

Link State Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 29, 2021

Y. Wang
T. Zhou
Huawei
F. Qin
China Mobile
H. Chen
China Telecom
R. Pang
China Unicom
July 28, 2020

IGP Extensions for In-situ Flow Information Telemetry (IFIT) Capability
Advertisement
[draft-wang-lsr-igp-extensions-ifit-01](#)

Abstract

This document extends Node and Link Attribute TLVs to Interior Gateway Protocols (IGP) 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 ingress routers to gather each router's IFIT capability for achieving the computation of Traffic Engineering (TE) paths or loose 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 29, 2021.

Copyright Notice

Copyright (c) 2020 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.	Terminology	3
3.	IFIT Capability	4
4.	Signaling IFIT Capability Using IS-IS	6
4.1.	IS-IS Node IFIT Sub-TLV	6
4.2.	IS-IS Link IFIT Sub-TLV	6
5.	Signaling IFIT Capability Using OSPF	7
5.1.	OSPF Node IFIT TLV	7
5.2.	OSPFv2 Link IFIT sub-TLV	8
5.3.	OSPFv3 Link IFIT Sub-TLV	9
6.	Application	9
7.	IANA Considerations	10
8.	Security Considerations	10
9.	Acknowledgements	10
10.	References	10
10.1.	Normative References	10
10.2.	Informative References	11
	Authors' Addresses	12

[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 ingress routers to gather each router's IFIT capability for achieving the computation of TE paths or loose TE paths that be able to fulfill the visibility of on-path OAM data.

BGP-LS defines a way to advertise topology and associated attributes and capabilities of the nodes in that topology to a centralized controller [[RFC7752](#)]. Typically, BGP-LS is configured on a small number of nodes that do not necessarily act as head-ends. In order for BGP-LS to signal IFIT node capabilities for all the devices in the network, IFIT node capabilities SHOULD be advertised by every IGP router in the network.

This document defines a mechanism to signal the supported IFIT capabilities at node and/or link granularity using IS-IS, OSPFv2 and OSPFv3.

2. Terminology

Following are abbreviations used in this document:

- o BGP-LS: Border Gateway Protocol - Link State
- o IS-IS: Intermediate System to Intermediate System
- o OSPF: Open Shortest Path First
- o IFIT: In-situ Flow Information Telemetry
- o TE: Traffic Engineering

- o IOAM: In-situ OAM
- o PBT: Postcard-Based Telemetry
- o DEX: IOAM Direct Export
- o EAM: Enhanced Alternate Marking
- o IGP: Interior Gateway Protocols
- o AS: Autonomous System
- o 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.

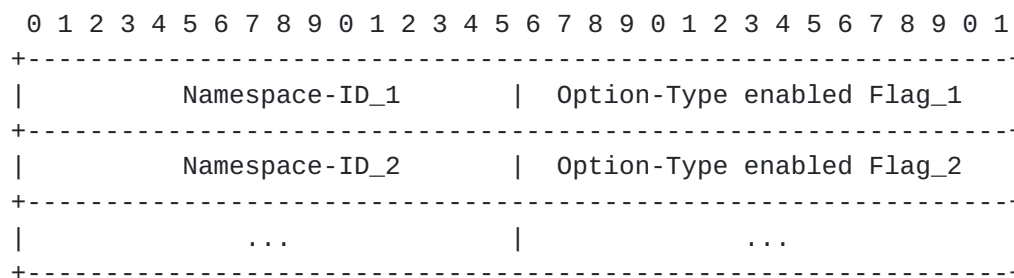


Fig. 1 IFIT Capability Format

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:

```
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-----+
|p|i|d|e|m|      Reserved      |
+-----+
```

Where:

- o 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](#)].
- o 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](#)].
- o 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](#)].
- o 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](#)].
- o 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](#)].
- o 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 IS-IS

The IS-IS Extensions for advertising Router Information TLV named IS-IS Router CAPABILITY TLV [[RFC7981](#)], which allows a router to announce its capabilities within an IS-IS level or the entire routing domain. And [[RFC5305](#)] describes extensions to IS-IS to support Traffic Engineering.

