IS-IS and OSPF Extension for Event Notification
draft-ppsenak-lsr-igp-event-notification-01

Abstract

   Link-state protocols like IS-IS and OSPF have been designed to
   distribute state information - state of the local adjacencies, state
   of the local prefix reachability, etc. Each state can have
   additional attributes associated with it, but all the attributes are
   only meaningful while the state exists.

   This document extends link-state IGPs to distribute event
   notifications. An event notification has a very limited lifetime.
   It is rapidly propagated across the network and leaves no state after
   its lifetime.

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

   Link-state IGP protocols like IS-IS and OSPF are primarily used to
   distribute routing information between routers belonging to a single
   Autonomous System (AS) and to calculate the reachability for IPv4 or
   IPv6 prefixes advertised by the individual nodes inside the AS. Each
   node advertises the state of its local adjacencies, connected
   prefixes, capabilities, etc. The collection of these states from all
   the routers inside the area form a link-state database (LSDB) that
   describes the topology of the area and holds additional state
   information about the prefixes, router capabilities, etc.

   Link-state protocols have been designed to distribute state
   information. More precisely, it's only the existent or steady state
   that is being advertised - local adjacencies in UP sate, local
   prefixes that are reachable, etc. When the state does not exist
   anymore (e.g., adjacency transition to DOWN state), it is simply
   removed by the advertising node. Each state can have additional
   attributes associated with it, but all the attributes are only
   meaningful while the state exists.

   There are certain types of events, that do not represent a steady
   state and therefore cannot be advertised, that may be useful for the
   network operation.

   This document introduces the capability in link-state protocols to
   propagate event notifications that have a short and limited lifetime
   and do not introduce a state into IS-IS, OSPFv2, and OSPFv3. These
   event notifications are referred to as pulses to reflect their short-
   lived nature. Pulses may be used to advertise many types of events
   including those that are positive or negative in nature and for which
   there is no associated state that is to be maintained by link-state
   protocols.

2. Requirements for Pulse Notification

   This section describes the basic requirements of the pulse based
   notification for link-state protocols.

      Pulse Processing - processing of the pulse on the router is
      OPTIONAL. It's the decision of the receiver of the pulse whether
      the pulse is processed and any action is taken.

      Reliability - the distribution of the pulses MUST be reliable.

      Separation from Link-State - receiving pulses MUST NOT result in
      any update of the link-state topology or in any route calculation.

Pulse advertisements MUST NOT be sent using existing protocol
link-state messages.

Limited Lifetime - pulses are short-lived. There is no flushing
or purging mechanism for pulses. They MUST be destroyed after
their flooding procedure is complete.

Limited Retransmissions - pulses MUST be retransmitted, as
required by the flooding procedure, only for a limited period.

Not Part of Database Sync - pulses MUST NOT be exchanged as part
of the initial or post Graceful Restart database synchronization
between adjacent peers.

Relevance to routing protocol - use of pulses is restricted to
information known to the routing protocol as part of its normal
operation

3. IS-IS Pulse PDUs

Two new IS-IS PDUs are introduced for pulse propagation:

Flooding Scoped Pulse LSP (FSP-LSP)

Flooding Scoped Pulse PSNP (FSP-PSNP)

3.1. IS-IS Flooding Scoped Pulse LSP

The format of an IS-IS Flooding Scoped Pulse LSP (FSP-LSP) is similar
to the format of the Flooding Scoped LSP defined in [RFC7356], with
the following fields being removed:

"Reserved"

"Remaining Lifetime"

"Reserved|LSPDBOL|IS Type"

FSP-LSP supports all flooding scopes defined in [RFC7356].

