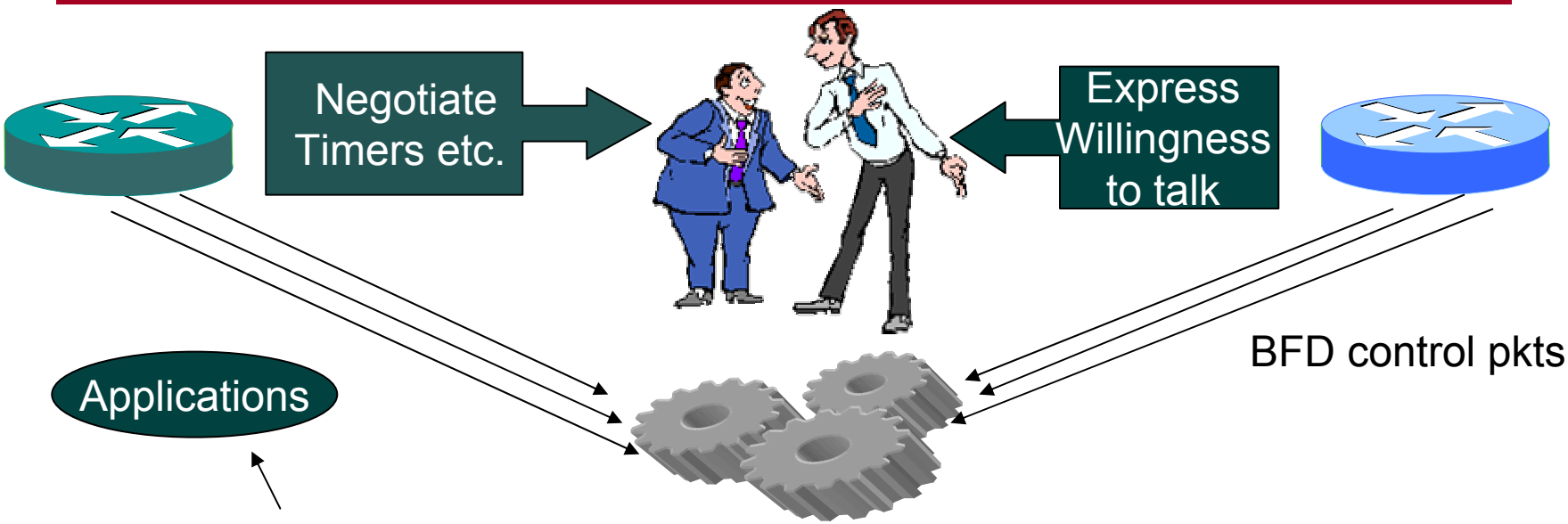
Desired Receive rate = 50 ms
Desired Transmit rate = 100 ms

