Internet Engineering Task Force
Internet-Draft
Updates: 5880 (if approved)
Intended status: Standards Track
Expires: December 14, 2014

N. Akiya
C. Pignataro
D. Ward
Cisco Systems
M. Bhatia
Alcatel-Lucent
P. K. Santosh
Juniper Networks
June 12, 2014

Seamless Bidirectional Forwarding Detection (S-BFD)
draft-ietf-bfd-seamless-base-00

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.

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

Copyright Notice

Table of Contents

1. Introduction 3
2. Seamless BFD Overview 3
3. Terminology 4
4. BFD Target Identifier Types 5
5. UDP Port 5
6. S-BFD Discriminators 5
7. Reflector BFD Session 7
8. State Variables 7
  8.1. New State Variables 7
  8.2. State Variable Initialization and Maintenance 8
9. Full Reachability Validations 8
  9.1. Initiator Behavior 8
    9.1.1. Initiator State machine 9
  9.2. Responder Behavior 10
    9.2.1. Responder Demultiplexing 10
    9.2.2. Reflector BFD Session Procedures 10
  9.3. Further Packet Details 12
  9.4. Diagnostic Values 12
  9.5. The Poll Sequence 13
  9.6. Control Plane Independent (C) 13
  9.7. Additional Initiator Behavior 13
  9.8. Additional Responder Behavior 13
10. Partial Reachability Validations 14
11. Scaling Aspect 14
12. Co-existence with Traditional BFD 15
13. BFD Echo 15
14. Security Considerations 15
15. IANA Considerations 16
16. Acknowledgements 16
17. Contributing Authors 16
18. References 17

18.1. Normative References 17
18.2. Informative References 17
Appendix A. Loop Problem 18
Authors' Addresses 19

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]. Because defined mechanism eliminates much of negotiation aspects of the BFD protocol, "Seamless BFD" (S-BFD) has been chosen as the name for this mechanism.

2. Seamless BFD Overview

Each protocol instance (e.g. OSPF/IS-IS) allocates one or more BFD discriminators on its network node, ensuring that BFD discriminators allocated are unique within the network domain. Allocated BFD discriminators may be advertised by the protocol. Required result is that a protocol possess the knowledge of mapping between network targets to BFD discriminators. Each network nodes will also create a BFD session instance that listens for incoming BFD control packets with "your discriminator" having protocol allocated values. The listener BFD session instance, upon receiving a BFD control packet targeted to one of local S-BFD discriminator values, will transmit a response BFD control packet back to the sender.

Once above setup is complete, any network node, understanding the mapping between network targets to BFD discriminators, can quickly perform reachability check to these network targets by simply sending BFD control packets with known BFD discriminator value as "your discriminator".

For example:

```
   <------- IS-IS Network ------->

             +---------+
             |         |
   A---------B---------C---------D
   ^                             ^
   |                             |
SystemID                      SystemID
  xxx                           yyy
BFD Discrim                   BFD Discrim
  123                           456
```

Figure 1: S-BFD for IS-IS Network

IS-IS with SystemID xxx allocates BFD discriminator 123, and advertises the BFD discriminator 123 in IS-IS TLV. IS-IS with SystemID yyy allocates BFD discriminator 456, and advertises the BFD discriminator 456 IS-IS TLV. Both network nodes (node A and node D) creates listener BFD session instance. When network node A wants to check a reachability to network node D, node A can send a BFD control packet, destined to node D, with "your discriminator" set as 456. If listener BFD on node D receives this BFD control packet, then response BFD control packet is sent back to node A, which allows node A to complete the reachability test.

Note that a protocol may create an explicit mapping between a protocol ID (e.g. System-ID, Router-ID) to a BFD discriminator. A protocol may also create an explicit mapping between a network target (e.g. IP address) to a BFD discriminator. A protocol may even function with implicit mapping between a network target (e.g. IPv4 address) to a BFD discriminator, i.e. IPv4 address is used as BFD discriminator value. Decisions and rules on how protocols allocate and distribute BFD discriminators are outside the scope of this document.

3. Terminology

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

- o BFD Target Identifier: Network entity that is provisioned as a target of Seamless BFD.

- o BFD Target Identifier Type: Type of network entity that is provisioned as a target of Seamless BFD.

- BFD Target Identifier Table: A table containing BFD target identifier type, BFD target identifier and corresponding BFD discriminator.

- Reflector BFD Session: A BFD session listening for incoming BFD control packets destined for local BFD target identifier(s).