An FS-Pulse-LSP has the following format:

```
                                      No. of octets
              +-------------------------+
              | Intradomain Routeing    |     1
              | Protocol Discriminator  |
              +-------------------------+
              | Length Indicator        |     1
              +-------------------------+
              | Version/Protocol ID     |     1
              | Extension               |
              +-------------------------+
              | ID Length               |     1
              +-------------------------+
              |R|R|R| PDU Type          |     1
              +-------------------------+
              |   Version               |     1
              +-------------------------+
              |P|  Scope                |     1
              +-------------------------+
              |   PDU Length            |     2
              +-------------------------+
              |   FSP-LSP ID            |    ID Length + 2
              +-------------------------+
              | Sequence Number         |     4
              +-------------------------+
              | Checksum                |     2
              +-------------------------+
              : Variable Length Fields  :    Variable
              +-------------------------+
```

PDU Type: 7 - (suggested - to be assigned by IANA)

All fields as defined in [RFC7356] for FS-LSPs

3.2. IS-IS Flooding Scoped Pulse PSNP

The format of an IS-IS Flooding Scoped Pulse PSNP (FSP-PSNP) is
similar to the format of the Flooding Scoped PSNP defined in
[RFC7356]

FSP-PSNP supports all flooding scopes defined in [RFC7356].

An FSP-PSNP has the following format:

```
                                         No. of octets
                   +-------------------------+
                   | Intradomain Routeing    |     1
                   | Protocol Discriminator  |
                   +-------------------------+
                   | Length Indicator        |     1
                   +-------------------------+
                   | Version/Protocol ID     |     1
                   | Extension               |
                   +-------------------------+
                   | ID Length               |     1
                   +-------------------------+
                   |R|R|R| PDU Type          |     1
                   +-------------------------+
                   |   Version               |     1
                   +-------------------------+
                   |   Reserved              |     1
                   +-------------------------+
                   |U|   Scope               |     1
                   +-------------------------+
                   |   PDU Length            |     2
                   +-------------------------+
                   |   Source ID             |   ID Length + 1
                   +-------------------------+
                   : Variable Length Fields  :    Variable
                   +-------------------------+
```

PDU Type: 8 (Suggested - to be assigned by IANA) defined in
[ISO10589].

All fields of the FSP-PSNP match the definition from Flooding
Scoped PSNP in [RFC7356].

Variable-length fields - list of TLVs.

This document defines a new TLV to be included in FSP-PSNPs: Flooding
Scope Pulse LSP Entries TLV (FSP-LSP Entries TLV) that has the
following format:

```
                                    No. of octets
                +------------------------+
                |  Type                  |    1
                +------------------------+
                |  Length                |    1
                +------------------------+
                |  FSP-LSP ID            |    ID Length + 2
                +------------------------+
                | Sequence Number        |    4
                +------------------------+
                | Checksum               |    2
                +------------------------+
                :                        :
                +------------------------+
                |  FSP-LSP ID            |    ID Length + 2
                +------------------------+
                | Sequence Number        |    4
                +------------------------+
                | Checksum               |    2
                +------------------------+
```

Type: 29 (Suggested - to be assigned by IANA)

Length: (ID Length + 8) * number of entries

FSP-LSP ID: The ID of the FSP-LSP being acknowledged

Sequence Number: Sequence number of the FSP-LSP being acknowledged

Checksum: Checksum reported in the FSP-LSP

3.3. Flooding Scope Update Process Operation

The Update Process in [ISO10589] is responsible for reliable flooding
of LSPs. In the case of FSP-LSPs, the lack of persistence introduces
some changes in how the Update Process operates.

Analagous to what is defined in [RFC7356], there is a separate
instance of the Update Process for each scope supported for FSP-LSPs.
The circuit(s) on which FSP-LSPs are flooded is limited to those
circuits that are participating in the given scope. Consistent
support of a given flooding scope on a circuit by all routers
operating on that circuit is required.

FSP-LSPs are not meant to be retained beyond the minimum time needed
to process the information and to provide a reasonable opportunity
for flooding the information to neighbors. FSP-LSPs are also not

synchronized on adjacency establishment and/or Graceful Restart
[RFC8706]. For this reason, an FSP Complete Sequence Number PDU is
NOT REQUIRED. Flooding of an FSP-LSP on a circuit ceases after a
configurable number of retries. Default number of retries is
RECOMMENDED to be 3.