Desired Receive rate = 60 ms
Desired Transmit rate = 40 ms

Green Transmits at 100ms

Blue transmits at 50ms

BFD Protocol Mechanics



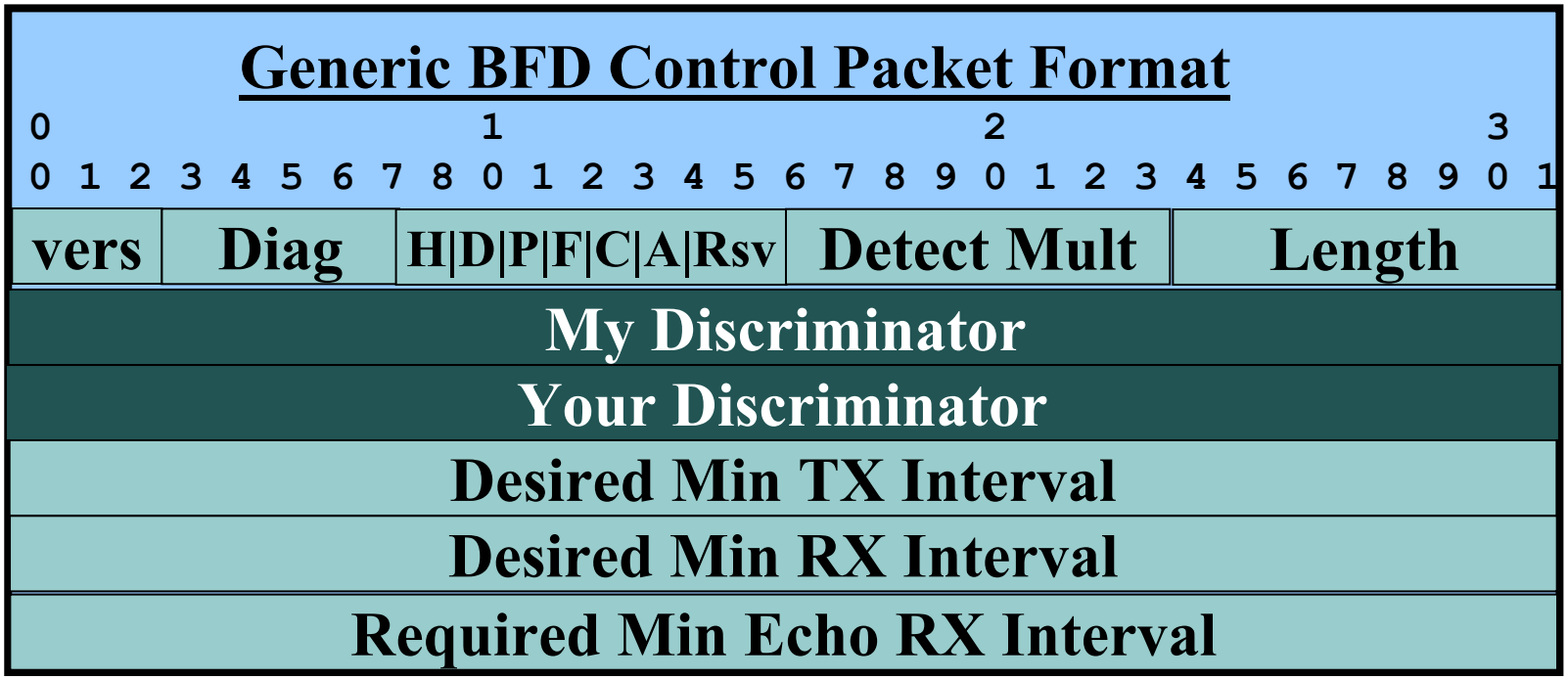
BFD Session is up and running

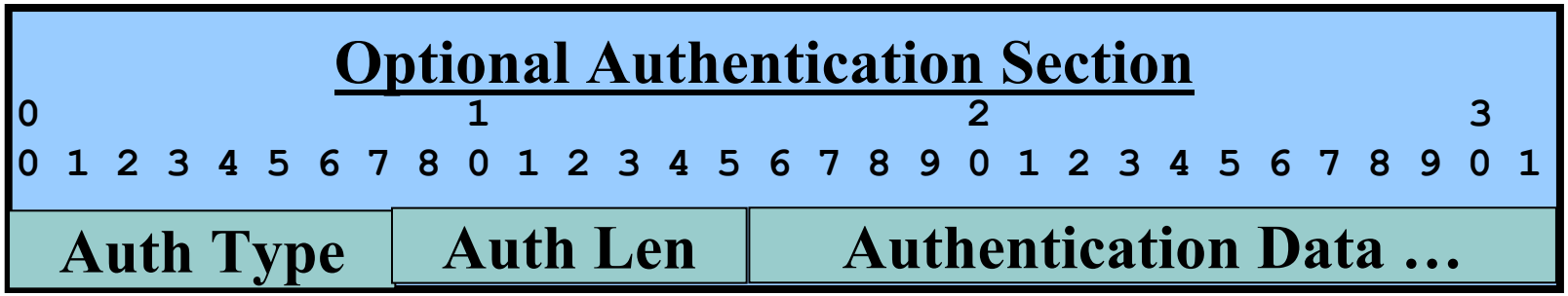
Periodic link
Health check

BFD

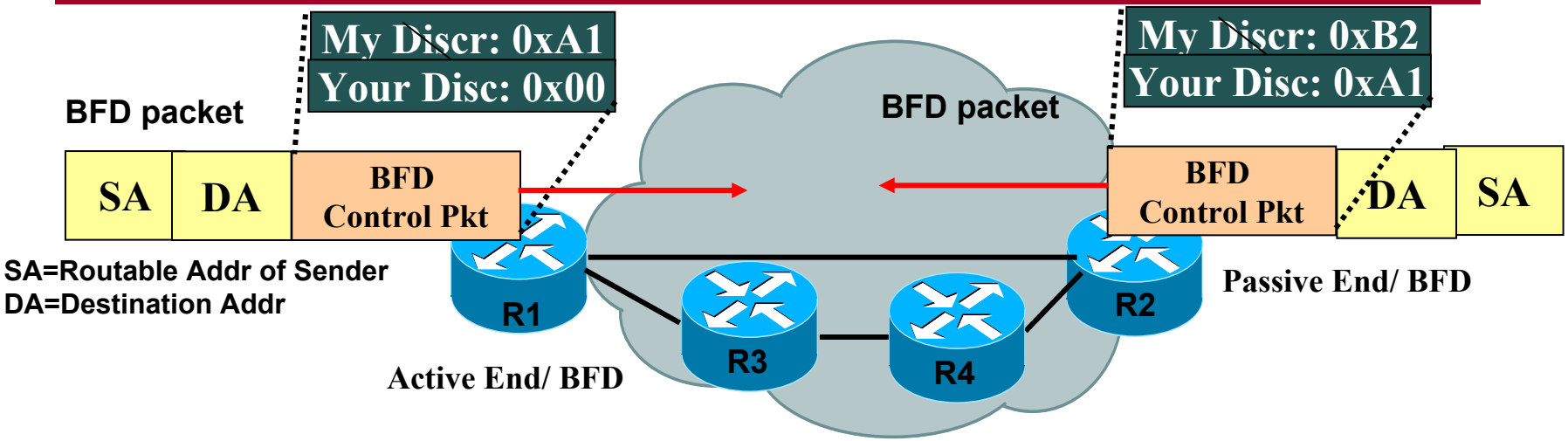
Forwarding
Goes Down

BFD Control Packet Format





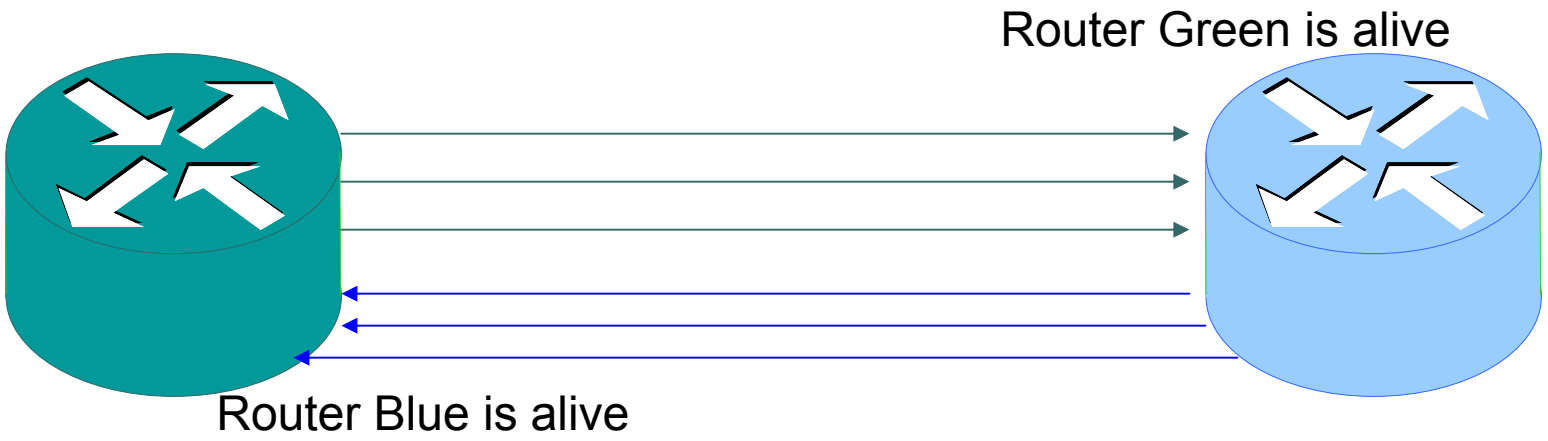
Bootstrapping of BFD Session



- Discriminator is a unique value in the system, which identifies a session
- The local discriminator is sent in 'My Discriminator field', and is echoed back in the 'Your Discriminator' field
