

Internet Engineering Task Force
Internet-Draft
Updates: 5880 (if approved)
Intended status: Standards Track
Expires: February 2, 2015

N. Akiya
C. Pignataro
D. Ward
Cisco Systems
M. Bhatia
Ionos Networks
P. K. Santosh
Juniper Networks
August 1, 2014

Seamless Bidirectional Forwarding Detection (S-BFD)
draft-ietf-bfd-seamless-base-02

Abstract

This document defines a simplified mechanism to use Bidirectional Forwarding Detection (BFD) with large portions of negotiation aspects eliminated, thus providing benefits such as quick provisioning as well as improved control and flexibility to network nodes initiating the path monitoring.

This document updates RFC5880.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 2, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Terminology	3
3. Seamless BFD Overview	4
4. S-BFD Discriminators	5
4.1. Discriminator Pools	5
4.2. S-BFD Discriminator Uniqueness	6
5. Reflector BFD Session	7
6. State Variables	7
6.1. New State Variables	7
6.2. State Variable Initialization and Maintenance	8
7. S-BFD Procedures	8
7.1. S-BFD Packet Demultiplexing	8
7.2. Initiator Procedures	8
7.2.1. SBFDInitiator State Machine	9
7.2.2. Details of S-BFD Packet Sent by SBFDInitiator	10
7.3. Responder Procedures	10
7.3.1. Responder Demultiplexing	10
7.3.2. Details of S-BFD Packet Sent by SBFDReflector	11
7.4. Diagnostic Values	11
7.5. The Poll Sequence	11
7.6. Control Plane Independent (C)	11
7.7. Additional SBFDInitiator Behaviors	12
7.8. Additional SBFDReflector Behaviors	12
8. Scaling Aspect	13
9. Co-existence with Classical BFD Sessions	13
10. S-BFD Echo Function	13
11. Security Considerations	14
12. IANA Considerations	15
13. Acknowledgements	15
14. Contributing Authors	15
15. References	16

15.1. Normative References	16
15.2. Informative References	16
Appendix A. Loop Problem	16
Authors' Addresses	18

1. Introduction

Bidirectional Forwarding Detection (BFD), [RFC5880] and related documents, has efficiently generalized the failure detection mechanism for multiple protocols and applications. There are some improvements which can be made to better fit existing technologies. There is a possibility of evolving BFD to better fit new technologies. This document focuses on several aspects of BFD in order to further improve efficiency, to expand failure detection coverage and to allow BFD usage for wider scenarios. This document extends BFD to provide solutions to use cases listed in [I-D.ietf-bfd-seamless-use-case].

One key aspect of the mechanism described in this document eliminates the time between a network node wanting to perform a continuity test and completing the continuity test. In traditional BFD terms, the initial state changes from DOWN to UP are virtually nonexistent. Removal of this seam (i.e. time delay) in BFD provides applications a smooth and continuous operational experience. Therefore, "Seamless BFD" (S-BFD) has been chosen as the name for this mechanism.

2. Terminology

The reader is expected to be familiar with the BFD, IP and MPLS terminologies and protocol constructs. This section describes several new terminologies introduced by S-BFD.

- o Classical BFD - BFD session types based on [RFC5880].
- o S-BFD - Seamless BFD.
- o S-BFD packet - a BFD control packet destined to or sourced from the well-known S-BFD port.
- o Entity - a function on a network node that S-BFD mechanism allows remote network nodes to perform continuity test to. An entity can be abstract (ex: reachability) or specific (ex: IP addresses, router-IDs, functions).
- o SBFDInitiator - an S-BFD session on a network node that performs a continuity test to a remote entity by sending S-BFD packets.

- o SBFDRReflector - an S-BFD session on a network node that listens for incoming S-BFD packets to local entities and generates response S-BFD packets.
- o Reflector BFD session - synonymous with SBFDRReflector.
- o S-BFD discriminator - a BFD discriminator allocated for a local entity and is being listened by an SBFDRReflector.
- o BFD discriminator - a BFD discriminator allocated for an SBFDRInitiator.
- o Initiator - a network node hosting an SBFDRInitiator.
- o Responder - a network node hosting an SBFDRReflector.