4. BFD Target Identifier Types

This document defines a generic mechanism where network nodes can send BFD control packets to specific network targets to perform various tasks. One task is to perform a reachability check (i.e requesting immediate response back). Details of this task is further defined in sections to follow. Further tasks (i.e. using BFD control packet to request specific services from specific network nodes) may be defined. Therefore, this document defines a code point for BFD Target Identifier. Each locally allocated S-BFD discriminator MUST be associated to BFD Target Identifier type, to allow demultiplexing to a specific task or service.

BFD Target Identifier types:

Value	BFD Target Identifier Type
0	Reserved
1	Network Target Discriminator

Procedures defined in this document are to be associated with BFD Target Identifier Type 1 (Network Target Discriminator).

Note that IP based BFD from [RFC5885] is supported by this specification, but non-IP based BFD is outside the scope of this document.

Further identifier types are to be defined as needed basis.

5. UDP Port

S-BFD functions on a well-known UDP port: TBD1.

6. S-BFD Discriminators

Protocols (i.e. client of S-BFD) may request an arbitrary BFD discriminator value, or protocols may request a specific BFD discriminator value. Therefore, it is RECOMMENDED for implementations to create a separate discriminator pool for S-BFD sessions to minimize the collision between existing BFD sessions and S-BFD sessions. In such case, incoming BFD control packets MUST be demultiplexed first with UDP port to identify the discriminator table to look up the session. Regardless of the approach, collision can happen with following scenarios.

- o Existing BFD session already using a discriminator value that collides with specific discriminator value requested for S-BFD session.

  * Implementation SHOULD allow migrating existing BFD sessions to free up the discriminator to accommodate specific discriminator value requested for S-BFD session.

- o S-BFD session already using a discriminator value, arbitrarily allocated, that collides with specific discriminator value requested for S-BFD session. The two S-BFD sessions are of different BFD Target Identifier type.

  * Protocol requesting arbitrary discriminator value MUST support migrating to another discriminator value, and implementations MUST allow migrating existing S-BFD sessions to free up the discriminator to accommodate specific discriminator value requested for S-BFD session.

- o S-BFD session already using a discriminator value, arbitrary allocated, that collides with specific discriminator value requested for S-BFD session. The two S-BFD sessions are of same BFD Target Identifier type.

  * No action is required, as the two can share the discriminator.

One important characteristics of S-BFD discriminator is that it MUST be network wide unique. If multiple network nodes allocated same S-BFD discriminator value, then S-BFD control packets falsely terminating on a wrong network node can result in reflector BFD session (described in Section 7) to generate a response back, due to "your discriminator" matching. This is clearly not desirable. If only IP based S-BFD is concerned, then it is possible for S-BFD reflector session to require demultiplexing of incoming S-BFD control packet 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, Segment Routing ([I-D.previdi-filsfils-isis-segment-routing]), etc. For other transports like MPLS, because of the need to use non-routable IP destination address, it is not possible for S-BFD reflector 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 network wide unique.

7. Reflector BFD Session

Each network node MUST create one or more reflector BFD sessions. This reflector BFD session is a session which transmits BFD control packets in response to received valid locally destined BFD control packets. Specifically, this reflector BFD session is to have following characteristics:

- o MUST NOT transmit any BFD control packets based on local timer expiry.

- o MUST transmit BFD control packet in response to a received valid locally destined BFD control packet.

- o MUST be capable of sending only two states: UP and ADMINDOWN.

One reflector BFD session MAY be responsible for handling received BFD control packets targeted to all local BFD target identifiers, or few reflector BFD sessions MAY each be responsible for subset of local BFD target identifiers. This policy is a local matter, and is outside the scope of this document.

Note that incoming BFD control packets destined to BFD target identifier types may be IPv4, IPv6 or MPLS based. For those BFD target identifier types, implementations MAY either allow the same reflector BFD session to handle all incoming BFD control packets in address family agnostic fashion, or setup multiple reflector BFD sessions to handle incoming BFD control packets with different address families. This policy is again a local matter, and is outside the scope of this document.

8. State Variables

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

8.1. 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: Any session on a network node that attempts to perform a path monitoring to any BFD target identifier on other network nodes.

* SBFDReflector: Any session on a network node, which receives BFD control packets transmitted by an initiator and responds back to initiator is referred as responder.

This variable MUST be initialized to the appropriate type when the session is created, according to the rules in section TBD.