- Once the remote end echoes back the local discriminator, all further received packets are demultiplexed based on the Your Discriminator field only
- Applicable to Single and Multi-hop
  - Single BFD session between a source/destination address pair or out-of-band bootstrapping of BFD Session.

Asynchronous Mode of Operation

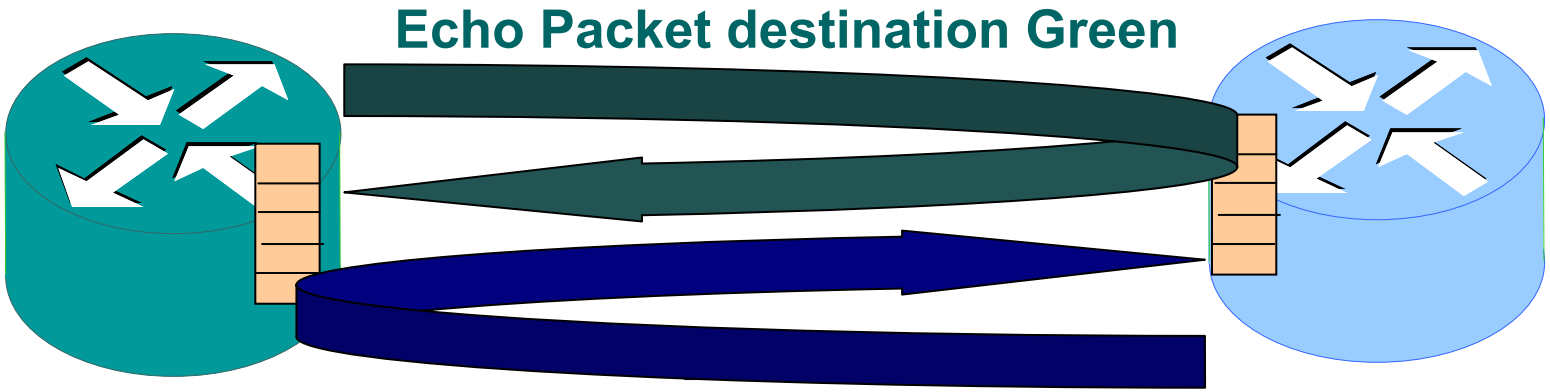
- Control packets flow in each direction



Systems periodically send BFD Control packets to one another, and if a number of those packets in a row are not received by the other system, the session is declared to be down.

