Inter-Domain Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 14, 2021

Y. Wang
T. Zhou
M. Liu
Huawei
R. Pang
China Unicom
H. Chen
China Telecom

July 13, 2020

BGP-LS Extensions for In-situ Flow Information Telemetry (IFIT) Capability Advertisement
draft-wang-idr-bgpls-extensions-ifit-00

Abstract

This document extends Node and Link Attribute TLVs to Border Gateway Protocol-Link State (BGP-LS) to advertise supported In-situ Flow Information Telemetry (IFIT) capabilities at node and/or link granularity. An ingress router cannot insert IFIT-Data-Fields for packets going into a path unless an egress router has indicated via signaling that it has the capability to process IFIT-Data-Fields. In addition, such advertisements would be useful for entities (e.g. Path Computation Element (PCE)) to gather each router's IFIT capability for achieving the computation of end-to-end Traffic Engineering (TE) paths that be able to fulfill the visibility of on-path OAM data.

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

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."

This Internet-Draft will expire on January 14, 2021.

Table of Contents

1. Introduction

IFIT provides a high-level framework and a reflection-loop solution for on-path telemetry [I-D.song-opsawg-ifit-framework]. At present, there is a family of emerging on-path telemetry techniques, including In-situ OAM (IOAM) [I-D.ietf-ippm-ioam-data], IOAM Direct Export (DEX) [I-D.ietf-ippm-ioam-direct-export], Enhanced Alternate Marking (EAM) [I-D.ietf-6man-ipv6-alt-mark], etc.

IFIT is a solution focusing on network domains. The "network domain" consists of a set of network devices or entities within a single

Autonomous System (AS). The part of the network which employs IFIT is referred to as the IFIT domain. One network domain may consist of multiple IFIT domains. An IFIT domain may cross multiple network domains. The family of emerging on-path telemetry techniques may be partially enabled in different vendors' devices as an emerging feature for various use cases of application-aware network operations. So that in order to dynamically enable IFIT functionality in a network domain, it is necessary to advertise the information of IFIT option types supported in each device.

An ingress router cannot insert IFIT-Data-Fields for packets going into a path unless an egress router has indicated via signaling that it has the capability to process IFIT-Data-Fields. In addition, such advertisements would be useful for entities (e.g. Path Computation Element (PCE)) to gather each router's IFIT capability for achieving the computation of end-to-end TE paths that be able to fulfill the visibility of on-path OAM data.

[RFC7752] describes a mechanism by which link-state and TE information can be collected from the network outside one Interior Gateway Protocols (IGP) area or Autonomous System (AS). This document extends Node and Link Attribute TLVs to BGP-LS to advertise supported IFIT capabilities at node and/or link granularity.

2. Terminology

Following are abbreviations used in this document:

- BGP-LS: Border Gateway Protocol - Link State
- IFIT: In-situ Flow Information Telemetry
- TE: Traffic Engineering
- IOAM: In-situ OAM
- PBT: Postcard-Based Telemetry
- DEX: IOAM Direct Export
- EAM: Enhanced Alternate Marking
- IGP: Interior Gateway Protocols
- AS: Autonomous System
- E2E: Edge-to-Edge

o NLRI: Network Layer Reachability Information

3. IFIT Capability

Each IFIT-capable node is configured with a node-id which uniquely identifies a node within the associated IFIT domain. To accommodate the different use cases or requirements of in-situ flow information telemetry, IFIT data fields updated by network nodes fall into different categories which are referred as different IFIT option types, including IOAM Trace Option-Types [I-D.ietf-ippm-ioam-data], IOAM Edge-to-Edge (E2E) Option-Type [I-D.ietf-ippm-ioam-data], IOAM DEX Option-Type [I-D.ietf-ippm-ioam-direct-export] and EAM Option-Type [I-D.ietf-6man-ipv6-alt-mark]. A subset or all the IFIT-Option-Types and their corresponding IFIT-Data-Fields can be associated to an IFIT-Namespace. Namespace identifiers allow a device which is IFIT-capable to determine whether IFIT-Option-Types need to be processed. So that IFIT-Option-Types and Namespace-IDs SHOULD be included in IFIT capability information.

This document defines the IFIT Capability information formed of one or more pairs of a 2-octet Namespace-ID and 16-bit Option-Type Flag.