8.2. 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 SBFDInitiator, and MUST be initialized to 0 for session type SBFDReflector.

9. Full Reachability Validations

9.1. Initiator Behavior

Any network node can attempt to perform a full reachability validation to any BFD target identifier on other network nodes, as long as destination BFD target identifier is provisioned to use this mechanism. BFD control packets transmitted by the initiator is to have "your discriminator" corresponding to destination BFD target identifier.

A node that initiates a BFD control packet MAY create an active BFD session to periodically send BFD control packets to a target, or a BFD control packet MAY be crafted and sent out on "as needed basis" (ex: BFD ping) without any session presence. In both cases, a BFD instance MUST have a unique "my discriminator" value assigned. If a node is to create multiple BFD instances to the same BFD target identifier, then each instance MUST have separate "my discriminator" values assigned. A BFD instance MUST NOT use a discriminator corresponding to one of local BFD target identifiers as "my discriminator". This is to prevent incoming response BFD control packets ("pong" packets) having "your discriminator" as a discriminator corresponding to the local BFD target identifier.

Below ASCII art describes high level concept of full reachability validations using this mechanism. R2 reserves value XX as BFD discriminator for its BFD target identifier. ASCII art shows that R1 and R4 performing full reachability validation to XX on R2.

```
 -- md=50/yd=XX (BFD ping) -->
<-- md=XX/yd=50 (BFD pong) --

                         [*]
 R1 --------------------- R2 ----------- R3 ----------- R4

                         |  ^
                         |  |
                         |  + - md=60/yd=XX (BFD ping) --
                         + - - -md=XX/yd=60 (BFD pong) -->

[*] Reflector BFD session on R2.
```

Figure 2: S-BFD path monitoring

If BFD control packet is to be sent via IP path, then:

- o Destination IP address MUST be an IP address corresponding to target identifier.
- o Source IP address MUST be a local IP address.
- o IP TTL MUST be 255 for full reachability validations. Partial reachability validations MAY use smaller TTL value (see Section 10).
- o Well-known UDP destination port(s) for IP based S-BFD.

If BFD control packet response is determined to explicitly be label switched, then:

- o BFD control packet MUST get imposed with a label stack that is expected to reach the target node.
- o MPLS TTL MUST be 255 for full reachability validations. Partial reachability validations MAY use smaller TTL value (see Section 10).
- o Destination IP address MUST be 127/8 for IPv4 and 0:0:0:0:0:FFFF:7F00/104 for IPv6.
- o Source IP address MUST be a local IP address.
- o IP TTL=1.
- o Well-known UDP destination port(s) for MPLS based S-BFD

9.1.1. Initiator State machine

The following diagram provides an overview of the initiator state machine. The notation on each arc represents the state of the remote system (as received in the State field in the BFD Control packet) or indicates the expiration of the Detection Timer.

```
                 +--+
ADMIN DOWN,      |  |
TIMER            |  V
              +------+   UP                +------+
              |      |-------------------->|      |----+
              | DOWN |                     |  UP  |    | UP
              |      |<--------------------|      |<---+
              +------+   ADMIN DOWN,       +------+
                         TIMER
```

Figure 3: S-BFD Initiator FSM

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 initiator of the S-BFD session. 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 initiator. The definitions of the states and the events have the same meaning as in the base BFD specification [RFC5880].

9.2. Responder Behavior

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

9.2.1. Responder Demultiplexing

When responder receives a BFD control packet, if "your discriminator" value is not one of local entries in the BFD target identifier table, then this packet MUST NOT be considered for this mechanism. If "your discriminator" value is one of local entries in the BFD target identifier table, then the packet is determined to be handled by a reflector BFD session responsible for specified BFD targeted identifier. 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 9.2.2, unless prohibited by local administrative or local policy reasons.

9.2.2. Reflector BFD Session Procedures

BFD target identifier type MUST be used to determine further information on how to reach back to the initiator.

In addition, destination IP address of received BFD control packet MUST be examined to determine how to construct response BFD control packet to send back to the initiator.

If destination IP address of received BFD control packet is not 127/8 for IPv4 or 0:0:0:0:0:FFFF:7F00/104 for IPv6, then:

- o Destination IP address MUST be copied from received source IP address.
- o Source IP address MUST be copied from received destination IP address if received destination IP address is a local address. Otherwise local IP address MUST be used.
- o IP TTL MUST be 255.

Response BFD control packet SHOULD be IP routed back, but MAY explicitly be label switched.