Receipt of an FSP-PSNP with a matching Flooding Scope Pulse LSP Entry
serves as an acknowledgment of receipt of an FSP-LSP on that circuit.

FSP-LSPs SHOULD be retained in the FSP Scope Specific LSDB for
ZeroAgeLifetime (60 seconds). This is done to support reliable
flooding of the FSP-LSP and to minimize the possibility of
reprocessing a previously received FSP-LSP.

3.3.1. FSP-LSP Generation Procedures

Although sequence numbers in FSP-LSPs are less important than in
traditional LSPs since FSP-LSPs are not retained for a significant
period and are not purged, they are still useful to identity a newer
version of a given FSP-LSP. Nodes which originate FSP-LSPs MUST
remember the last sequence number used for a given FSP-LSP and
increment the sequence number when generating a new version.

FSP-LSP generation SHOULD utilize the "next" FSP-LSP ID each time new
pulse information needs to be advertised i.e., if the most recent
FSP-LSP ID used was A-00.n, the next set of pulse information SHOULD
be advertised using FSP-LSP.ID A-00.n+1. This minimizes the
possibility of confusion if other routers in the network have not yet
removed A-00.n from their LSPDB.

3.3.2. FSP-LSP Acknowledgement Behavior

Determining whether a received FSP-LSP is newer than a previously
received copy follows the rules defined for LSPs defined in
[ISO10589].

Received FSP-LSPs which are either newer or the same as an existing
entry in the LSPDB are acknowledged using FSP-PSNPs.

Received FSP-LSPs which are older than existing entries in the LSPDB
are ignored. The sender of the FSP-LSP will in any case cease
flooding such an FSP-LSP after a modest number of retries.

4. Use Case: Supporting BGP-PIC at scale

In this section we present one practical use case of event based
notification in link-state routing protocols.

The growth of networks running a link-state routing protocol results
in the addition of more state that presents itself in the form of
scalability and convergence challenges. The organization of networks
into levels/areas and IGP domains helps limit the scope of link-state
information within certain boundaries. However, the state related to
prefix reachability often requires propagation across a multi-area/
level and/or multi-domain IGP network.

Techniques such as summarization have been used traditionally to
address the scale challenges associated with advertising prefix state
outside of local area/domain.

However, this results in suppression of the individual prefix state
that is useful for triggering fast-convergence mechanisms outside of
the IGPs - e.g., BGP PIC Edge [I-D.ietf-rtgwg-bgp-pic].

In such a scenario, it is desirable to enable the notification of
events, such as an individual prefix becoming unreachable, outside of
the local area/domain and across the network in a manner that does
not leave behind any persistent state in the link-state database.

4.1. Use of Summarization

Deployment of large networks may utilize a significant number of
discrete IGP areas. Advertisement of inter-area prefixes is limited
to summaries to reduce the number of prefix advertisements which need
to be flooded domain-wide.

Considere a network consisting of 100 areas with 1K prefixes/area.
In the absence of summarization, there are 100 K prefixes which would
be advertised domain-wide. If in each area there are two Area Border
Routers (ABRs) - for redundancy purposes - each of which is
advertising 1K intra-area prefixes into other areas, there would then
be 100*2*1K = 200K prefix advertisements sent domain-wide.

If a single summary address is used to represent reachability for the
1K prefixes within an area, the number of prefix advertisements
flooded domain-wide becomes 100*2*1 = 200 summary prefix
adverytisements.

The use of summarization dramatically reduces the scale of network-
wide flooding, but it also means that a change in reachability to any
specific destination covered by a summary is not known to routers
outside a given area.

4.2. Use of Pulse in combination w Summarization

   Pulse can be used to signal loss of reachability to an individual
   destination covered by a summary. In the event that a single node
   becomes unreachable, this would result in the flooding of 2 pulses
   (one by each ABR in the impacted area. If we generalize this to loss
   of reachability to N nodes throughout the network, the total number
   of additional advertisements is 2N. The received pulses can then be
   used to trigger BGP-PIC fast convergence.