Echo Mode of Operation

Routers send packets with destination IP address as their own address.



Echo Packets loop through the remote system.

The session is established using control session and then echo mode can be Negotiated between the systems.

On-Demand Mode of Operation

- Both systems must agree on running demand mode.
- Relies on alternate mechanisms to imply continuing connectivity
- Demand mode MAY be enabled or disabled at any time by setting or clearing the Demand (D) bit in the BFD Control packet, without affecting the BFD session state.
- Periodically, at a negotiated rate, the system trying to verify connectivity sends a packet with 'P' bit set and the neighbor replies with 'F' bit set.
- If no response is received to a Poll, the Poll is repeated until the detection time expires, at which point the session is declared to be down
- Demand mode is especially useful for a system with a very large number of BFD sessions.

Outline

- History of BFD
- Brief overview of BFD Protocol
- Applicability of BFD in MPLS
- BFD updates from IETF
- Engineering Considerations
- Summary

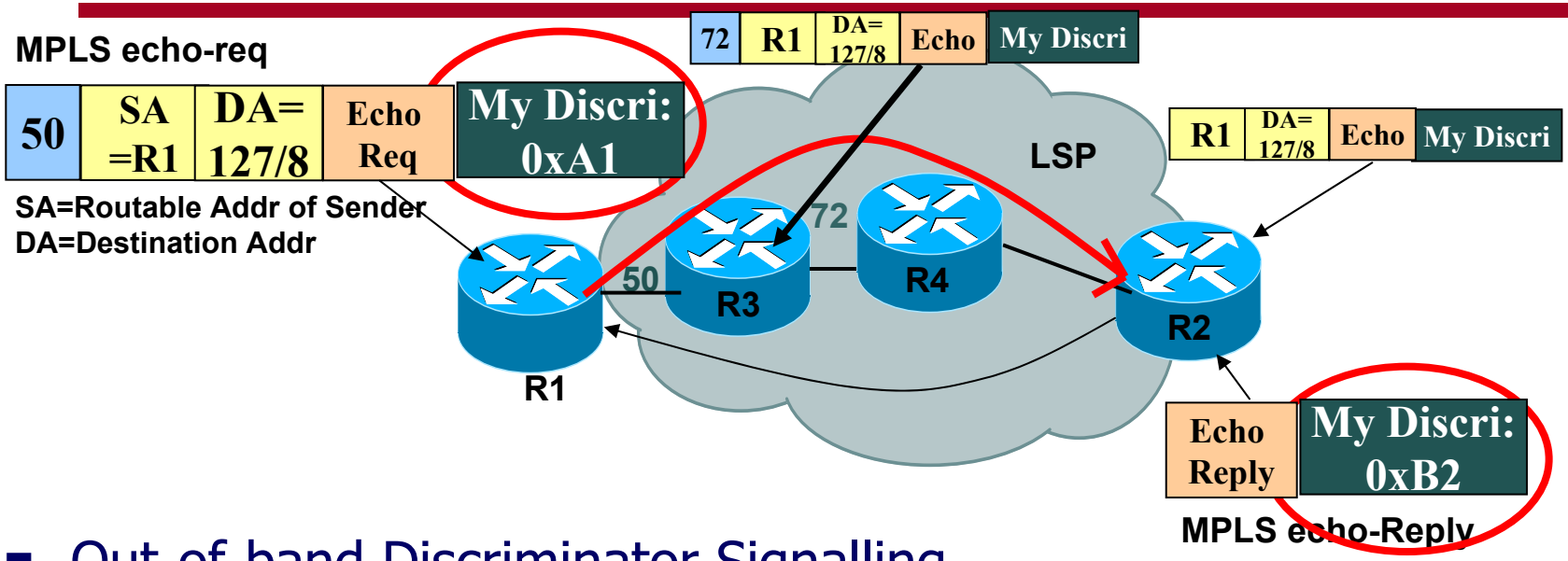
MPLS BFD Vs. LSP Ping

Method	Data Plane Failure Detection	Control Plane Consistency	Protocol Overhead
LSP Ping	**YES**	**YES**	**Higher than BFD**
MPLS-BFD	**YES**	**NO**	**Low**

