

Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 14, 2014

U. Chunduri
W. Lu
A. Tian
Ericsson Inc.
N. Shen
Cisco Systems, Inc.
February 10, 2014

IS-IS Extended Sequence number TLV
draft-ietf-isis-extended-sequence-no-tlv-02

Abstract

This document defines Extended Sequence number TLV to protect Intermediate System to Intermediate System (IS-IS) PDUs from replay attacks.

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 <http://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 August 14, 2014.

Copyright Notice

Copyright (c) 2014 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 (<http://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
1.1. Requirements Language	3
1.2. Acronyms	3
2. Replay attacks and Impact on IS-IS networks	4
2.1. IIHs	4
2.2. LSPs	4
2.3. SNPs	4
3. Extended Sequence Number TLV	4
3.1. Sequence Number Wrap	5
4. Mechanism and Packet Encoding	6
4.1. IIHs	6
4.2. SNPs	6
5. Backward Compatibility and Deployment	6
5.1. IIH and SNPs	7
6. IANA Considerations	7
7. Security Considerations	7
8. Contributors	7
9. Acknowledgements	7
10. Appendix A	8
10.1. Appendix A.1	8
10.2. Appendix A.2	8
11. Appendix B	9
11.1. Operational/Implementation consideration	9
12. References	9
12.1. Normative References	9
12.2. Informative References	9
Authors' Addresses	10

1. Introduction

With the rapid development of new data center infrastructures, due to its flexibility and scalability attributes, IS-IS has been adopted widely in various L2 and L3 routing deployment of the data centers for critical business operations. At the meantime the SDN-enabled networks even though put more power to Internet applications and also make network management easier, it does raise the security requirement of network routing infrastructure to another level.

A replayed IS-IS PDU can potentially cause many problems in the IS-IS networks ranging from bouncing adjacencies to black hole or even some form of Denial of Service (DoS) attacks as explained in Section 2. This problem is also discussed in security consideration section, in

the context of cryptographic authentication work as described in [RFC5304] and in [RFC5310].

Currently, there is no mechanism to protect IS-IS HELLO PDUs (IIHs) and Sequence number PDUs (SNPs) from the replay attacks. However, Link State PDUs (LSPs) have sequence number in the LSP header as defined in [ISO10589], with which it can effectively mitigate the intra-session replay attacks. But, LSPs are still susceptible to inter-session replay attacks.

This document defines Extended Sequence number (ESN) TLV to protect Intermediate System to Intermediate System (IS-IS) PDUs from replay attacks.

The new ESN TLV defined here thwart these threats and can be deployed with authentication mechanism as specified in [RFC5304] and in [RFC5310] for a more secure network.

Replay attacks can be effectively mitigated by deploying a group key management protocol (being developed as defined in [I-D.yeung-g-ikev2] and [I-D.hartman-karp-mrkmp]) with a frequent key change policy. Currently, there is no such mechanism defined for IS-IS. Even if such a mechanism is defined, usage of this TLV can be helpful to avoid replays before the keys are changed.

Also, it is believed, even when such key management system is deployed, there always will be some manual key based systems that co-exist with KMP (Key Management Protocol) based systems. The ESN TLV defined in this document is more helpful for such deployments.

1.1. 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].

1.2. Acronyms

CSNP	-	Complete Sequence Number PDU
ESN	-	Extended Sequence Number
IIH	-	IS-IS Hello PDU
KMP	-	Key Management Protocol (auto key management)
LSP	-	IS-IS Link State PDU

MKM - Manual Key management Protocols

PDU - Protocol Data Unit

PSNP - Partial Sequence Number PDU

SNP - Sequence Number PDU

2. Replay attacks and Impact on IS-IS networks

Replaying a captured protocol packet to cause damage is a common threat for any protocol. Securing the packet with cryptographic authentication information alone cannot mitigate this threat completely. This section explains the replay attacks and the applicability of the same for each IS-IS PDU.

2.1. IIHs

At the time of adjacency bring up an IS sends IIH packet with empty neighbor list (TLV 6) and with or without the authentication information as per provisioned authentication mechanism. If this packet is replayed later on the broadcast network all ISes in the broadcast network can bounce the adjacency to create a huge churn in the network.