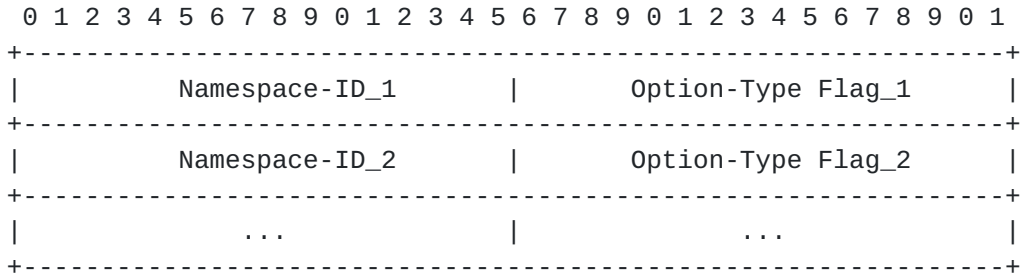
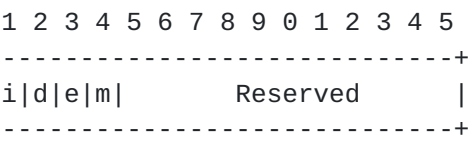


Fig. 1 IFIT Capability Format

Where:

o Namespace-ID: A 2-octet identifier, which must be present and populated in all IFIT-Option-Types. The definition is the same as described in [I-D.ietf-ippm-ioam-data].

o Option-Type Flag: A 16-bit bitmap, which is defined as:



Where:

- p-Flag: IOAM Pre-allocated Trace Option Type flag. When set, this indicates that the router is capable of IOAM Pre-allocated Trace [I-D.ietf-ippm-ioam-data].

- i-Flag: IOAM Incremental Trace Option Type flag. When set, this indicates that the router is capable of IOAM Incremental Tracing [I-D.ietf-ippm-ioam-data].

- d-Flag: IOAM DEX Option Type flag. When set, this indicates that the router is capable of IOAM DEX [I-D.ietf-ippm-ioam-direct-export].

- e-Flag: IOAM E2E Option Type flag. When set, this indicates that the router is capable of IOAM E2E processing [I-D.ietf-ippm-ioam-data].

- m-Flag: EAM flag. When set, this indicates that the router is capable of processing Enhanced Alternative Marking packets [I-D.ietf-6man-ipv6-alt-mark].

- Reserved: Must be set to zero upon transmission and ignored upon receipt.

An IFIT node MAY be capable of more than one IFIT option types. In this case, Option-Type Flag can has more than one bit being set.

In this document, Link IFIT Capability is defined as the supported IFIT-Option-Types of the interface associated with the link. When all interfaces associated with links support the same IFIT-Option-Type, the Node IFIT Capability SHOULD represent the Link IFIT Capability. Both of Node and Link IFIT Capability information are formed of one or more pairs of Namespace-ID and Option-Type Flag.

When both of Node and Link IFIT Capability are advertised, the Link IFIT Capability information MUST take precedence over the Node IFIT Capability. Besides, when a Link IFIT Capability is not signaled, then the Node IFIT Capability SHOULD be considered to be the IFIT Capability for this link.

4. Signaling IFIT Capability Using BGP-LS

4.1. Node IFIT TLV

The Node IFIT TLV is an optional extension to Node Attribute TLVs that may be encoded in the BGP-LS Attribute associated with a Node NLRI [RFC7752]. The following format is used:

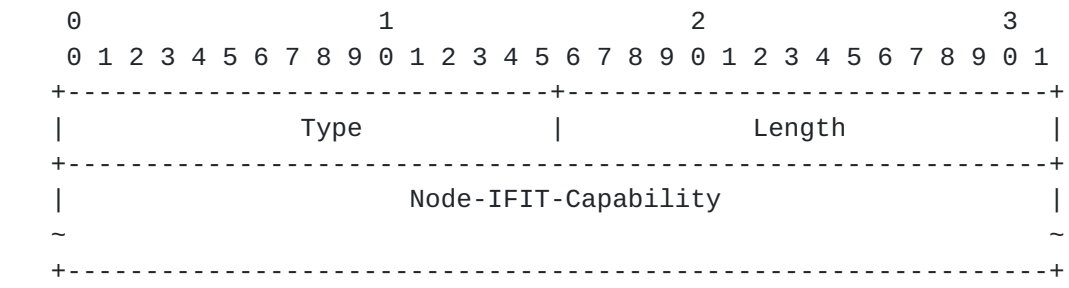


Fig. 7 BGP-LS Node IFIT TLV

Where:

- Type: To be assigned by IANA
- Length: A 2-octet field that indicates the length of the value.
- Node-IFIT-Capability: A multiple of 4-octet field, which is as defined in Section 3.