Below figure describes the relationship between S-BFD terminologies.

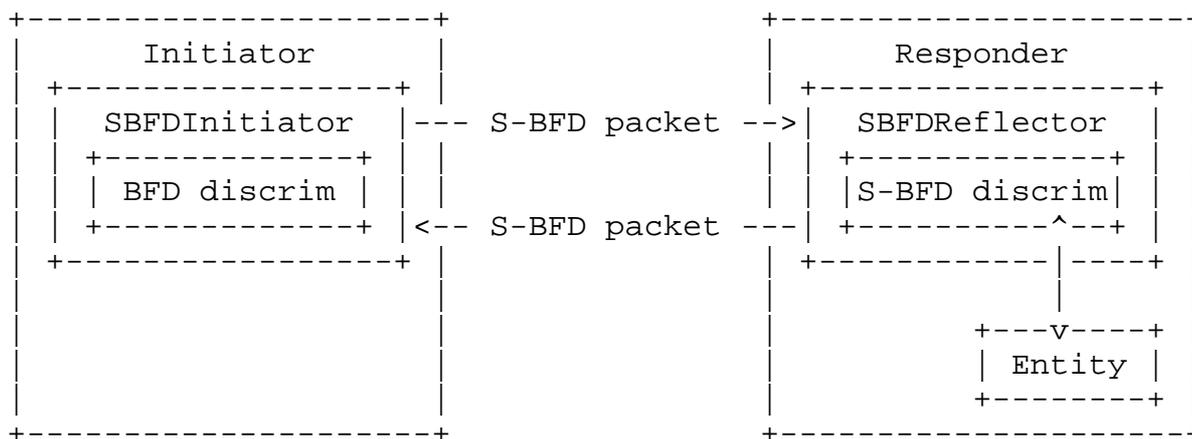


Figure 1: S-BFD Terminology Relationship

3. Seamless BFD Overview

An S-BFD module on each network node allocates one or more S-BFD discriminators for local entities, and creates a reflector BFD session. Allocated S-BFD discriminators may be advertised by applications (ex: OSPF/IS-IS). Required result is that applications, on other network nodes, possess the knowledge of the mapping from remote entities to S-BFD discriminators. The reflector BFD session is to, upon receiving an S-BFD packet targeted to one of local S-BFD discriminator values, transmit a response S-BFD packet back to the initiator.

Once above setup is complete, any network nodes, having the knowledge of the mapping from a remote entity to an S-BFD discriminator, can

- o SBFDRReflector is to allocate a discriminator from the S-BFD discriminator pool. The S-BFD discriminator pool SHOULD be a separate pool than the BFD discriminator pool.

Remainder of this subsection describes the reasons for above suggestions.

Locally allocated S-BFD discriminator values for entities, listened by SBFDRReflector sessions, may be arbitrary allocated or derived from values provided by applications. These values may be protocol IDs (ex: System-ID, Router-ID) or network targets (ex: IP address). To avoid derived S-BFD discriminator values already being assigned to other BFD sessions (i.e. SBFDRInitiator sessions and classical BFD sessions), it is RECOMMENDED that discriminator pool for SBFDRReflector sessions be separate from other BFD sessions.

Even when following the separate discriminator pool approach, collision is still possible between one S-BFD application to another S-BFD application, that may be using different values and algorithms to derive S-BFD discriminator values. If the two applications are using S-BFD for a same purpose (ex: network reachability), then the colliding S-BFD discriminator value can be shared. If the two applications are using S-BFD for a different purpose, then the collision must be addressed. How such collisions are addressed is outside the scope of this document.