2.2. LSPs

Normal operation of the IS-IS update Process as specified in [ISO10589] provides timely recovery from all LSP replay attacks. Therefore the use of the extensions defined in this document are prohibited in LSPs. Further discussion of the vulnerability of LSPs to replay attacks can be found in [I-D.ietf-karp-isis-analysis].

2.3. SNPs

A replayed CSNP can result in the sending of unnecessary PSNPs on a given link. A replayed CSNP or PSNP can result in unnecessary LSP flooding on the link.

3. Extended Sequence Number TLV

The Extended Sequence Number (ESN) TLV is composed of 1 octet for the Type, 1 octet that specifies the number of bytes in the Value field and a 12 byte Value field. This TLV is defined only for IIH and SNP PDUs.

x CODE - TBD.

x LENGTH - total length of the value field, which is 12 bytes.

x Value - 64-bit Extended Session Sequence Number (ESSN), which is followed by a 32 bit monotonically increasing per Packet Sequence Number (PSN).

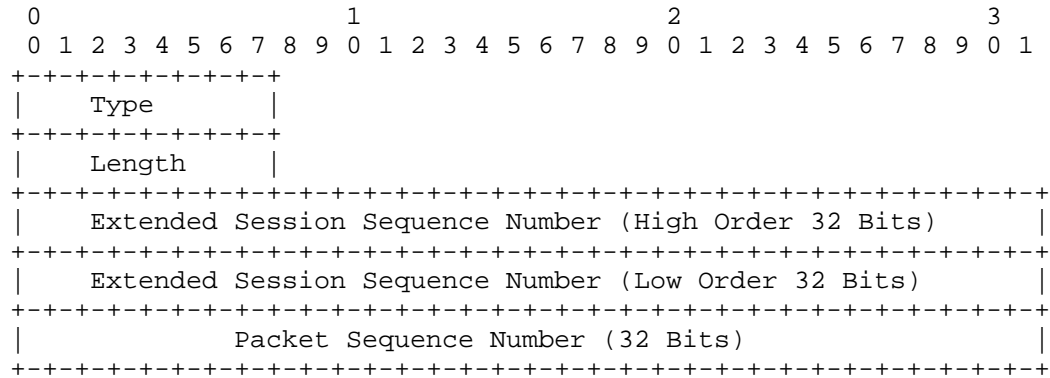


Figure 1: Extended Sequence Number (ESN) TLV

The ESN TLV defined here is optional. Though this is an optional TLV, this can be mandatory on a link when 'verify' mode is enabled as specified in Section 5.1. The ESN TLV MAY be present only in any IIH and SNP PDUs. A PDU with multiple ESN TLVs is invalid and MUST be discarded on receipt.

The 64 bit ESSN MUST be non-zero and MUST contain ever increasing number whenever it is changed due any situation as specified in Section 3.1. For each PDU which contains the ESN TLV the 96 bit unsigned integer value consisting of the 64 bit ESSN and 32 bit Packet Sequence Number (PSN) - where ESSN is the 64 MSBs - MUST be greater than the most recently received value in a PDU of the same type originated by the same IS.

3.1. Sequence Number Wrap

If the 32-bit Packet Sequence Number in ESN TLV wraps or for the cold restart of the router, the 64-bit ESSN value MUST be set higher than the previous value. IS-IS implementations MAY use guidelines provided in Section 10 for accomplishing this.

4. Mechanism and Packet Encoding

The encoding of ESN TLV in each IS-IS PDU is applicable is detailed below. Also refer, when to ignore processing of the ESN TLV as described in Section 5 for appropriate operation in the face of legacy node(s) in the network, which do not support the extensions defined in this document. If the received PDU with ESN TLV is accepted then the stored value for the corresponding originator, PDU type and level MUST be updated with the latest value received.

4.1. IIHs

ESN TLV information is maintained for each type of IIH PDU being sent on a given circuit. The procedures for encoding, verification and sequence number wrap scenarios are explained in Section 3.

4.2. SNPs

A separate CSNP/PSNP ESN TLV information is maintained per PDU type and per link. The procedures for encoding, verification and sequence number wrap scenarios are explained in Section 3.