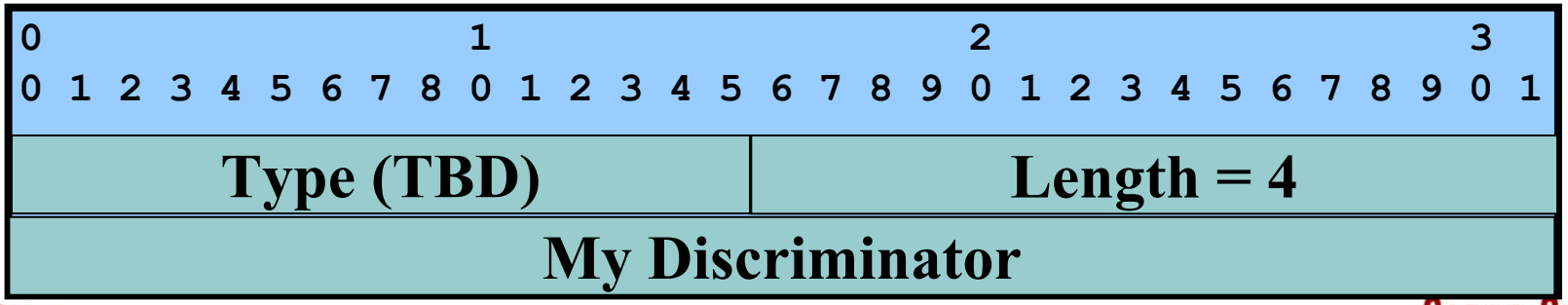
MPLS-BFD can complement LSP Ping to detect a data plane failure in the forwarding path of a MPLS LSP

Supported FECs:
RSVP IPv4/IPv6 Session, LDP IPv4/IPv6 prefix
VPN IPv4/IPv6 prefix, Layer 2 VPN, Layer 2 Circuit ID

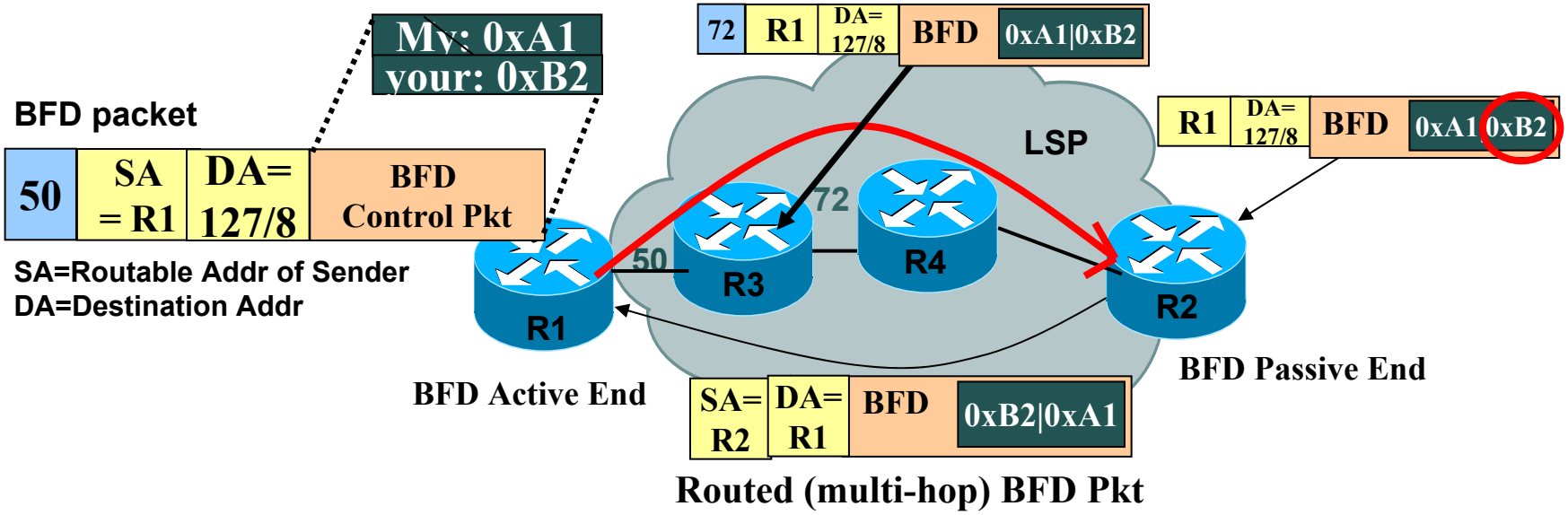
Boot-strapped of BFD Session using LSP-Ping



- Out-of-band Discriminator Signalling
- Echo Request and Reply carries BFD local discriminator values in "*BFD discriminator TLV*".