4.2. S-BFD Discriminator Uniqueness

One important characteristics of an S-BFD discriminator is that it MUST be unique within an administrative domain. If multiple network nodes allocated a same S-BFD discriminator value, then S-BFD packets falsely terminating on a wrong network node can result in a reflector BFD session to generate a response back, due to "your discriminator" matching. This is clearly not desirable. If only IP based S-BFD is considered, then it is possible for the reflector BFD session to require demultiplexing of incoming S-BFD packets with combination of destination IP address and "your discriminator". Then S-BFD discriminator only has to be unique within a local node. However, S-BFD is a generic mechanism defined to run on wide range of environments: IP, MPLS, etc. For other transports like MPLS, because of the need to use non-routable IP destination address, it is not possible for reflector BFD session to demultiplex using IP destination address. With PHP, there may not be any incoming label stack to aid in demultiplexing either. Thus, S-BFD imposes a requirement that S-BFD discriminators MUST be unique within an administrative domain.

5. Reflector BFD Session

Each network node creates one or more reflector BFD sessions. This reflector BFD session is a session which transmits S-BFD packets in response to received S-BFD packets with "your discriminator" having S-BFD discriminators allocated for local entities. Specifically, this reflector BFD session is to have following characteristics:

- o MUST NOT transmit any S-BFD packets based on local timer expiry.
- o MUST transmit an S-BFD packet in response to a received S-BFD packet having a valid S-BFD discriminator in the "your discriminator" field, unless prohibited by local policies (ex: administrative, security, rate-limiter, etc).
- o MUST be capable of sending only two states: UP and ADMINDOWN.

One reflector BFD session may be responsible for handling received S-BFD packets targeted to all locally allocated S-BFD discriminators, or few reflector BFD sessions may each be responsible for subset of locally allocated S-BFD discriminators. This policy is a local matter, and is outside the scope of this document.

Note that incoming S-BFD packets may be IPv4, IPv6 or MPLS based. How such S-BFD packets reach an appropriate reflector BFD session is also a local matter, and is outside the scope of this document.

6. State Variables

S-BFD introduces new state variables, and modifies the usage of existing ones.

6.1. New State Variables

A new state variable is added to the base specification in support of S-BFD.

- o `bfd.SessionType`: The type of this session. Allowable values are:
 - * `SBFDInitiator` - an S-BFD session on a network node that performs a continuity test to a target entity by sending S-BFD packets.
 - * `SBFDReflector` - an S-BFD session on a network node that listens for incoming S-BFD packets to local entities and generates response S-BFD packets.

bfd.SessionType variable MUST be initialized to the appropriate type when an S-BFD session is created.

6.2. State Variable Initialization and Maintenance

Some state variables defined in section 6.8.1 of the BFD base specification need to be initialized or manipulated differently depending on the session type.

- o bfd.DemandMode: This variable MUST be initialized to 1 for session type SBFDDiscriminator, and MUST be initialized to 0 for session type SBFDDiscriminator.

7. S-BFD Procedures

7.1. S-BFD Packet Demultiplexing

Received BFD control packet MUST first be demultiplexed with information from the lower layer (ex: destination UDP port, associated channel type). If the packet is determined to be for an SBFDDiscriminator, then the packet MUST be looked up to locate a corresponding SBFDDiscriminator session based on the value from the "your discriminator" field in the table describing S-BFD discriminators. If the packet is determined not to be for SBFDDiscriminator, then the packet MUST be looked up to locate a corresponding SBFDDiscriminator session or classical BFD session based on the value from the "your discriminator" field in the table describing BFD discriminators. If the located session is a SBFDDiscriminator, then destination of the packet (i.e. destination IP address) SHOULD be validated to be for self.

Details of the initial BFD control packet demultiplexing are described in relevant S-BFD data plane documents.

7.2. Initiator Procedures

S-BFD packets transmitted by an SBFDDiscriminator MUST set "your discriminator" field to an S-BFD discriminator corresponding to the remote entity.

Every SBFDDiscriminator MUST have a locally unique "my discriminator" allocated from the BFD discriminator pool.

Below ASCII art describes high level concept of continuity test using S-BFD. R2 allocates XX as the S-BFD discriminator for its network reachability purpose, and advertises XX to neighbors. ASCII art shows R1 and R4 performing a continuity test to R2.

Note that the above state machine is different from the base BFD specification [RFC5880]. This is because the INIT state is no longer applicable for the SBFDInitiator. Another important difference is the transition of the state machine from the DOWN state to the UP state when a packet with State UP is received by the SBFDInitiator. The definitions of the states and the events have the same meaning as in the base BFD specification [RFC5880].

