Link State Routing
Internet-Draft
Intended status: Standards Track
Expires: May 4, 2020

K. Talaulikar
P. Psenak
Cisco Systems, Inc.
A. Fu
Bloomberg
M. Rajesh
Juniper Networks
November 1, 2019

OSPF Strict-Mode for BFD
draft-ketant-lsr-ospf-bfd-strict-mode-03

Abstract

This document specifies the extensions to OSPF that enables a router and its neighbor to signal their intention to use Bidirectional Forwarding Detection (BFD) for their adjacency using link-local advertisement between them. The signaling of this BFD enablement, allows the router to block and not allow the establishment of adjacency with its neighbor router until a BFD session is successfully established between them. The document describes this OSPF "strict-mode" of BFD establishment as a prerequisite to adjacency formation.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Table of Contents

1. Introduction

Bidirectional Forwarding Detection (BFD) [RFC5880] enables routers to monitor dataplane connectivity over links between them and to detect faults in the bidirectional path between them. This capability is leveraged by routing protocols like Open Shortest Path First (OSPFv2) [RFC2328] and OSPFv3 [RFC5340] to detect connectivity failures for their adjacencies and trigger the rerouting of traffic around this failure more quickly than their periodic hello messaging based detection mechanism.

The use of BFD for monitoring routing protocols adjacencies is described in [RFC5882]. When BFD monitoring is enabled for OSPF

adjacencies, the BFD session is bootstrapped based on the neighbor address information discovered by the exchange of OSPF hello messages. Faults in the bidirectional forwarding detected via BFD then result in the bringing down of the OSPF adjacency. Note that it is possible in some failure scenarios for the network to be in a state such that the OSPF adjacency is capable of coming up, but the BFD session cannot be established, and, more particularly, data cannot be forwarded. In certain other scenarios, a degraded or poor quality link may result in OSPF adjacency formation to succeed only to result in BFD session establishment not being successful or the BFD session going down frequently due to its faster detection mechanism.

To avoid such situations which result in routing churn in the network, it would be beneficial not to allow OSPF to establish a neighbor adjacency until the BFD session is successfully established and stabilized. However, this would preclude the OSPF operation in an environment in which not all OSPF routers support BFD and are enabled for BFD monitoring. A solution would be to block the establishment of OSPF adjacencies if both systems are willing to establish a BFD session but a BFD session cannot be established. Such a mode of BFD use by OSPF is referred to as "strict-mode" wherein BFD session establishment becomes a prerequisite for OSPF adjacency coming up.

This document specifies the OSPF protocol extensions using link-local signaling (LLS) [RFC5613] for a router to indicate to its neighbor the willingness to establish a BFD session in the "strict-mode". It also introduces an extension for OSPFv3 link-local signaling of interface IPv4 address when used for IPv4 address-family (AF) instance to enable discovery of the IPv4 addresses for BFD session setup.

A similar functionality for IS-IS is specified [RFC6213].

2. LLS B-bit Flag

A new B-bit is defined in the LLS Type 1 Extended Options and Flags field. This bit is defined for the LLS block included in Hello packets and indicates that BFD is enabled on the link and that the router supports BFD strict-mode. Section 7 describes the position of this new B-bit.

A router MUST include the LLS block with the LLS Type 1 Extended Options and Flags TLV with the B-bit set its Hello messages when BFD is enabled on the link.

3. Local Interface IPv4 Address TLV

The Local Interface IPv4 Address TLV is a new LLS TLV meant for OSPFv3 protocol operations for IPv4 AF instances [RFC5838]. It has following 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Local Interface IPv4 Address                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBD, suggested value 21

Length: 4 octet

Local Interface IPv4 Address: The primary IPv4 address of the local interface.

4. Procedures

A router supporting BFD strict-mode advertises this capability through its hello messages as described in Section 2 above. When a router supporting BFD strict-mode, detects a new neighbor router that also supports BFD strict-mode, then it proceeds to establish adjacency with that neighbor as described further in this section.

This document updates the OSPF neighbor state machine as described in [RFC2328] specifically the operations related to the Init state as below when BFD strict-mode is used:

Init (without BFD strict-mode)

In this state, an Hello packet has recently been seen from the neighbor. However, bidirectional communication has not yet been established with the neighbor (i.e., the router itself did not appear in the neighbor's Hello packet). All neighbors in this state (or higher) are listed in the Hello packets sent from the associated interface.