4.2. Link IFIT TLV

The Link IFIT TLV is an optional extension to Link Attribute TLVs that may be encoded in the BGP-LS Attribute associated with a Link NLRI [RFC7752]. The following format is used:

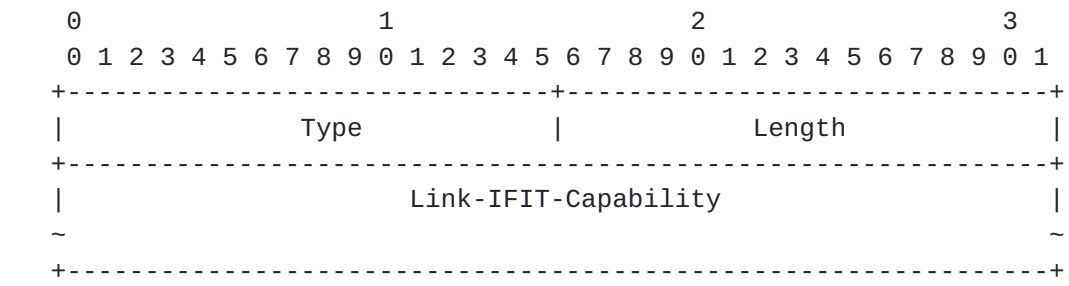


Fig. 8 BGP-LS Link IFIT TLV

Where:

- Type: To be assigned by IANA
- Length: A 2-octet field that indicates the length of the value.
- Link-IFIT-Capability: A multiple of 4-octet field, which is as defined in Section 3.

5. Application

As any packet with IFIT-Data-Fields must not leak out from the IFIT domain, the IFIT decapsulating node must be able to capture packets with IFIT-specific header and metadata and recover their format before forwarding them out of the IFIT domain. Thus, an ingress router cannot insert IFIT-Data-Fields for packets going into a path unless an egress router has indicated via signaling that it has the capability to process IFIT-Data-Fields. In this case, such advertisements are helpful for avoiding the leak of IFIT-specific header and metadata.

In addition, such advertisements would be useful for entities (e.g. Path Computation Element (PCE)) to gather each router's IFIT capability for achieving the computation of end-to-end TE paths that be able to fulfill the visibility of on-path OAM data. For example, for achieving the computation of low-latency SR-TE path, latency is expected to be collected at every node that a packet traverses to ensure performance visibility into the entire path. IOAM Trace Option-Types is a desired option to have a hop-by-hop latency measurement. If not all nodes on this path are IOAM Trace Option-Type capable, an incomplete measurement can have negative impacts on SR-TE path computation and adjustment for low-latency assurance.

6. IANA Considerations

IANA is requested to allocate values for the following TLV Type from the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry.

Code Point	Description	Reference
TBA1	Node IFIT	This document
TBA2	Link IFIT	This document

7. Security Considerations

This document introduces new BGP-LS Node and Link Attribute TLVs for the IFIT Capability advertisements at node and/or link granularity. It does not introduce any new security risks to BGP-LS.

8. Acknowledgements

The authors would like to thank Acee Lindem, Christian Hopps, Robert Raszuk, Les Ginsberg, Jeff Tantsura, Rakesh Gandhi and Greg Mirsky for the comments and advices.

9. References

9.1. 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>.

[RFC7752] "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", <https://datatracker.ietf.org/doc/rfc7752/>.

9.2. Informative References

[I-D.ietf-6man-ipv6-alt-mark] "IPv6 Application of the Alternate Marking Method", <https://datatracker.ietf.org/doc/draft-ietf-6man-ipv6-alt-mark/>.

[I-D.ietf-ippm-ioam-data] "Data Fields for In-situ OAM".

[I-D.ietf-ippm-ioam-direct-export] "In-situ OAM Direct Exporting", <https://datatracker.ietf.org/doc/draft-ietf-ippm-ioam-direct-export/>.

[I-D.song-opsawg-ifit-framework] "In-situ Flow Information Telemetry Framework", <https://datatracker.ietf.org/doc/draft-song-opsawg-ifit-framework/>.

Authors' Addresses

Yali Wang
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: wangyali11@huawei.com

Tianran Zhou
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: zhoutianran@huawei.com

Min Liu
Huawei
156 Beiqing Rd., Haidian District
Beijing
China

Email: lucy.liumin@huawei.com

Ran Pang
China Unicom
9 Shouti South Rd., Haidian District
Beijing
China

Email: pangran@chinaunicom.cn

Huanan Chen
China Telecom
109 West Zhongshan Ave.
Guangzhou, Guangdong
China

Email: chenhuan6@chinatelecom.cn