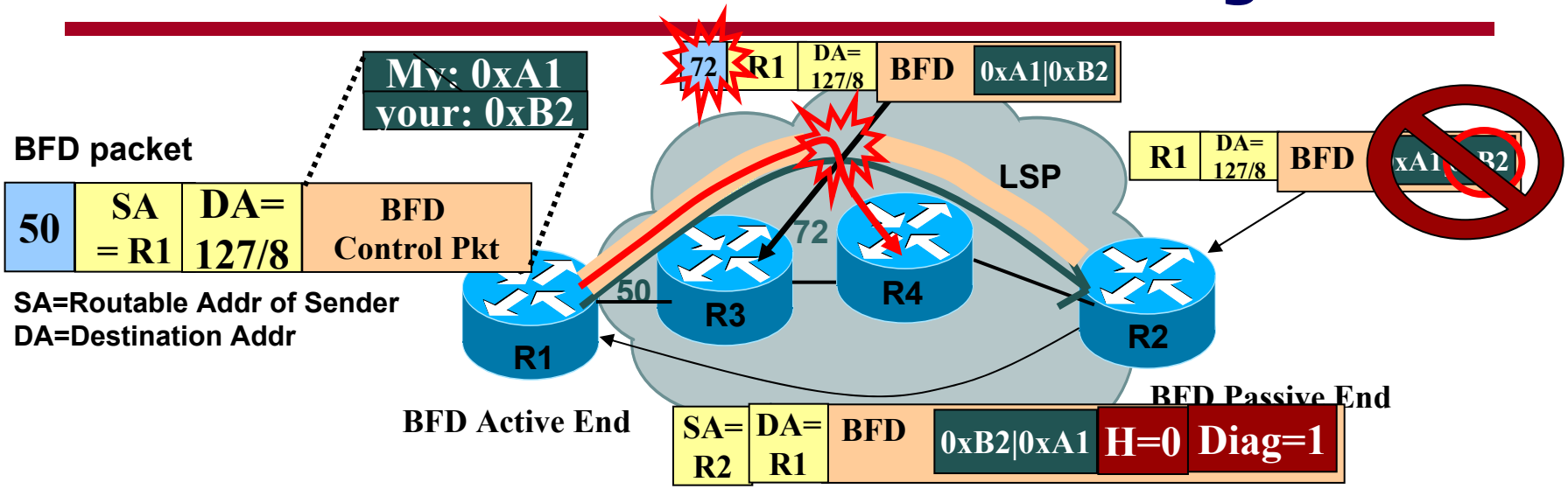


BFD Over Uni-directional MPLS LSP



- BFD control packets are encapsulated in the MPLS label stack that corresponds to the FEC under fault detection.
- The destination IP address is an address from 127/8.
- Egress LSR is single hop from BFD prospective,
  - The IP TTL is set to 1.
- Egress LSR routes BFD packet based on the destination IP address.

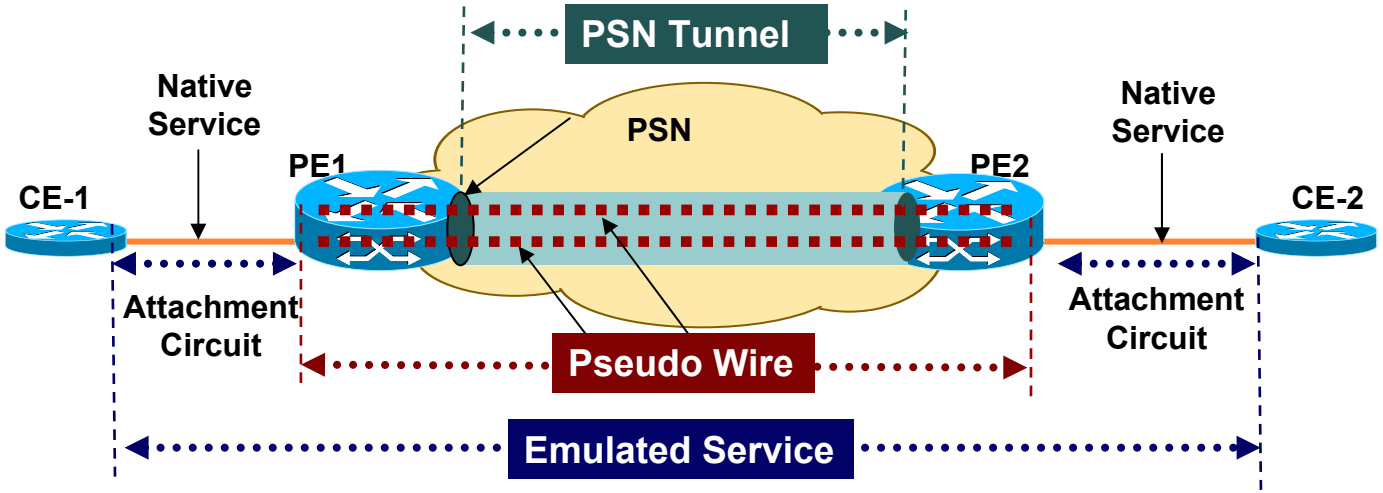
Data Plane Failure Detection Using BFD



I don't hear you; diag: "Control Detection Time Expired"