4.1. IS-IS Node IFIT Sub-TLV

According to the format of IS-IS Router CAPABILITY TLV [[RFC7981](#)], the Node IFIT sub-TLV within the body of the IS-IS router CAPABILITY TLV is composed of three fields, a one-octet Type field, a one-octet Length field, and a multiple of 4-octet Value field. The following format is used:

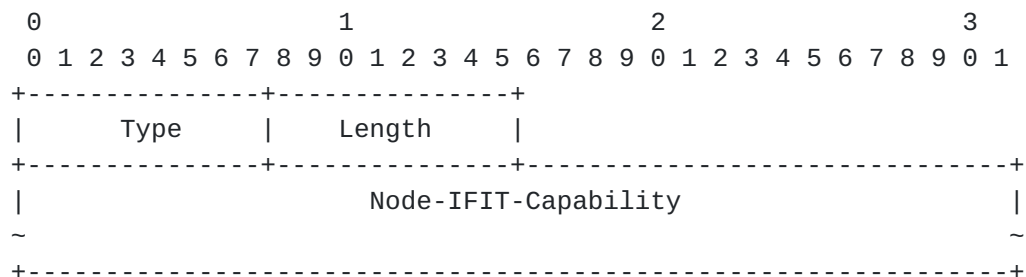


Fig. 2 IS-IS Node IFIT Sub-TLV Format

Where:

- o Type: To be assigned by IANA
- o Length: A one-octet field that indicates the length of the value portion in octets.
- o Node-IFIT-Capability: A multiple of 4-octet field, which is same as defined in [Section 3](#).

4.2. IS-IS Link IFIT Sub-TLV

The Link IFIT sub-TLV is defined to carry the IFIT-Capability information of the interface associated with the link, which is formed of three fields, a one-octet Type field, a one-octet Length field, and a multiple of 4-octet Value field. The following format is used:

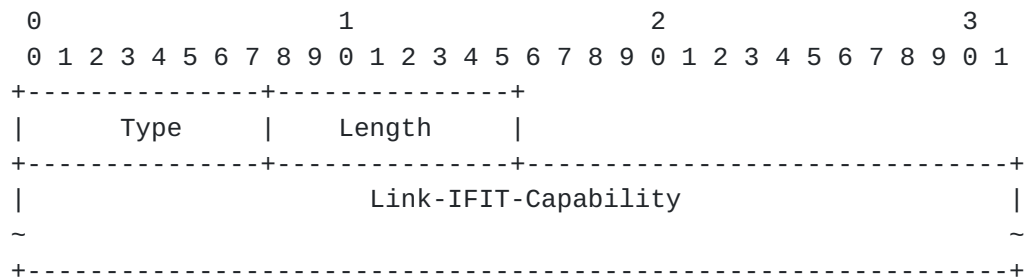


Fig. 3 IS-IS Link IFIT Sub-TLV Format

Where:

- o Type: To be assigned by IANA
- o Length: A one-octet field that indicates the length of the value portion in octets.
- o Link-IFIT-Capability: A multiple of 4-octet field, which is same as defined in [Section 3](#).

5. Signaling IFIT Capability Using OSPF

Given that OSPF uses the options field in LSAs and hello packets to advertise optional router capabilities [[RFC7770](#)], this document defines a new IFIT Node TLV within the body of the OSPF RI Opaque LSA [[RFC7770](#)] to carry the IFIT Capabilities of the router originating the RI LSA.

This document defines the Link IFIT sub-TLV to carry the IFIT-Capability information of the interface associated with the link. For OSPFv2, the link-level IFIT capability information is advertised as a sub-TLV of the OSPFv2 Extended Link TLV as defined in [[RFC7684](#)]. For OSPFv3, the link-level IFIT capability information is advertised as a sub-TLV of the E-Router-LSA TLV as defined in [[RFC8362](#)].

5.1. OSPF Node IFIT TLV

The Node IFIT TLV is composed of three fields, a 2-octet Type field, a 2-octet Length field, and a multiple of 4-octet Value field. The following format is used:

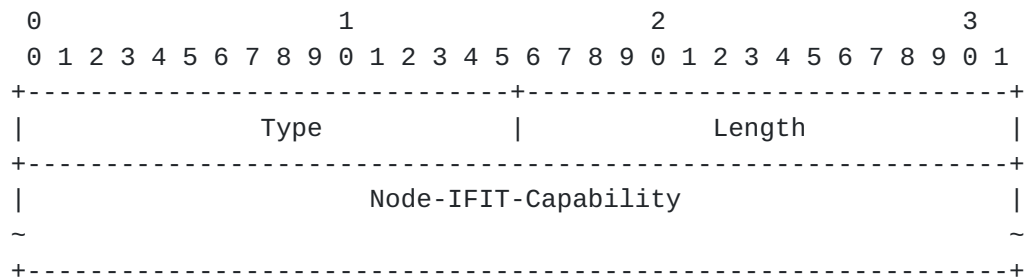


Fig. 4 OSPF Node IFIT TLV

Where:

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

5.2. OSPFv2 Link IFIT sub-TLV

The Link IFIT sub-TLV encoded in the OSPFv2 Extended Link TLV is constructed of three fields, a 2-octet Type field, a 2-octet Length field, and a multiple of 4-octet Value field. The following format is used:

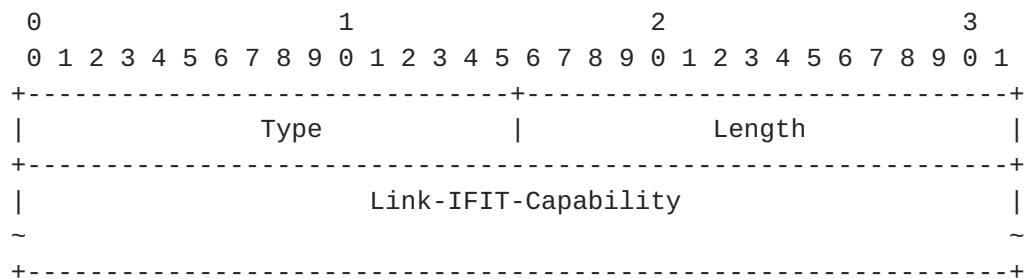


Fig. 5 OSPFv2 Link IFIT sub-TLV

Where:

- o Type: To be assigned by IANA
- o Length: A 2-octet field that indicates the length of the value field.

- o Link-IFIT-Capability: A multiple of 4-octet field, which is as defined in [Section 3](#).

5.3. OSPFv3 Link IFIT Sub-TLV

The Link IFIT sub-TLV encoded in the OSPFv3 E-Router-LSA TLV is constructed of three fields, a 2-octet Type field, a 2-octet Length field, and a multiple of 4-octet Value field. The following format is used:

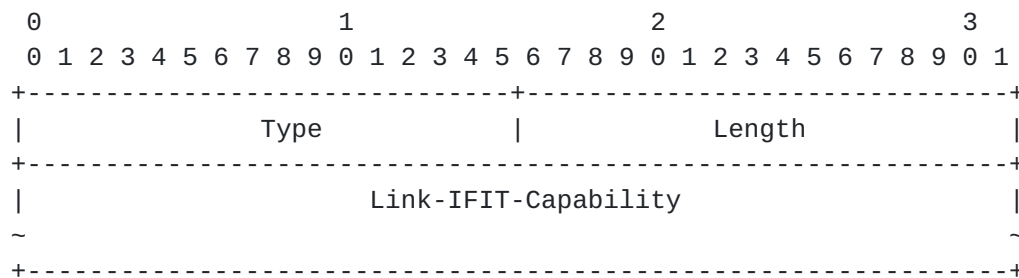


Fig. 6 OSPFv3 Link IFIT sub-TLV

Where:

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

6. 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 ingress routers to gather each router's IFIT capability for achieving the computation of TE paths or loose 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.

7. IANA Considerations

IANA is requested to allocate values for the following new TLV and sub-TLVs.

+-----+-----+	
Type	Description
+-----+-----+	
TBD	IS-IS Node IFIT Sub-TLV
TBD	IS-IS Link IFIT Sub-TLV
+-----+-----+	
+-----+-----+	
Type	Description
+-----+-----+	
TBD	OSPF Node IFIT Capability TLV
TBD	OSPFv2 Link IFIT Capability sub-TLV
TBD	OSPFv3 Link IFIT Capability sub-TLV
+-----+-----+	

8. Security Considerations

This document introduces new IGP Node and Link Attribute TLVs and sub-TLVs for the IFIT Capability advertisements at node and/or link granularity. It does not introduce any new security risks to IS-IS, OSPFv2 and OSPFv3.

9. 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.

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>>.
- [RFC5305] "IS-IS Extensions for Traffic Engineering", <<https://www.rfc-editor.org/info/rfc5305>>.
- [RFC7684] "OSPFv2 Prefix/Link Attribute Advertisement", <<https://www.rfc-editor.org/info/rfc7684>>.
- [RFC7752] "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", <<https://datatracker.ietf.org/doc/rfc7752/>>.
- [RFC7770] "Extensions to OSPF for Advertising Optional Router Capabilities", <<https://www.rfc-editor.org/info/rfc7770>>.
- [RFC7981] "IS-IS Extensions for Advertising Router Information", <<https://www.rfc-editor.org/info/rfc7981>>.
- [RFC8362] "OSPFv3 Link State Advertisement (LSA) Extensibility", <<https://www.rfc-editor.org/info/rfc8362>>.

10.2. 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

Fengwei Qin
China Mobile
32 Xuanwumenxi Ave.
Beijing
China

Email: qinfengwei@chinamobile.com

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

Email: chenhuan6@chinatelecom.cn

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

Email: pangran@chinaunicom.cn