Init (with BFD strict-mode)

In this state, an Hello packet has recently been seen from the neighbor. However, bidirectional communication has not yet been

established with the neighbor (i.e., the router itself did not appear in the neighbor's Hello packet). A BFD session establishment to the neighbor is requested, if not already done (e.g. in the event of transition from 2-way state). All neighbors in higher than Init state and those in Init state with BFD session up are listed in the Hello packets sent from the associated interface.

Whenever the neighbor state transitions to Down state, the removal of the BFD session associated with that neighbor SHOULD be requested by OSPF and the session re-setup SHOULD similarly be requested by OSPF after transitioning into Init state. This may result in the deletion and creation of BFD session respectively when OSPF is the only client interested in BFD session to the neighbor address.

An implementation MUST NOT wait for BFD session establishment in Init state unless BFD strict-mode is enabled on the router and the specific neighbor indicates BFD strict-mode capability via its Hello messages. When BFD is enabled, but the strict-mode of operation cannot be used, then an implementation SHOULD start the BFD session establishment only in 2-Way or higher state. This makes it possible for router to operate a mix of BFD operation in strict-mode or normal mode across different interfaces or even different neighbors on the same multi-access LAN interface.

Once the OSPF state machine has moved beyond the Init state, any change in the B-bit advertised in subsequent Hello messages MUST NOT result in any trigger in either the OSPF adjacency or the BFD session management (i.e. the B-bit is considered only when in the Init state). The disabling of BFD (or BFD strict-mode) on a router would result in its not setting the B-bit in its subsequent Hello messages. The disabling of BFD strict-mode has no change on the BFD operations and would not result in bringing down of any established BFD session. The disabling of BFD would result in the BFD session brought down due to Admin reason and hence would not bring down the OSPF adjacency.

When BFD is enabled on an interface over which we already have an existing OSPF adjacency, it would result in the router setting the B-bit in its subsequent Hello messages. If the adjacency is already up (i.e. in its terminal state of Full or 2-way with non-DR routers on a LAN) with a neighbor that also support BFD strict-mode, then an implemantion SHOULD NOT bring this adjacency down and instead use the BFD strict-mode of operations after the next transition into Init state. However, if the adjacency is not up, then an implementation MAY bring such an adjacency down so it can use the BFD strict-mode for its bring up.

4.1. OSPFv3 IPv4 Address-Family Specifics

The multiple AF support in OSPFv3 [RFC5838] requires the use of IPv6 link-local address as source address for hello packets even when forming adjacencies for IPv4 AF instances. In most deployments of OSPFv3 IPv4 AF, it is required that BFD be used to monitor and verify the IPv4 data plane connectivity between the routers on the link and hence the BFD session is setup using IPv4 neighbor addresses. The IPv4 neighbor address on the interface is learnt only later in the adjacency formation phase when the neighbor's Link-LSA is received. This results in the setup of the BFD session either after the adjacency is established or much later in the adjacency formation sequence.

To enable the BFD operations in strict-mode, it is necessary for a router to learn it's neighbor's IPv4 link address during the Init state of adjacency formation (ideally when it receives the first hello). The use of the Local Interface IPv4 Address TLV (as defined in Section 3) in the LLS block of the OSPFv3 Hello messages for IPv4 AF instances makes this possible. Implementations that support strict-mode of BFD operations for OSPFv3 IPv4 AF instances MUST include the Local Interface IPv4 Address TLV in the LLS block of their hello messages whenever the B-bit is set. A receiver MUST ignore the B-bit (i.e. not operate in BFD strict mode) unless the Local Interface IPv4 Address TLV is present in OSPFv3 Hello message for IPv4 AF instances.

4.2. Graceful Restart Considerations

An implementation needs to handle scenarios where both graceful restart (GR) and the strict-mode of BFD operations are deployed together. The GR aspects discussed in [RFC5882] also apply with strict-mode of operations. In addition to that, since the OSPF adjacency formation is held up until the BFD session establishment in the strict-mode of operation, the resultant delay in adajcency formation may affect or break the GR based recovery. In such cases, it is RECOMMENDED that the GR timers are setup such that they provide sufficient time to cover for normal BFD session establishment delays.

5. Operations & Management Considerations

An implementation SHOULD report the BFD session status along with the OSPF Init adjacency state when operating in BFD strict-mode and perform logging operations on state transitions to include the BFD events. This allows an operator to detect scenarios where an OSPF adjacency may be stuck waiting for BFD session establishment.

