IS-IS Flooding Speed advertisement
draft-decraene-lsr-isis-flooding-speed-01

Abstract

This document proposes a mechanism that can be used to increase the speed at which link state information is flooded across a network when multiple LSPDUs need to be flooded, such as in case of a node failure. It also reduces the likelihood of overloading the downstream flooding neighbors. This document defines a new TLV to be advertised in IS-IS Hello messages. This TLV carries two parameters indicating the performance capacity to receive LSPDUs: the minimum delay between two consecutive LSPDUs and the number of LSPDUs which can the received back to back.

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 RFC 2119 [RFC2119] RFC 8174 [RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

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

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

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

This Internet-Draft will expire on January 6, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction .............................................. 2
2. Flooding Speed TLV .................................... 4
3. Operation ................................................ 5
4. Interaction with other LSPDU rate limiting mechanisms ..... 6
5. Determining values to be advertised in the Flooding Speed TLV 6
6. IANA Considerations ....................................... 7
7. Security Considerations ................................... 7
8. Normative References ................................... 8
Appendix A. Changes / Author Notes ........................ 8
Authors’ Addresses .......................................... 8

1. Introduction

IGP flooding is paramount for Link State IGP as routing computations assume that the Link State DataBases (LSDBs) are always in sync across all nodes in the flooding domain.

Slow flooding directly translates to delayed network reaction to failure and LSDB inconsistencies across nodes. The former increases packets losses. The latter translates to routing inconsistencies and micro-loops leading to packets losses, link(s) overload, and jitter for all classes of services. Note that the link(s) affected by those forwarding issues may be any link in the network and not necessarily the links whose IGP status has changed.
IGP flooding is hard. One would want fast flooding when the network is stable and slow enough flooding to not overload the neighbor(s) when the network is less stable. Since flooding is performed hop by hop, not overloading the adjacent neighbors is sufficient. This document addresses these requirements by shaping LSPDUs with a minimal delay between two consecutive LSPDUs. This flooding behavior is already largely implemented by most implementations and hence doesn’t require changes in the basic flooding behavior. However, existing implementations rely on a shaping delay configured on the node sending the LSPDUs. But since the need is to not overload the downstream flooding node, the need is for the upstream flooding node to know the receiving speed of its downstream flooding neighbors. Although in theory this parameter could be configured on each upstream flooding node, given the knowledge of each of its neighbor in the topology and the receiving speed of those neighbors, in practice this configuration is difficult to maintain over time as the network topology change. In addition, as things currently stand, each network operator needs to evaluate the receiving capacity of each type of platform, depending on its hardware, software version and number of IS-IS adjacencies. Such platform performance is better known by its designer (the vendor) and even if validation tests are required, one single validation test by the vendor is more effective than N validations from N network providers. Finally, the reasoning behind the original choices of default value is not clear. Default values have largely remained unchanged over many years, despite very large increases in interface speeds and processing speed. This has resulted in default values that are likely to be very sub-optimal. For example, typical default values are one LSPDU per 33ms or 100ms, resulting in the ability to only send 30 or 10 LSPDUs per second. However, the same vendors recommend setting a BGP DDoS policer to 10,000 packets per second or more on the same control plane hardware, indicating that the control plane is capable of processing BGP packets at a rate of 10,000 packets per second.

This document proposes that the downstream flooding node advertises its LSPDU receiving speed for a single interface to the upstream flooding node in IS-IS hellos. This allows the sender to take into account the actual speed of the receiver. It also creates an incentive for vendors to improve this speed over time and to innovate to advertise a value reflecting the speed in the deployed environment (for example, by taking into account the number of IS-IS neighbors, which may send LSPDUs at the same time).

In addition, one single event in the network can require the flooding of multiple LSPDUs. The typical case is a node failure which requires the flooding of at least one LSPDU per neighbor of the failed node. Hence, if a node has N IGP neighbors, the failure of this node requires the advertisement and flooding of at least N
Internet-Draft     IS-IS Flooding Speed advertisement          July 2019

LSPDUs. The network won’t be able to converge to the new topology until all N LSPDUs are received by all nodes. Hence there is a need to be able to quickly flood N LSPDUs. This document addresses this requirement by allowing the fast flooding of some number of consecutive LSPDUs.

