Network Working Group   L. Dunbar
Internet Draft   Huawei
                                                               N. So
                                                             Verizon
Intended status: Standard Track
Expires: July 2010

January 19, 2010

Path Condition Change Marking using BFD

draft-dunbar-bfd-path-condition-change-marking-00.txt

Status of this Memo

Abstract

There are times that an LSP source node needs to know change(s) has occurred along the originally established LSP path. This is especially true in the Mobile Backhaul environment where microwave transport is widely deployed. The bandwidth provided by the microwave transport can change with weather. The source LSR, e.g. LTE's eNodeB, needs to adjust its admission rate or shift more load to alternative paths when the backhaul transport path condition is changed.

This draft describes a simple mechanism for transit LSRs to notify Source LSR of the path condition changes.

Conventions used in this document

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 0.

Table of Contents

1. Motivation

   LSP's source node(s) can have multiple paths to the corresponding
   destination node(s). Those paths can have different performance
   behavior such as delay, delay variation, bandwidth, and so on. The
   LSP(s)' can carry traffic with various Class of Services (CoS). Some
   of the premium CoS require strict performance objectives to be met at
   all time, so it is desirable to have LSP's source node(s) to be
   notified if there is any condition change along the LSP path. That
   way the source node(s), which is aware of the performance objectives,
   can make the proper decision if an alternative path needs to be
   established, or the admission rate needs to be changed for the
   incoming traffic.

   A common example for the LSP condition change is the bandwidth
   fluctuation in Mobile Backhaul network, where Microwave transport is
   widely deployed. Most Microwave transport nodes adjust its bandwidth
   based on the weather. Even though there is RSVP-TE for individual
   links to advertise its available bandwidth to all the nodes in the
   routing domain, RSVP-TE might not be possible in some Mobile Backhaul
   environment where there might be multiple routing domains from base
   stations to MSO. If Source Nodes, i.e. LTE's eNodeB, are aware of the
   bandwidth change, they can adjust services accepted to the network,
   request other base stations to accept new calls, or trigger (or
   increase the frequency of) Performance Monitoring scheme.

   In other applications, some source LSRs want to get notified when
   there is congestion condition or port changes on the transit nodes,
   so that proper actions can be taken. MPLS-ECN (RFC 5129) specifies a
   mechanism for transit nodes to mark EXP bits when congestion happens.
   However, many deployed MPLS networks already use EXP bits to mark
   priority, making it not possible to use MPLS-ECN (RFC 5129) mechanism
   for the purpose of LSP change notification.

2. Analysis of potential methods

   There is MPLS-ECN (RFC 5129) marking on the MPLS header's EXP field
   when transit nodes encounter congestion. The problem with MPLS-ECN
   (RFC 5129) is that many deployed MPLS networks already use EXP bits
   to represent priority. Another issue with MPLS-ECN (RFC 5129) is that
   the congestion marking doesn't occur until congestion happens. When a
   transit link bandwidth is reduced, such as microwave transport link
   bandwidth reduction due to weather, queues on the transit node can

quickly build up. Even if MPLS-ECN (RFC 5129) scheme is used, the
queue on the transit node may already been overrun by the time the
egress node recognizing the congestion and notifies the source node.
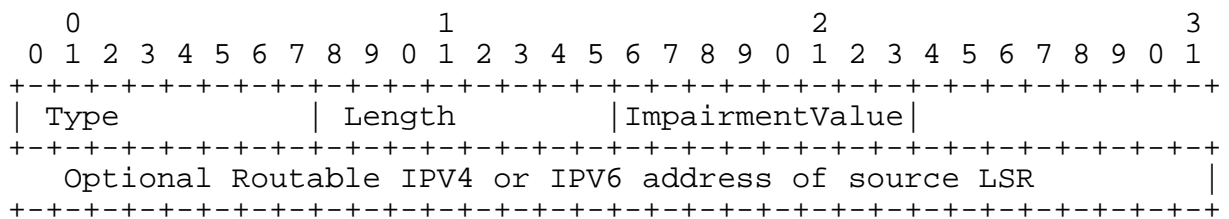
When there is a hard event change on a transit LSR, such as bandwidth
reduction or port change, an alternative approach is for the transit
LSR to send a simple notification to the source node indicating the
change occurred. However, the transit nodes may not know the source
nodes of the LSPs (e.g. LDP LSPs). Although the transit LSR can get
the source LSR's IP address from the MPLS-Ping Request message, the
MPLS-Ping may not be triggered when there is condition change on the
LSP path.

Another alternative is to mark on the MPLS header to indicate the
hard impairment event occurred on the path. It is the similar
approach as the MPLS-ECN (RFC5129). However, there are no available
bits on the MPSL header for such marking.