- Presence of the 127/8 address in the IP header destination address field causes the packet to be consumed by any routers trying to forward the packet using the ip header.
- In this case R4 would not forward the BFD packet to R2 but rather consumes the packet.
- R1 and R2 would declare the BFD session down.
- BFD Session down event would be notified to MPLS application

Virtual Circuit Connection Verification (VCCV)



- **Motivation**
  - Need mechanisms for connectivity verification of PW
  - 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

VCCV BFD Vs. VCCV Ping

Method	Data Plane Failure Detection	Control Plane Consistency	Protocol Overhead
VCCV Ping	**YES**	**YES**	**Higher than BFD**
VCCV-BFD	**YES**	**NO**	**Low**

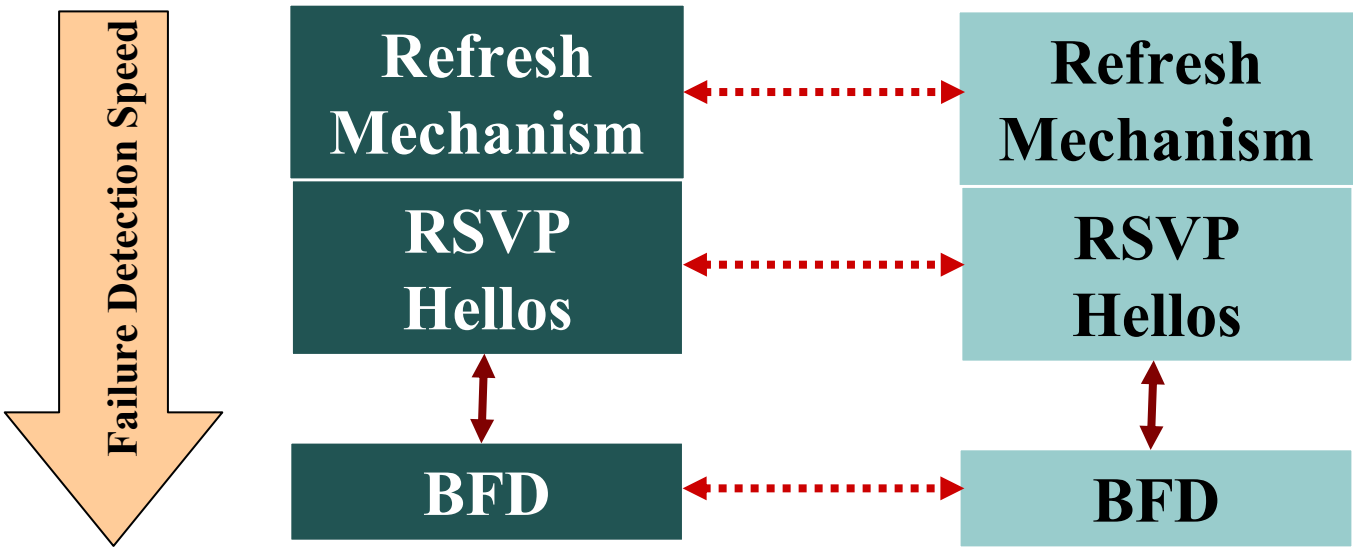
VCCV-BFD can complement VCCV-LSP Ping to detect a data plane failure in the forwarding path of a PW

VCCV-BFD works over MPLS or IP networks; Multiple PSN Tunnel Type MPLS, IPSEC, L2TP, GRE, etc.

VCCV-BFD

- Bootstrapping
  - BFD Capability Exchange: A new CV type for BFD is defined to allow exchange of BFD capability.
  - The use of the control channel provides the context required to bind the BFD session to a particular pseudo wire (FEC).
  - LSP Ping can be used to bootstrap BFD session.
- Encapsulation
  - BFD control packets sent by both ingress as well as egress LSR are encapsulated in the MPLS label stack that corresponds to the FEC of PW under connectivity verification.
- Failures in each direction can independently be monitored using the same BFD session.
- Asynchronous Mode of Operation (only)

BFD for MPLS TE

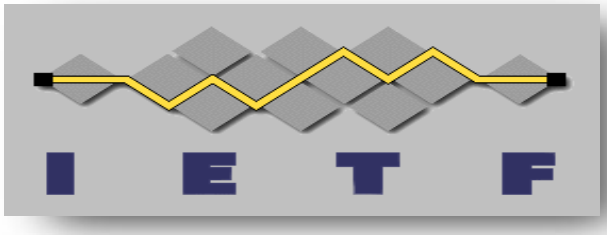


- BFD can be used to complement or replace use of RSVP Hellos for
  - MPLS FRR Link/ Node Protection
  - Health check for (PSC) FA-LSP

Outline

- History of BFD
- Brief overview of BFD Protocol
- Applicability of BFD in MPLS
- BFD updates from IETF
- Engineering Considerations
- Summary