5. Backward Compatibility and Deployment

The implementation and deployment of the ESN TLV can be done to support backward compatibility and gradual deployment in the network without requiring a flag day. This feature can also be deployed for the links in a certain area of the network where the maximum security mechanism is needed, or it can be deployed for the entire network.

The implementation SHOULD allow the configuration of ESN TLV feature on each IS-IS link level. The implementation SHOULD also allow operators to control the configuration of 'send' and/or 'verify' the feature of IS-IS PDUs for the links and for the node. In this document, the 'send' operation is to include the ESN TLV in its own IS-IS PDUs; and the 'verify' operation is to process the ESN TLV in the receiving IS-IS PDUs from neighbors.

In the face of an adversary doing an active attack, it is possible to have inconsistent data view in the network, if there is a considerable delay in enabling ESN TLV 'verify' operation from first node to the last node in the network. This can happen primarily because, replay PDUs can potentially be accepted by the nodes where 'verify' operation is still not provisioned at the time of the attack. To minimize such a window it is recommended that provisioning of 'verify' SHOULD be done in a timely fashion by the network operators.

5.1. IIH and SNPs

On the link level, ESN TLV involves the IIH PDUs and SNPs (both CSNP and PSNP). The "send" and "verify" modes described above can be set independently on each link and in the case of a broadcast network independently for each level.

To introduce ESN support without disrupting operations, ISs on a given interface are first configured to operate in 'send' mode. Once all routers operating on an interface are operating in 'send' mode 'verify' mode can be enabled on each IS. Once 'verify' mode is set for an interface all the IIH and SNP PDUs being sent on that interface MUST contain the ESN TLV. Any such PDU received without an ESN TLV MUST be discarded when 'verify' mode is enabled

6. IANA Considerations

This document requests that IANA allocate from the IS-IS TLV Codepoints Registry a new TLV, referred to as the "Extended Sequence Number" TLV, with the following attributes:

Type	Description	IIH	LSP	SNP	Purge
----	-----	---	---	---	-----
TBD	ESN TLV	Y	N	Y	N

Figure 2: IS-IS Codepoints Registry Entry

7. Security Considerations

This document describes a mechanism to the replay attack threat as discussed in the Security Considerations section of [RFC5304] and in [RFC5310]. This document does not introduce any new security concerns to IS-IS or any other specifications referenced in this document.

8. Contributors

Authors would like to thank Les Ginsberg for his significant contribution in detailed reviews and suggestions.

9. Acknowledgements

As some sort of sequence number mechanism to thwart protocol replays is a old mechanism, authors of this document do not make any claims on the originality of the overall protection idea described. Authors are thankful for the review and the valuable feedback provided by Acee Lindem and Joel Halpern.

10. Appendix A

IS-IS nodes implementing this specification SHOULD use available mechanisms to preserve the 64-bit Extended Session Sequence Number's strictly increasing property, whenever it is changed for the deployed life of the IS-IS node (including cold restarts).

This Appendix provides only guidelines for achieving the same and implementations can resort to any similar method as far as strictly increasing property of the 64-bit ESSN in ESN TLV is maintained.

10.1. Appendix A.1

One mechanism for accomplishing this is by encoding 64-bit ESSN as system time represented in 64-bit unsigned integer value. This MAY be similar to the system timestamp encoding for NTP long format as defined in Appendix A.4 of [RFC5905]. New current time MAY be used when the IS-IS node loses its sequence number state including in Packet Sequence Number wrap scenarios.

Implementations MUST make sure while encoding the 64-bit ESN value with current system time, it should not default to any previous value or some default node time of the system; especially after cold restarts or any other similar events. In general system time must be preserved across cold restarts in order for this mechanism to be feasible. One example of such implementation is to use a battery backed real-time clock (RTC).

10.2. Appendix A.2

One other mechanism for accomplishing this would be similar to the one as specified in [I-D.ietf-ospf-security-extension-manual-keying], to use the 64-bit ESSN as a wrap/boot count stored in non-volatile storage. This value is incremented anytime the IS-IS node loses its sequence number state including in Packet Sequence Number wrap scenarios.