If BFD control packet response is determined to be IP routed, then:

- o Destination IP address MUST be copied from received source IP address.
- o Source IP address MUST be a local address.
- o IP TTL MUST be 255.

If BFD control packet response is determined to explicitly be label switched, then:

- o BFD control packet MUST get label switched back to the initiator. Determining the label stack to be imposed on a response BFD control packet is outside the scope of this document.
- o MPLS TTL MUST be 255.
- o Destination IP address MUST be 127/8 for IPv4 and 0:0:0:0:0:FFFF:7F00/104 for IPv6.
- o Source IP address MUST be a local IP address.
- o IP TTL MUST be 1.

Regardless of the response type, BFD control packet being sent by the responder MUST perform following procedures:

- o Copy "my discriminator" from received "your discriminator", and "your discriminator" from received "my discriminator".
- o UDP destination port MUST be same as received UDP destination port.

9.3. Further Packet Details

Further details of BFD control packets sent by initiator (ex: active BFD session):

- o Well-known UDP destination port assigned for S-BFD.
- o UDP source port as per described in [RFC5881]/[RFC5883]/[RFC5884]/[RFC5885].
- o "my discriminator" assigned by local node.
- o "your discriminator" corresponding to an identifier of target node.
- o "State" MUST be set to a value reflecting local state.
- o "Desired Min TX Interval" MUST be set to a value reflecting 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 reflecting locally used multiplier value.
- o "Demand bit (D)" MUST be set by the initiator.

Further details of BFD control packets sent by responder (reflector BFD session):

- o Well-known UDP destination port assigned for S-BFD.
- o UDP source port as described in [RFC5881]/[RFC5883]/[RFC5884]/[RFC5885].
- 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 9.8.
- 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 reflecting how many incoming control packets this reflector BFD session can handle.
- o "Required Min Echo RX Interval" SHOULD be set to zero.
- o "Detection Multiplier" MUST be copied from received "Detection Multiplier".
- o "Demand bit (D)" MUST be cleared by the reflector.

9.4. 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.

9.5. The Poll Sequence

The Poll sequence MUST operate in accordance with [RFC5880].

9.6. Control Plane Independent (C)

Control plane independent (C) bit for BFD instances speaking 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 BFD control packets transmitted MUST have control plane independent (C) bit set. If reflector BFD session implementation shares fate with control plane, then response BFD control packets transmitted MUST NOT have control plane independent (C) bit set.

9.7. Additional Initiator Behavior

- If initiator receives valid BFD control packet in response to transmitted BFD control packet, then initiator SHOULD conclude that packet reached intended target.

- When a sufficient number of BFD control packets have not arrived as they should, the initiator could declare loss of reachability. The criteria for declaring loss of reachability and the action that would be triggered as a result are outside the scope of this specification.

- Relating to above bullet item, it is critical for an implementation to understand the latency to/from reflector BFD session on target node. In other words, for very first BFD control packet transmitted, an implementation MUST NOT expect response BFD control packet to be received for time equivalent to sum of latencies: initiator node to target node and target node back to initiator node.

- If initiator receives a packet with D bit set, the packet MUST be discarded.

9.8. Additional Responder Behavior

- BFD control packets transmitted by a reflector BFD session MUST have "Required Min RX Interval" set to a value which reflects how many incoming control packets this reflector BFD session can handle. Responder can control how fast initiators will be sending BFD control packets to self by ensuring "Required Min RX Interval" reflects a value based on current load.

- o If a reflector BFD session wishes to communicate to some or all initiators that monitored BFD target identifier is "temporarily out of service", then BFD control packets with "state" set to ADMINDOWN are sent to those initiators. Initiators, upon reception of such packets, MUST NOT conclude loss of reachability to corresponding BFD target identifier, and MUST back off packet transmission interval to corresponding BFD target identifier an interval no faster than 1 second. If a reflector BFD session is generating a response BFD control packet for BFD target identifier that is in service, then "state" in response BFD control packets MUST be set to UP.

- o If a reflector receives a packet with D bit cleared, the packet MUST be discarded.

10. Partial Reachability Validations

Same mechanism as described in "Full Reachability Validations" section will be applied with exception of following differences on initiator.

- o When initiator wishes to perform a partial reachability validation towards identifier X upto identifier Y, number of hops to identifier Y is calculated.

- o TTL value based on this calculation is used as the IP TTL or MPLS TTL on top most label, and "your discriminator" of transmitted BFD control packet will carry BFD discriminator corresponding to target transit identifier Y.