This document defines a new TLV for IS-IS hello that allows a given node to be able to advertise the rate at which the node can be expected to safely receive and process IS-IS LSPDUs from a given upstream flooding neighbor on a given interface. Each upstream flooding neighbor listens to the value advertised by the downstream flooding neighbor. This allows the fast flooding of LSPDUs while at the same time protecting the downstream flooding neighbor from receiving more LSPDUs than it can safely process in the event of network instability.

2. Flooding Speed TLV

This document defines a new TLV called "Flooding Speed TLV" to be included in IIH PDUs. All IIHs transmitted by a router that support this capability MUST include this TLV.

Type is TBD1.

Length is 4 octets.

Value field has two two-octets fields:

- minimumInterfaceLSPTransmissionInterval: the minimum interval, in milliseconds, between two LSPDUs consecutively sent by the upstream flooding node on a single interface;

- maximumInterfaceLSPTransmissionBurst: the maximum number of un-acknowledged LSPDUs that can be transmitted by the upstream flooding node on a single interface with a separation interval shorter than minimumInterfaceLSPTransmissionInterval (or even ‘back to back’).

```
+-----------------------------------------------------+
| minimumInterfaceLSPTransmissionInterval  (2 octets) |
+-----------------------------------------------------+
+-----------------------------------------------------+
| maximumInterfaceLSPTransmissionBurst     (2 octets) |
+-----------------------------------------------------+
```

Figure 1: Flooding Speed TLV
3. Operation

By sending the Flooding Speed TLV the node advertises to its IS-IS neighbor(s) its ability to receive, from the upstream flooding neighbor receiving this Flooding Speed TLV:

- minimumInterfaceLSPTransmissionInterval: the minimum interval, in milliseconds, between two LSPDUs consecutively sent by the upstream flooding node on a single interface;

- maximumInterfaceLSPTransmissionBurst: the maximum number of un-acknowledged LSPDUs that can be transmitted by the upstream flooding node on a single interface with a separation interval shorter than minimumInterfaceLSPTransmissionInterval (or even ‘back to back’).

The node sending the Flooding Speed TLV is the downstream flooding neighbor. It MUST be prepared to sustain, for a long duration, the reception of one LSPDU every minimumInterfaceLSPTransmissionInterval milliseconds. In addition, it MUST be capable of receiving maximumInterfaceLSPTransmissionBurst un-acknowledged LSPDUs with a shorter separation interval, provided than no more than 1000/minimumInterfaceLSPTransmissionInterval un-acknowledged LSPDUs are transmitted in any one second period.

Note that if the above two "MUST" cannot be fulfilled because of transient conditions, this cause no severe harm to the operation of IS-IS as this condition is already accounted for in [ISO10589]. As per [ISO10589], after a few seconds, respectively 2 and 10 by default in [ISO10589], neighbors will exchange PSNP (for point to point interface) or CSNP (for broadcast interface) and recover from the lost LSPDUs.

Note that, as per [ISO10589], the downstream flooding node dynamically acknowledges the received LSPDUs by sending CSNP or PSNP. By acknowledging the LSPDUs before the maximumInterfaceLSPTransmissionBurst is exhausted, the downstream flooding neighbor can achieve dynamic flow control and increase the flooding speed with its upstream flooding node without risking to overload any IS-IS router. If the maximumInterfaceLSPTransmissionBurst is large enough, on a point to point interface the downstream flooding node can acknowledge a set of multiple LSPDUs up to the maximum size of a PSNP (up to 90 LPDUs) which allows dynamic flow control without even increasing the number of PSNPs.

The node receiving the Flooding Speed TLV is the upstream flooding neighbor. The upstream flooding neighbor MUST NOT transmit LSPDUs at
a sustained rate greater than one LSPDU every
minimumInterfaceLSPTransmissionInterval milliseconds. The upstream
flooding neighbor MAY transmit maximumInterfaceLSPTransmissionBurst
un-acknowledged LSPDUs with a shorter separation interval, provided
than no more than 1000/minimumInterfaceLSPTransmissionInterval LSPDUs
are transmitted in any one second period.