BFD Updates from IETF



- BFD WG was formed before IETF 60.
- BFD is under IETF Routing Area.
- The following WG documents are quite stable
  - Base Spec (Single Hop): draft-ietf-bfd-base-00.txt
  - BFD Multi-hop: draft-ietf-bfd-multihop-00.txt
  - BFD MPLS: draft-ietf-bfd-mpls-00.txt
  - MIB: draft-ietf-bfd-mib-00.txt.
  - BFD for IPv4 and IPv6 (Single Hop): draft-ietf-bfd-v4v6-1hop-00.txt
- Mailing List
  - General Discussion: rtg-bfd@ietf.org
    To Subscribe: rtg-bfd-request@ietf.org
    In Body: With a subject line: subscribe
    Archive: ftp://ftp.ietf.org/ietf-mail-archive/rtg-bfd/

Outline

- History of BFD
- Brief overview of BFD Protocol
- Applicability of BFD in MPLS
- BFD updates from IETF
- Engineering Considerations
- Summary

Engineering Considerations

- BFD for VCCV and MPLS LSP induces a requirement on BFD to scale to a large number of sessions.
- Distributed BFD implementations are needed to deal with the scalability requirement.
- Aggressive detection interval with large number of BFD sessions increases the chance of False-Positives.
- If MPLS LSPs is fast reroutable, BFD fault detection interval should be greater than the fast-reroute switchover time.

Fast detection of BFD for MPLS LSPs and VCCV needs more study and operational input

Outline

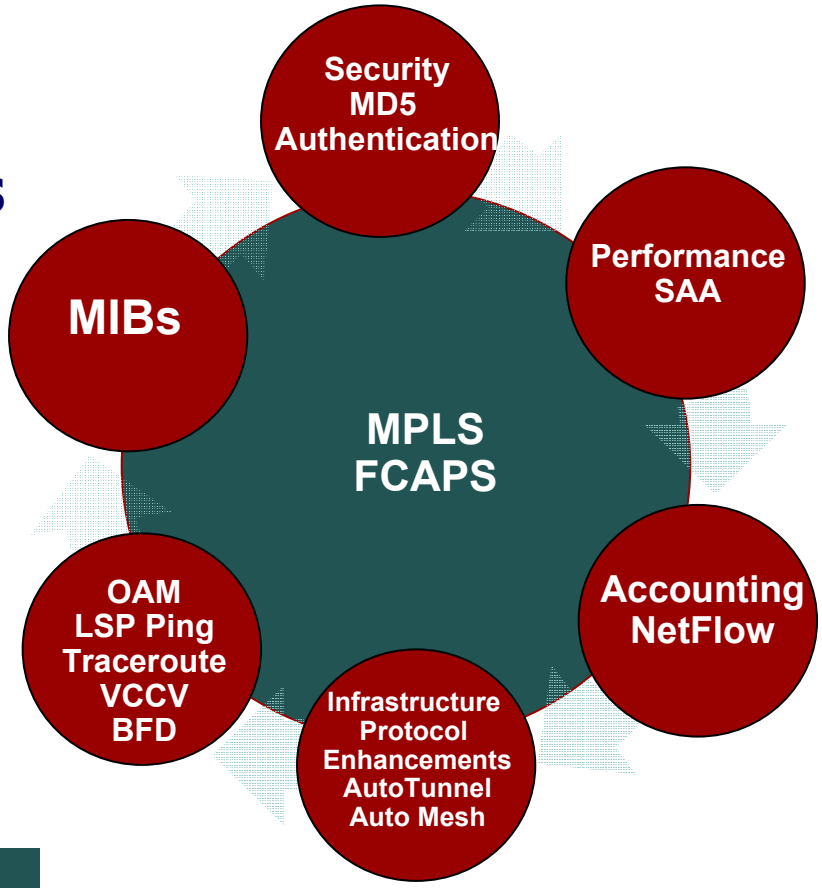
- History of BFD
- Brief overview of BFD Protocol
- Applicability of BFD in MPLS
- BFD updates from IETF
- Engineering Considerations
- Summary

Putting to Together

Media Type	Data Plane Connectivity Check	Loopback Test
MPLS LDP	MPLS BFD	LSP Ping
MPLS TE		VCCV Ping
MPLS PW	VCCV BFD	VCCV Ping
IPv4/ IPv6	BFD	IP Ping

Conclusion

- A collection of tools and techniques are required to address O&M requirements of MPLS.
- These tool must address FCAPS of MPLS O&M
  - Fault-Management
  - Configuration Management
  - Accounting
  - Performance
  - Security
- BFD is one of the tools that can be used for data plan connectivity for physical links, MPLS LSP or PW.
- Light-weight nature of BFD makes it attractive for these applications.



Fast detection of BFD for MPLS LSPs and VCCV needs more study and operational input


CISCO SYSTEMS