7.2.2. Details of S-BFD Packet Sent by SBFDInitiator

S-BFD packets sent by an SBFDInitiator is to have following contents:

- o "my discriminator" assigned by local node.
- o "your discriminator" corresponding to a remote entity.
- o "State" MUST be set to a value describing local state.
- o "Desired Min TX Interval" MUST be set to a value describing local desired minimum transmit interval.
- o "Required Min RX Interval" MUST be zero.
- o "Required Min Echo RX Interval" SHOULD be zero.
- o "Detection Multiplier" MUST be set to a value describing locally used multiplier value.
- o Demand (D) bit MUST be set.

7.3. Responder Procedures

A network node which receives S-BFD packets transmitted by an initiator is referred as responder. The responder, upon reception of S-BFD packets, is to perform necessary relevant validations described in [RFC5880], [RFC5881], [RFC5883], [RFC5884] and [RFC5885].

7.3.1. Responder Demultiplexing

When a responder receives an S-BFD packet, if the value in the "your discriminator" field is not one of S-BFD discriminators allocated for local entities, then this packet MUST NOT be considered for this mechanism. If the value in the "your discriminator" field is one of S-BFD discriminators allocated for local entities, then the packet is determined to be handled by a reflector BFD session responsible for the S-BFD discriminator. If the packet was determined to be processed further for this mechanism, then chosen reflector BFD session is to transmit a response BFD control packet using procedures described in Section 7.3.2, unless prohibited by local policies (ex: administrative, security, rate-limiter, etc).

7.3.2. Details of S-BFD Packet Sent by SBFDRReflector

S-BFD packets sent by an SBFDRReflector is to have following contents:

- o "my discriminator" MUST be copied from received "your discriminator".
- o "your discriminator" MUST be copied from received "my discriminator".
- o "State" MUST be UP or ADMINDOWN. Clarification of reflector BFD session state is described in Section 7.8.
- o "Desired Min TX Interval" MUST be copied from received "Desired Min TX Interval".
- o "Required Min RX Interval" MUST be set to a value describing how many incoming control packets this reflector BFD session can handle. Further details are described in Section 7.8.
- o "Required Min Echo RX Interval" SHOULD be set to zero.
- o "Detection Multiplier" MUST be copied from received "Detection Multiplier".
- o Demand (D) bit MUST be cleared.

7.4. Diagnostic Values

Diagnostic value in both directions MAY be set to a certain value, to attempt to communicate further information to both ends. However, details of such are outside the scope of this specification.

7.5. The Poll Sequence

Poll sequence MAY be used in both directions. The Poll sequence MUST operate in accordance with [RFC5880]. An SBFDRReflector MAY use the Poll sequence to slow down that rate at which S-BFD packets are generated from an SBFDRInitiator. This is done by the SBFDRReflector using procedures described in Section 7.8 and setting the Poll (P) bit in the reflected S-BFD packet. The SBFDRInitiator is to then send the next S-BFD packet with the Final (F) bit set. If an SBFDRReflector receives an S-BFD packet with Poll (P) bit set, then the SBFDRReflector MUST respond with an S-BFD packet with Poll (P) bit cleared and Final (F) bit set.

7.6. Control Plane Independent (C)

Control plane independent (C) bit for an SBFDRInitiator sending S-BFD packets to a reflector BFD session MUST work according to [RFC5880]. Reflector BFD session also MUST work according to [RFC5880]. Specifically, if reflector BFD session implementation does not share fate with control plane, then response S-BFD packets transmitted MUST have control plane independent (C) bit set. If reflector BFD session implementation shares fate with control plane, then response S-BFD

packets transmitted MUST NOT have control plane independent (C) bit set.