4. Interaction with other LSPDU rate limiting mechanisms

[ISO10589] describes a mechanism that limits the rate at which LSPDUs
from the same source system are sent out all interfaces. (See the
description of the parameter minimumLSPTransmissionInterval in
sections 7.3.21 and 7.3.15.5 of [ISO10589] .) In practice, however,
router vendors have implemented mechanisms that limit the rate of
LSPDUs sent on a given interface. This is often configurable on a
per-interface basis using ‘lsp-interval’ or ‘lsp-pacing-interval’ CLI
configuration.) The mechanism described in the current document
extends the practice of limiting the rate of LSPDUs sent on a given
interface, by using parameters advertised by the downstream flooding
neighbor. When the mechanism described in the current document is
used, the mechanism described in section 7.3.15.5 of [ISO10589] is
not used.

5. Determining values to be advertised in the Flooding Speed TLV

The values that a downstream flooding neighbor advertises in the
Flooding Speed TLV should not change often. For example, in order to
compute the values in the Flooding Speed TLV, a reasonable choice
might be for a node to use a formula based on an off line tests of
the overall LSPDU processing speed for a particular set of hardware
and the number of interfaces configured for IS-IS. With such a
formula, the values advertised in the Flooding Speed TLV would only
change when additional IS-IS interfaces are configured. On the other
hand, it would be undesirable to use a formula that depends, for
example, on an active measurement of the CPU load to modify the
values advertised in the Flooding Speed TLV. This could introduce
feedback into the IGP flooding process that could produce unexpected
behavior. Since correct IGP flooding is so fundamental to network
operation, we do not want to introduce new dynamic behavior to it.
By requiring that the values advertised in the Flooding Speed TLV not
change very often, we expect to produce overall flooding behavior
similar to what might be achieved by manually configuring per-
interface LSPDU rate limiting on all interfaces in the network.
6. IANA Considerations

IANA is requested to allocate one TLV from the IS-IS TLV codepoint registry.

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>IIH</th>
<th>LSP</th>
<th>SNP</th>
<th>Purge</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD</td>
<td>Flooding Speed TLV</td>
<td>y</td>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
</tbody>
</table>

Figure 2

7. Security Considerations

Any new security issues raised by the procedures in this document depend upon the ability of an attacker to inject a false but apparently valid IIH, the ease/difficulty of which has not been altered.

As with others TLV advertisements, the use of a cryptographic authentication as defined in [RFC5304] or [RFC5310] allows the authentication of the peer and the integrity of the message. As this document defines a TLV for IS-IS Hello message (IIH), the relevant cryptographic authentication is for IS-IS Hello message (IIH).

In the absence of cryptographic authentication, as IS-IS does not run over IP but directly over the link layer, it’s considered difficult to inject false IIH without having access to the link layer.

If a false IIH is sent with a Flooding Speed TLV set to low values, the attacker can reduce the flooding speed between the two adjacent neighbors which can result in LSDB inconsistencies and transient forwarding loops. However, is not significantly different than filtering or altering LSPDUs which would also be possible with access to the link layer. In addition, if the downstream flooding neighbor has multiple IGP neighbors, which is typically the case for reliability or topological reasons, it would receive LSPDUs at a regular speed from its other neighbors and hence would maintain LSDB consistency.

If a false IIH is sent with a Flooding Speed TLV set to high values, the attacker can increase the flooding speed which can either overload a node or more likely generate loss of LSPDUs. However, is not significantly different than sending many LSPDUs which would also be possible with access to the link layer, even with cryptographic authentication enabled. In addition, IS-IS has procedures to detect the loss of LSPDUs and recover.
This TLV advertisement is not flooded across the network but only sent between two adjacent IS-IS neighbors. This would limit the consequences in case of forged messages, and also limits the dissemination of such information.

8. Normative References


Appendix A. Changes / Author Notes

[RFC Editor: Please remove this section before publication]

00: Initial version.

Authors’ Addresses

Bruno Decraene
Orange

Email: bruno.decrapere@orange.com
Chris Bowers  
Juniper Networks, Inc.  
1194 N. Mathilda Avenue  
Sunnyvale, CA  94089  
USA  

Email: cbowers@juniper.net

Jayesh J  
Juniper Networks, Inc.  
1194 N. Mathilda Avenue  
Sunnyvale, CA  94089  
USA  

Email: jayeshj@juniper.net

Tony Li  
Arista Networks  
5453 Great America Parkway  
Santa Clara, California  95054  
USA  

Email: tony.li@tony.li

Gunter Van de Velde  
Nokia  
Copernicuslaan 50  
Antwerp  2018  
Belgium  

Email: gunter.van_de_velde@nokia.com