- o Imposed label stack or IP destination address will continue to be of identifier X.

11. Scaling Aspect

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

If traditional BFD technology was used on a network comprised of N nodes, and each node monitored M unidirectional paths/LSPs, then total number of BFD sessions in such network will be:

(((N - 1) x M) x 2)

Assuming that each network node creates one reflector BFD session to handle all local BFD target identifiers, then total number of BFD sessions in same scenario will be:

(((N - 1) x M) + N)

12. Co-existence with Traditional BFD

This mechanism has no issues being deployed with traditional BFDs ([RFC5881]/[RFC5883]/[RFC5884]/[RFC5885]) because BFD discriminators which allow this mechanism to function are explicitly reserved and separate UDP port values are used with S-BFD.

13. BFD Echo

BFD echo is outside the scope of this document.

14. 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 Implementations MUST provide filtering capability based on source IP addresses or source node segment IDs of received BFD control packets: [RFC2827].

- o Implementations MUST NOT act on received BFD control packets containing Martian addresses as source IP addresses.

- o Implementations MUST ensure response target IP addresses or node segment IDs are reachable.

- o Initiator MAY pick crypto sequence number based on authentication mode configured.

- o The reflector MUST NOT look at the crypto sequence number before accepting the packet.

- o Reflector MAY look at the Key ID [I-D.ietf-bfd-generic-crypto-auth] in the incoming packet and verify the authentication data.

- o Reflector MUST accept the packet if authentication is successful.

- Reflector 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.
- Initiator MUST accept the S-BFD packet if it either comes with the same sequence number as it had sent or its within the window that it finds acceptable (described in detail in [I-D.ietf-bfd-generic-crypto-auth])

Using the above method,

- Reflectors continue to remain stateless despite using security.
- Reflectors are not susceptible to replay attacks as they always respond to S-BFD packets irrespective of the sequence number carried.
- An attacker cannot impersonate the Reflector since the Initiator will only accept S-BFD packets that come with the sequence number that it had originally used when sending the S-BFD packet.

15. IANA Considerations

BFD Target Identifier types:

Value	BFD Target Identifier Type
0	Reserved
1	Network Target Discriminator

New UDP port number, TBD1, will be requested for S-BFD.

16. Acknowledgements

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

17. 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

18. References

18.1. Normative References

[I-D.ietf-bfd-seamless-use-case]
Aldrin, S., Bhatia, M., Mirsky, G., Kumar, N., and S. Matsushima, "Seamless Bidirectional Forwarding Detection (BFD) Use Case", draft-ietf-bfd-seamless-use-case-00 (work in progress), June 2014.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

[RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881, June 2010.

[RFC5883] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for Multihop Paths", RFC 5883, June 2010.

[RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, June 2010.

18.2. Informative References

[I-D.ietf-bfd-generic-crypto-auth]
Bhatia, M., Manral, V., Zhang, D., and M. Jethanandani, "BFD Generic Cryptographic Authentication", draft-ietf-bfd-generic-crypto-auth-06 (work in progress), April 2014.

[I-D.previdi-filsfils-isis-segment-routing]
Previdi, S., Filsfils, C., Bashandy, A., Horneffer, M., Decraene, B., Litkowski, S., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., and J. Tantsura, "Segment Routing with IS-IS Routing Protocol", draft-previdi-filsfils-isis-segment-routing-02 (work in progress), March 2013.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[RFC5885] Nadeau, T. and C. Pignataro, "Bidirectional Forwarding Detection (BFD) for the Pseudowire Virtual Circuit Connectivity Verification (VCCV)", RFC 5885, June 2010.

Appendix A. Loop Problem

Consider a scenario where we have two nodes and both are S-BFD capable.

```
     Node A (IP 1.1.1.1) ---------------- Node B (IP 2.2.2.2)
                                |
                                |
                        Man in the Middle (MiM)
```

Assume node A reserved a discriminator 0x01010101 for target identifier 1.1.1.1 and has a reflector session in listening mode. Similarly node B reserved a discriminator 0x02020202 for its target identifier 2.2.2.2 and also has a reflector session in listening mode.

Suppose MiM sends a spoofed packet with MyDisc = 0x01010101, YourDisc = 0x02020202, source IP as 1.1.1.1 and dest IP as 2.2.2.2. 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 MyDdisc=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
Alcatel-Lucent

Email: manav.bhatia@alcatel-lucent.com

Santosh
Juniper Networks

Email: santoshpk@juniper.net