7.7. Additional SBFDDInitiator Behaviors

- o If the SBFDDInitiator receives a valid S-BFD packet in response to transmitted S-BFD packet to a remote entity, then the SBFDDInitiator SHOULD conclude that S-BFD packet reached the intended remote entity.
- o When a sufficient number of S-BFD packets have not arrived as they should, the SBFDDInitiator SHOULD declare loss of reachability to the remote entity. The criteria for declaring loss of reachability and the action that would be triggered as a result are outside the scope of this document.
- o Relating to above bullet item, it is critical for an implementation to understand the latency to/from the reflector BFD session on the responder. In other words, for very first S-BFD packet transmitted by the SBFDDInitiator, an implementation MUST NOT expect response S-BFD packet to be received for time equivalent to sum of latencies: initiator to responder and responder back to initiator.
- o If the SBFDDInitiator receives an S-BFD packet with Demand (D) bit set, the packet MUST be discarded.

7.8. Additional SBFDDReflector Behaviors

- o S-BFD packets transmitted by the SBFDDReflector MUST have "Required Min RX Interval" set to a value which expresses how many incoming S-BFD packets this SBFDDReflector can handle. The SBFDDReflector can control how fast SBFDDInitiators will be sending S-BFD packets to self by ensuring "Required Min RX Interval" indicates a value based on the current load.
- o If the SBFDDReflector wishes to communicate to some or all SBFDDInitiators that monitored local entity is "temporarily out of service", then S-BFD packets with "state" set to ADMINDOWN are sent to those SBFDDInitiators. The SBFDDInitiators, upon reception of such packets, MUST NOT conclude loss of reachability to corresponding remote entity, and MUST back off packet transmission interval for the remote entity to an interval no faster than 1 second. If the SBFDDReflector is generating a response S-BFD packet for a local entity that is in service, then "state" in response BFD control packets MUST be set to UP.

- o If an SBFDRreflector receives an S-BFD packet with Demand (D) bit cleared, the packet MUST be discarded.

8. Scaling Aspect

This mechanism brings forth one noticeable difference in terms of scaling aspect: number of SBFDRreflector. This specification eliminates the need for egress nodes to have fully active BFD sessions when only one side desires to perform continuity tests. With introduction of reflector BFD concept, egress no longer is required to create any active BFD session per path/LSP/function basis. Due to this, total number of BFD sessions in a network is reduced.

9. Co-existence with Classical BFD Sessions

Initial packet demultiplexing requirement is described in Section 7.1. Because of this, S-BFD mechanism can co-exist with classical BFD sessions.

10. S-BFD Echo Function

The concept of the S-BFD Echo function is similar to the BFD Echo function described in [RFC5880], packets are self-generated and self-terminated after traversing a link/path. S-BFD echo packets are expected to u-turn on the target node in the data plane and MUST NOT be processed by any reflector BFD sessions on the target node.

When using the S-BFD Echo function, it is RECOMMENDED that:

- o Both S-BFD packets (with BFD control header) and S-BFD echo packets (implementation specific) be sent.
- o Both S-BFD packets and S-BFD echo packets have the same semantics in the forward direction to reach the target node.

In other words, it is not preferable to send just S-BFD echo packets. There are two reason behind this suggestion:

- o S-BFD packets can verify reachability to intended target node, which allows one to conclude that S-BFD echo packets are u-turning on the expected target node.
- o S-BFD packets can detect when the target node is going out of service (i.e. via receiving back ADMINDOWN state).

Implementations MAY set "Required Min Echo RX Interval" field to indicate the rate which SBFDRInitiator is sending S-BFD Echo packets

(in ping) or the rate which SBFDReflector wants SBFDDInitiators to send S-BFD Echo packets (in pong). However, this is likely more than necessary for the S-BFD Echo function to operate. Therefore, it is RECOMMENDED that "Required Min Echo RX Interval" field simply be set to zero in both directions.

Additionally, following aspects are left as implementation details, and are outside the scope of this document:

- o Format of the S-BFD Echo packet (ex: data beyond UDP header).
- o Procedures on when and how to use the S-BFD Echo function.