   The economy provided by the use of pulses in this use case diminishes
   linearly with the number of nodes which fail within a given time
   interval. In the event of a catastrophic network failure where many
   nodes fail within a given pulse interval, the number of pulses
   present in the network could begin to approach the number of
   individual prefixes present in the domain - which would effectively
   eliminate the scale benefits of the summary. Therefore, when using
   pulse for this use case, implementations SHOULD limit the number of
   pulses which are advertised in a given time interval.

4.3. IS-IS Summary Component Reachability Loss Pulse TLV

   IS-IS Summary Component Reachability Loss Pulse (SCRLP) TLV MAY be
   sent in an FSP-LSP. It is used by the IS-IS L1/L2 routers or by IS-
   IS Autonomous Boundary Routers (ASBR) that are performing prefix
   summarization at the area or domain boundary, to inform other nodes
   in the attached area(s) or domain(s) about the loss of the
   reachability to a previously reachable component of the summary-
   prefix inside the area or domain from which the summary-prefix is
   originated.

   The IS-IS SCRLP TLV has the 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    |      Flags    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R|R|R|      MT ID           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S|Summ-Pfx Len |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Summary Prefix (variable) ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Sub-TLV-len  |        Sub-TLVs (variable) . . .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S| Comp-Pfx Len|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Component Prefix (variable) ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Sub-TLV-len  |        Sub-TLVs (variable) . . .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S| Comp-Pfx Len|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Component Prefix (variable) ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Sub-TLV-len  |        Sub-TLVs (variable) . . .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              ...                              |
```

where:

   Type: 1

   Length: variable

   Flags: 1 octet. The following flags are defined:

```
       0
       0 1 2 3 4 5 6 7
      +-+-+-+-+-+-+-+-+
      |D|F|  Reserved |
      +-+-+-+-+-+-+-+-+
```

      D-flag: Same as described in section 4.1. of [RFC5305].

      F-flag: If unset, then the Summary Prefix and Component
      Prefix(es) are IPv4 prefixes. If set, then the Summary Prefix
      and Component Prefix(es) are IPv6 prefixes.

The remaining bits are reserved for future use. They MUST be
set to zero on transmission and MUST be ignored on receipt.

R bits: reserved for future use. They MUST be set to zero on
transmission and MUST be ignored on receipt.

MT ID: Multitopology Identifier as defined in [RFC5120]. Note
that the value 0 is legal.

Summ-Pfx Length + Flag: 1 octet

Summ-Pfx Length: Length of the Summary Prefix in bits. Valid
values are (0-31) when F-flag is unset, (0-127) when F-flag is
set.

S-bit: MUST be set when Sub-TLVs are present for Summary
Prefix, otherwise MUST NOT be set.

Summary Prefix: variable. IPv4 or IPv6 Summary Prefix.

Sub-TLV-length: 1 octet. Number of octets used by Summary Prefix
Sub-TLVs. Only present when S-bit is set.

Optional Sub-TLVs: No Sub-TLVs are defined by this document.

Comp-Pfx Length + Flag: 1 octet

Comp-Pfx Len.: 1 octet. Length of the Component Prefix in
bits. Valid values are (1-32) when F-flag is unset, (1-128)
when F-flag is set. Comp-Pfx Len MUST be > Summ-Pfx Length.

S-bit: MUST be set when Sub-TLVs are present for Component
Prefix, otherwise MUST NOT be set.

Component Prefix: variable. IPv4 or IPv6 Component Prefix.

Sub-TLV-length: 1 octet. Number of octets used by Component
Prefix sub-TLVs. Only present when S-bit is set.

Optional sub-TLVs: No Sub-TLVs are defined by this document.

When an IS-IS L1/L2 router or an IS-IS Autonomous Boundary Router
(ASBR) is performing prefix summarization and it loses the
reachability to one or more previously reachable component(s) of the
summary-prefix inside the area or domain for which the summarization
is done, it MAY originate the SCRLP TLV to inform routers in other
areas or domains about such summary component-prefix reachability
loss.

An originator of the SCRLP TLV chooses to advertise it in FSP-LSP
with L1 flooding scope and/or FSP-LSP with L2 flooding scope.

The IS-IS SCRLP TLV MAY be leaked between levels on L1/L2 router,
subject to local policy of such L1/L2 router.

IS-IS SCRLP TLV MUST NOT be leaked inside the area if the summary
prefix carried in IS-IS SCRLP TLV (Summary Prefix, Summ-Pfx Length)
is advertised from such area by L1/L2 router.

When the router receives the SCRLP TLV it MAY choose to inform the
BGP component on the router. BGP component on the router MAY trigger
BGP Prefix Independent Convergence (PIC) as specified in
[I-D.ietf-rtgwg-bgp-pic] as a result of such notification.

The mechanism on how the IS-IS passes the information from IS-IS
SCRLP TLV to the BGP component or how the BGP component uses this
information to trigger the PIC is implementation-specific and outside
of the scope of this specification.

The IS-IS SCRLP TLV may be used by other applications on the
receiving node that wish to be notified about the loss of summary
component-prefix reachability. The details of such usage are outside
of the scope of this specification.

5. Handling of the Control Plane Restart and ISSU

An egress PE may undergo a control-plane or protocol restart, or and
In-Service Software Upgrade. If these events are performed using
Nonstop Forwarding (NSF) as specified in [RFC3847] or Nonstop Routing
(NSR) procedures, the egress PE reachability inside its area is
preserved and no Pulse would be generated as a result of these
events.

If the IS-IS protocol restart, or route-processor fail-over on the
egress PE is performed using cold-restart procedures, the egress PE
reachability will be lost and Pulse will be generated by the ABRs
connected to the area. This is an expected behavior, as in case of
the cold-restart recovery the traffic is expected to be dropped if
forwarded to the egress PE and using an alternate BGP path is
desirable.

6. OSPF Pulse Notification

TBD.

7. OSPFv3 Pulse Notification

   TBD.

8. IANA Considerations

8.1. New IS-IS PDU Types

   This document includes the definition of two new PDU types that are
   reflected in the "IS-IS PDU Registry":

Value	Description
7	FSP-LSP
8	FSP-PSNP

8.2. Revised sub-TLV table

   IANA is requested to modify the table in "TLV Codepoints Registry" by
   adding columns for FSP-LSP and FSP-PSNP and set FSP-LSP:n and FSP-
   PSNP:n for all existing TLVs with the exception of 10
   (Authentication) and 11 (ESN).

8.3. IS-IS Flooding Scope Pulse LSP Entries TLV

   This document makes the following registrations in the IS-IS TLV
   Codepoints registry:

Type	Description	IIH	LSP	SNP	Purge	FSP-LSP	FSP-PSNP
29	FS-LSP Entries TLV	n	n	n	n	n	y

8.4. IS-IS Summary Component Reachability Loss Pulse TLV

   This document makes the following registrations in the IS-IS TLV
   Codepoints registry:

Type	Description	IIH	LSP	SNP	Purge	FSP-LSP	FSP-PSNP
30	Summary Component Reachability Loss Pulse	n	n	n	n	y	n

9. Security Considerations for ISIS

   The introduction of new PDU types introduces the possibility that an