In network deployments with noisy links or those with packet loss, BFD sessions may flap frequently. In such scenarions, OSPF strict-mode for BFD may be deployed in conjunction with an BFD dampening or hold-down mechanism to help avoid frequent adjacency flaps due BFD causing routing churn.

6. Backward Compatibility

An implementation MUST support OSPF adjacency formation and operations with a neighbor router that does not advertise the BFD strict-mode capability - both when that neighbor router does not support BFD and when it does support BFD but not in the strict-mode of operation as described in this document. Implementations MAY provide an option to specifically enable BFD operations only in the strict-mode in which case, OSPF adjacency with a neighbor that does not support BFD strict-mode would not be established successfully. Implementations MAY provide an option to disable BFD strict-mode which results in the router not advertising the B-bit and BFD operations being performed in the same way as before this specification.

The signaling specified in this document happens at a link-local level between routers on that link. A router which does not support this specification would ignore the B-bit in the LLS block of hello messages from its neighbors and continue to bootstrap BFD sessions, if enabled, without holding back the OSPF adjacency formation. Since the router which does not support this specification would not have set the B-bit in the LLS block of its own hello messages, its neighbor routers that support this specification would not use BFD strict-mode with it. As a result, the behavior would be the same as before this specification. Therefore, there are no backward compatibility related issues or considerations that need to be taken care of when implementing this specification.

7. IANA Considerations

This specification updates Link Local Signaling TLV Identifiers registry.

Following values are requested for allocation:

o B-bit from "LLS Type 1 Extended Options and Flags" registry at bit position 0x00000010.

o TBD (Suggested value 21) - Local Interface IPv4 Address TLV

8. Security Considerations

The security considerations for "OSPF Link-Local Signaling" [RFC5613] also apply to the extension described in this document. Inappropriate use of the B-bit in the LLS block of an OSPF hello message could prevent an OSPF adjacency from forming or lead to failure to detect bidirectional forwarding failures. If authentication is being used in the OSPF routing domain [RFC5709][RFC7474], then the Cryptographic Authentication TLV [RFC5613] SHOULD also be used to protect the contents of the LLS block.

9. Acknowledgements

The authors would like to acknowledge the review and inputs from Acee Lindem, Manish Gupta, Balaji Ganesh and Rajesh M.

The authors would like to acknowledge Dylan van Oudheusden for highlighting the problems in using strict-mode for BFD session for IPv4 AF instance with OSPFv3 and Baalajee S for his suggestions on the approach to address it.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <https://www.rfc-editor.org/info/rfc2328>.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <https://www.rfc-editor.org/info/rfc5340>.

[RFC5613] Zinin, A., Roy, A., Nguyen, L., Friedman, B., and D. Yeung, "OSPF Link-Local Signaling", RFC 5613, DOI 10.17487/RFC5613, August 2009, <https://www.rfc-editor.org/info/rfc5613>.

[RFC5838] Lindem, A., Ed., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC 5838, DOI 10.17487/RFC5838, April 2010, <https://www.rfc-editor.org/info/rfc5838>.

[RFC5882] Katz, D. and D. Ward, "Generic Application of Bidirectional Forwarding Detection (BFD)", RFC 5882, DOI 10.17487/RFC5882, June 2010, <https://www.rfc-editor.org/info/rfc5882>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

10.2. Informative References

[RFC5709] Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M., Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic Authentication", RFC 5709, DOI 10.17487/RFC5709, October 2009, <https://www.rfc-editor.org/info/rfc5709>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/info/rfc5880>.

[RFC6213] Hopps, C. and L. Ginsberg, "IS-IS BFD-Enabled TLV", RFC 6213, DOI 10.17487/RFC6213, April 2011, <https://www.rfc-editor.org/info/rfc6213>.

[RFC7474] Bhatia, M., Hartman, S., Zhang, D., and A. Lindem, Ed., "Security Extension for OSPFv2 When Using Manual Key Management", RFC 7474, DOI 10.17487/RFC7474, April 2015, <https://www.rfc-editor.org/info/rfc7474>.

Authors' Addresses

Ketan Talaulikar
Cisco Systems, Inc.
India

Email: ketant@cisco.com

Peter Psenak
Cisco Systems, Inc.
Apollo Business Center
Mlynske nivy 43
Bratislava 821 09
Slovakia

Email: ppsenak@cisco.com

Albert Fu
Bloomberg
USA

Email: afu14@bloomberg.net

Rajesh M
Juniper Networks
India

Email: mrajesh@juniper.net