11. Security Considerations

Same security considerations as [RFC5880], [RFC5881], [RFC5883], [RFC5884] and [RFC5885] apply to this document. Additionally, implementing the following measures will strengthen security aspects of the mechanism described by this document:

- o SBFDDInitiator MAY pick crypto sequence number based on authentication mode configured.
- o SBFDReflector MUST NOT look at the crypto sequence number before accepting the packet.
- o SBFDReflector MAY look at the Key ID [I-D.ietf-bfd-generic-crypto-auth] in the incoming packet and verify the authentication data.
- o SBFDReflector MUST accept the packet if authentication is successful.
- o SBFDReflector MUST compute the Authentication data and MUST use the same sequence number that it received in the S-BFD packet that it is responding to.
- o SBFDDInitiator MUST accept the S-BFD packet if it either comes with the same sequence number as it had sent or it's within the window that it finds acceptable (described in detail in [I-D.ietf-bfd-generic-crypto-auth])

Using the above method,

- o SBFDReflector continue to remain stateless despite using security.

- o SBFDDReflector are not susceptible to replay attacks as they always respond to S-BFD packets irrespective of the sequence number carried.
- o An attacker cannot impersonate the responder since the SBFDDInitiator will only accept S-BFD packets that come with the sequence number that it had originally used when sending the S-BFD packet.

12. IANA Considerations

No action is required by IANA for this document.

13. Acknowledgements

Authors would like to thank Jeffrey Haas, Greg Mirsky and Marc Binderberger for performing thorough reviews and providing number of suggestions. Authors would like to thank Girija Raghavendra Rao, Les Ginsberg, Srihari Raghavan, Vanitha Neelamegam and Vengada Prasad Govindan from Cisco Systems for providing valuable comments. Authors would also like to thank John E. Drake and Pablo Frank for providing comments and suggestions.

14. Contributing Authors

Tarek Saad
Cisco Systems
Email: tsaad@cisco.com

Siva Sivabalan
Cisco Systems
Email: msiva@cisco.com

Nagendra Kumar
Cisco Systems
Email: naikumar@cisco.com

Mallik Mudigonda
Cisco Systems
Email: mmudigon@cisco.com

Sam Aldrin
Huawei Technologies
Email: aldrin.ietf@gmail.com

Assume node A reserved a discriminator 0x01010101 for target identifier 192.0.2.1 and has a reflector session in listening mode. Similarly node B reserved a discriminator 0x02020202 for its target identifier 192.0.2.2 and also has a reflector session in listening mode.

Suppose MiM sends a spoofed packet with MyDisc = 0x01010101, YourDisc = 0x02020202, source IP as 192.0.2.1 and dest IP as 192.0.2.2. When this packet reaches Node B, the reflector session on Node B will swap the discriminators and IP addresses of the received packet and reflect it back, since YourDisc of the received packet matched with reserved discriminator of Node B. The reflected packet that reached Node A will have MyDisc=0x02020202 and YourDisc=0x01010101. Since YourDisc of the received packet matched the reserved discriminator of Node A, Node A will swap the discriminators and reflects the packet back to Node B. Since reflectors MUST set the TTL of the reflected packets to 255, the above scenario will result in an infinite loop with just one malicious packet injected from MiM.

FYI: Packet fields do not carry any direction information, i.e., if this is Ping packet or reply packet.

Solutions

The current proposals to avoid the loop problem are:

- o Overload "D" bit (Demand mode bit): Initiator always sets the 'D' bit and reflector clears it. This way we can identify if a received packet was a reflected packet and avoid reflecting it back. However this changes the interpretation of 'D' bit.
- o Use of State field in the BFD control packets: Initiator will always send packets with State set to DOWN and reflector will send back packets with state field set to UP. Reflectors will never reflect any received packets with state as UP. However the only issue is the use of state field differently i.e. state in the S-BFD control packet from initiator does not reflect the local state which is anyway not significant at reflector.
- o Use of local discriminator as My Disc at reflector: Reflector will always fill in My Discriminator with a locally allocated discriminator value (not reserved discriminators) and will not copy it from the received packet.

Authors' Addresses

Nobo Akiya
Cisco Systems

Email: nobo@cisco.com

Carlos Pignataro
Cisco Systems

Email: cpignata@cisco.com

Dave Ward
Cisco Systems

Email: wardd@cisco.com

Manav Bhatia
Ionos Networks

Email: manav@ionosnetworks.com

Santosh
Juniper Networks

Email: santoshpk@juniper.net