   attacker could inject a false but apparently valid PDU. The use of

cryptographic authentication as defined in [RFC5304] or [RFC5310] minimizes the possibility of such occurrences.

Replay attacks could still be possible. Prevention of replay attacks can be done by including the Extended Sequence Numbers(ESN) TLV [RFC7602] in FSP-LSPs and FSP-PSNPs. Note, however, that the use of ESN MUST be done independently for each FSP-LSP ID. It is not safe to use a single ESN for the FSP-LSP PDU Type (as is done with hellos and SNPs) since we cannot guarantee the order in which multiple FSP-LSPs from the same source may arrive at a given node.

If a false PDU were to be injected, invalid SCRLP information could falsely trigger BGP-PIC behavior.

10. Security Considerations for OSPF

TBD.

11. Contributors

TBD

12. References

12.1. Normative References

[ISO10589]
          International Organization for Standardization, "Intermediate system to Intermediate system intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)", Nov 2002.

[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>.

[RFC3847] Shand, M. and L. Ginsberg, "Restart Signaling for Intermediate System to Intermediate System (IS-IS)", RFC 3847, DOI 10.17487/RFC3847, July 2004, <https://www.rfc-editor.org/info/rfc3847>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

   [RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic
              Authentication", RFC 5304, DOI 10.17487/RFC5304, October
              2008, <https://www.rfc-editor.org/info/rfc5304>.

   [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic
              Engineering", RFC 5305, DOI 10.17487/RFC5305, October
              2008, <https://www.rfc-editor.org/info/rfc5305>.

   [RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R.,
              and M. Fanto, "IS-IS Generic Cryptographic
              Authentication", RFC 5310, DOI 10.17487/RFC5310, February
              2009, <https://www.rfc-editor.org/info/rfc5310>.

   [RFC7356] Ginsberg, L., Previdi, S., and Y. Yang, "IS-IS Flooding
              Scope Link State PDUs (LSPs)", RFC 7356,
              DOI 10.17487/RFC7356, September 2014,
              <https://www.rfc-editor.org/info/rfc7356>.

   [RFC7602] Chunduri, U., Lu, W., Tian, A., and N. Shen, "IS-IS
              Extended Sequence Number TLV", RFC 7602,
              DOI 10.17487/RFC7602, July 2015,
              <https://www.rfc-editor.org/info/rfc7602>.

   [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>.

   [RFC8706] Ginsberg, L. and P. Wells, "Restart Signaling for IS-IS",
              RFC 8706, DOI 10.17487/RFC8706, February 2020,
              <https://www.rfc-editor.org/info/rfc8706>.

12.2. Informative References

   [I-D.ietf-rtgwg-bgp-pic]
              Bashandy, A., Filsfils, C., and P. Mohapatra, "BGP Prefix
              Independent Convergence", draft-ietf-rtgwg-bgp-pic-17
              (work in progress), October 2021.

Appendix A. BGP Pulse Handling

   Handling of the Pulse by the receiving application is out of scope of
   this document. This section provides some informational, high-level
   description on how a BGP on an ingress PE (Provider Edge) device may
   use the Pulse to trigger the BGP PIC (Prefix Independent
   Convergence). Note that PIC is a local behavior on ingress PE, which
   is implementation specific and nothing in this section mandates the
   implementation to follow what is described here to any degree.

Assuming the BGP multi-path destination prefix on an ingress PE, the
arrival of the Pulse, that indicates the loss of reachability of the
BGP next-hop for the primary path, can trigger the BGP PIC for such
prefix. This is similar in nature to what happens on ingress PE
without the use of summarization when the BGP next-hop for primary
path becomes unreachable.

In case the egress PE associated with the primary BGP path went down,
BGP on the ingress PE would eventually receive a withdrawal of such
path and would re-converge to the alternate path out of multi-paths.
Note that this happens independently of and after the BGP PIC was
triggered previously.

If the loss of reachability signaled by the Pulse is short-lived, it
is desirable that BGP reconverge to the state prior to receipt of the
Pulse. However, determination of the transient nature of the loss of
reachability depends on the absence of BGP updates which would be
expected following the loss of reachability to the egress PE. This
can be determined by triggering a timer on receipt of the Pulse. If
that timer expires without receipt of the expected BGP updates, then
BGP can reconverge to the pre-pulse state. The timer duration needs
to be long enough to allow for the expected BGP convergence to take
place in the case where the loss of reachability to the egress PE is
not transient.

Authors' Addresses

Peter Psenak (editor)
Cisco Systems
Pribinova Street 10
Bratislava 81109
Slovakia

Email: ppsenak@cisco.com

Les Ginsberg
Cisco Systems
821 Alder Drive
Milpitas, CA 95035
USA

Email: ginsberg@cisco.com

   Ketan Talaulikar
   Individual Contributor

   Email: ketant.ietf@gmail.com