3. BFD protocol extension for path condition impairment notification

   When periodic BFD is enabled between a pair of LSRs (the Source and
   the destination), the BFD frequency is based on a fixed interval,
   usually in the magnitude of milliseconds. Since MPLS-BFD is intended
   to traverse along the LSP path, a transit node has the information on
   the port to the downstream LSR and is aware of the downstream link
   status, including impairment status, such as bandwidth being reduced,
   port being altered, or congested. Therefore, BFD is a good choice for
   path condition impairment notification when BFD is enabled on the
   LSP.

   This draft suggests adding an optional Impairment section to the BFD
   Control Frame:

```
         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            |ImpairmentValue|
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
           Optional Routable IPV4 or IPV6 address of source LSR        |
        +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   If a source LSR of a LSP cannot do anything when the LSP path is
   impaired, then the source LSR SHOULD NOT include this optional
   section in the BFD control frame. When the Impairment section is not
   present in the BFD control frame, the transit LSR does not need to
   mark the path condition change indication.

If the Source LSR can do something when the normal LSP path is impaired or altered, the Source LSR CAN include the optional Impairment section in the BFD Control Frame. If the Impairment section is attached, the transit LSR CAN do the following when experiencing downstream link impairment:

- The transit LSR SHOULD mark on the Impairment field of the condition of downstream link, and/or

- The transit LSR SHALL send a Path Condition Impairment Notification (PCIN) back to the source LSR if it knows the source LSR.

The Impairment Value field can take one of the following values:

Value	Meaning
1	port towards downstream LSR is congested
2	Bandwidth of the link towards downstream LSR is reduced
3	Port towards downstream LSR has been altered

The optional Routable IP address in the Impairment Section is for transit LSR to send the Path Condition Impairment Notification back to the source LSR. Since BFD control frames between a pair of LSRs are exchanged frequently, it is not necessary for the transit node to send the Path Condition Impairment Notification every time there is BFD traversed. In order to minimize work required on transit node, source node takes the responsibility to indicate if it needs a notification from transit node when the transit node experiences downstream link impairment. When the Routable IP address is included in the Impairment Section of BFD control frame, transit node SHALL send back a Path Condition Impairment Notification to the Source LSR when impairment is encountered.

The source LSR CAN perform the following actions upon receiving the Path Condition Impairment Notification from transit LSR.

- send more sophisticated inquiry messages to the transit LSR to diagnose what happened

- trigger performance monitoring scheme to measure the quality of the path

- re-adjust load balancing among the multiple paths from Source LSR to the Destination LSR

- Re-signal LSP to alternative path

- activate the secondary/protection path

- reduce admission rate (in the case of LTE eNodeB).

When a transit node receives a BFD and its downstream link is impaired or altered, it CAN perform the following actions:

1. If the BFD doesn't have the Impairment section, do nothing. Simply forward the BFD to the next hop.

   Otherwise:

2. The transit node SHALL set the ImpairmentValue field accordingly and then forward the BFD to the next hop.

3. If Routable IP address is included in the Impairment section, the transit LSR SHALL construct the Path Condition Impairment Notification message and send it to the Source LSR.

4. Path Condition Impairment Notification

   Though a new message format can be specified for transit node to send the Path Condition Impairment Notification, it is always easier for LSR to use an existing message type, like LSP-Ping Echo Reply [RFC4379], with some minor modification to do the job.
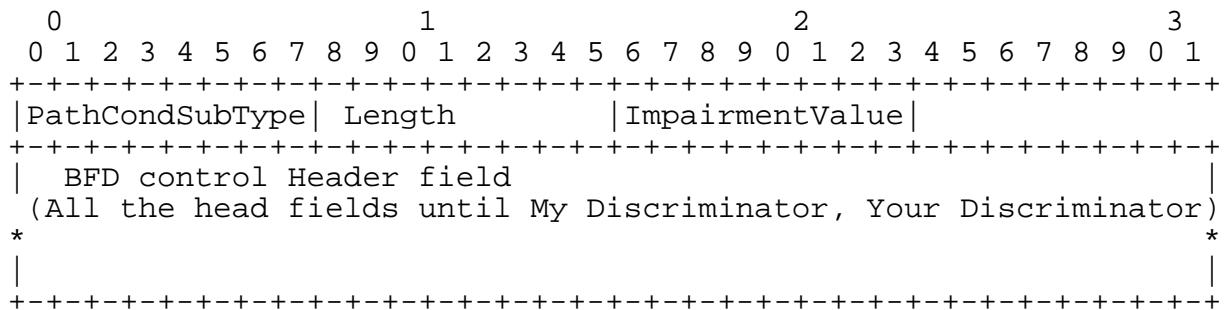
   This draft suggests a message type similar to the MPLS-Ping Echo reply which is specified in Section 3 of RFC 4379 to indicate the Path Condition Impairment Notification with the following changes:

   RFC 4379 specifies that the Message Type of the LSP-Ping has one of the two values. This draft suggests adding a new value to indicate that the message is for Path Condition Impairment Notification in responding to BFD:

```
   Value      Meaning
  --------    --------
     1        MPLS echo request [RFC 4379]
     2        MPLS echo reply [RFC 4379]
     3        Path Condition Impairment Notification in responding to
              BFD
```

[MPLS-Ping-Enhanced] introduced Downstream Detailed Mapping TLV. This draft suggests adding a new "Downstream Path Condition[1]" sub-TLV to reflect the condition of the downstream link. When used alone, path condition impairment notification SHALL be activated upon receiving a BFD control Frame with the optional Routable IP Address included in the Impairment Section. In this case, the Downstream Path Condition sub-TLV SHALL be the only sub-TLV in the Downstream Detailed Map [MPLS-Ping-Enhanced]. When the downstream path condition is included as a sub-type, the Return Code of the echo response message SHALL be set to "See DDM TLV for Return Code and Return SubCode".

Downstream Path Condition Impairment sub-TLV SHOULD have the following field:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |PathCondSubType| Length        |ImpairmentValue|
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   BFD control Header field                                   |
    (All the head fields until My Discriminator, Your Discriminator)
   *                                                             *
   |                                                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The ImpairmentValue is same as the value used in the BFD control (section 3 above).

5. Receiving Path Condition Impairment Notification

   It is out of the scope of this draft on what Source LSR will do upon receiving the Path Condition Impairment Notification.

6. Manageability Considerations

   This document does not add additional manageability considerations.

---

[1] The Downstream Path Condition sub-TLV can also be included in the normal LSP-Ping response to indicate that the downstream link, even though its connectivity is up, is impaired. This will be a separate draft amending the MPLS-Ping-Enhanced scheme.

7. Security Considerations

   This document has no requirement for a change to the security models
   within BFD.

8. IANA Considerations

   A future revision of this document will present requests to IANA for
   codepoint allocation.

9. Acknowledgments

   This document was prepared using 2-Word-v2.0.template.dot.

10. References

10.1. Normative References

   [ECN]   B. Davie, et. al., "Explicit Congestion Marking in MPLS",
              RFC 5129, Jan. 2008.

   [MPLS-PING] K. Kompella, et. al., "Detecting MPLS Data Plane
              Failures", RFC 4379, February 2006.

   [BFD-MPLS] Aggarwal, R., "draft-ietf-bfd-mpls-07.txt", work in
              progress.

   [BFD]   Katz, D. and Ward, D., "Bidirectional Forwarding Detection",
              draft-ietf-bfd-base-09.txt

   [LSP-Ping-Enhanced] N. Bahadur, et. al.,"Mechanism for performing
              LSP-Ping over MPLS tunnels", "draft-ietf-mpls-lsp-ping-
              enhanced-dsmap-04", work in progress.

10.2. Informative References

   [IEEE802.1au] IEEE802.1au Virtual Bridged Local Area Networks—
              Amendment: Congestion Notification

Authors' Addresses

    Linda Dunbar (Ed.)
    Huawei Technologies
    1700 Alma Drive, Suite 500
    Plano, TX 75075, USA
    Phone: (972) 543 5849
    Email: ldunbar@huawei.com

    Ning So (Ed.)
    Enterprise Data Network and Traffic Planning
    2400 N. Glenville Dr.
    Richardson, TX 75082, USA
    Phone: (972) 729 7905
    Email: ning.so@verizonbusiness.com

Intellectual Property Statement

    The IETF Trust takes no position regarding the validity or scope of
    any Intellectual Property Rights or other rights that might be
    claimed to pertain to the implementation or use of the technology
    described in any IETF Document or the extent to which any license
    under such rights might or might not be available; nor does it
    represent that it has made any independent effort to identify any
    such rights.

    Copies of Intellectual Property disclosures made to the IETF
    Secretariat and any assurances of licenses to be made available, or
    the result of an attempt made to obtain a general license or
    permission for the use of such proprietary rights by implementers or
    users of this specification can be obtained from the IETF on-line IPR
    repository at http://www.ietf.org/ipr

    The IETF invites any interested party to bring to its attention any
    copyrights, patents or patent applications, or other proprietary
    rights that may cover technology that may be required to implement
    any standard or specification contained in an IETF Document. Please
    address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

Acknowledgment