The drawback of this approach per Section 6 of [I-D.ietf-ospf-security-extension-manual-keying], if used is applicable here. The only drawback is, it requires the IS-IS implementation to be able to save its boot count in non-volatile storage. If the non-volatile storage is ever repaired or upgraded such that the contents are lost, keys MUST be changed to prevent replay attacks.

11. Appendix B

11.1. Operational/Implementation consideration

Since the ESN is maintained per interface, per level and per PDU type, this scheme can be useful for monitoring the health of the IS-IS adjacency. A Packet Sequence Number skip on IIH can be recorded by the neighbors which can be used later to correlate with adjacency state changes over the interface. For instance in a multi-access media, all the neighbors have the skips from the same IIH sender or only one neighbor has the Packet Sequence Number skips can indicate completely different issues on the network. Effective usage of the TLV defined in this document for operational issues MAY also need more system information before making concrete conclusions and defining all that information is beyond the scope of this document.

12. References

12.1. Normative References

- [ISO.10589.1992]
International Organization for Standardization,
"Intermediate system to intermediate system intra-domain-
routing routine information exchange protocol for use in
conjunction with the protocol for providing the
connectionless-mode Network Service (ISO 8473)", ISO/
IEC 10589:2002, Second Edition, Nov. 2002.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network
Time Protocol Version 4: Protocol and Algorithms
Specification", RFC 5905, June 2010.

12.2. Informative References

- [I-D.hartman-karp-mrkmp]
Hartman, S., Zhang, D., and G. Lebovitz, "Multicast Router
Key Management Protocol (MaRK)", draft-hartman-karp-
mrkmp-05 (work in progress), September 2012.
- [I-D.ietf-karp-isis-analysis]
Chunduri, U., Tian, A., and W. Lu, "KARP IS-IS security
analysis", draft-ietf-karp-isis-analysis-02 (work in
progress), February 2014.

- [I-D.ietf-ospf-security-extension-manual-keying]
Bhatia, M., Hartman, S., Zhang, D., and A. Lindem,
"Security Extension for OSPFv2 when using Manual Key
Management", draft-ietf-ospf-security-extension-manual-
keying-06 (work in progress), November 2013.
- [I-D.weis-gdoi-mac-tek]
Weis, B. and S. Rowles, "GDOI Generic Message
Authentication Code Policy", draft-weis-gdoi-mac-tek-03
(work in progress), September 2011.
- [RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic
Authentication", RFC 5304, October 2008.
- [RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R.,
and M. Fanto, "IS-IS Generic Cryptographic
Authentication", RFC 5310, February 2009.
- [RFC6518] Lebovitz, G. and M. Bhatia, "Keying and Authentication for
Routing Protocols (KARP) Design Guidelines", RFC 6518,
February 2012.

Authors' Addresses

Uma Chunduri
Ericsson Inc.
300 Holger Way,
San Jose, California 95134
USA

Phone: 408 750-5678
Email: uma.chunduri@ericsson.com

Wenhu Lu
Ericsson Inc.
300 Holger Way,
San Jose, California 95134
USA

Email: wenhu.lu@ericsson.com

Albert Tian
Ericsson Inc.
300 Holger Way,
San Jose, California 95134
USA

Phone: 408 750-5210
Email: albert.tian@ericsson.com

Naiming Shen
Cisco Systems, Inc.
225 West Tasman Drive,
San Jose, California 95134
USA

Email: naiming@cisco.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

Z. Li
Q. Zhao
Huawei Technologies
October 21, 2013

IS-IS Extensions for MPLS Multi-Topology
draft-li-isis-mpls-multi-topology-00

Abstract

MPLS plays a key role in the process of implementing network virtualization. [I-D.li-mpls-network-virtualization-framework] proposes the framework to implement MPLS virtual network basedn on the architecture of central controller IGP. This document defines the corresponding IS-IS protocol extension and procedures to support MPLS Multi-Topology.

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 <http://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 April 24, 2014.

Copyright Notice

Copyright (c) 2013 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

(<http://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	2
3. Overview	3
3.1. Application for MRT FRR	3
4. IS-IS Extensions	4
4.1. IS-IS Label Mapping TLV	4
4.2. Label Sub-TLV	5
4.3. MPLS Multi-Topology Sub-TLV	6
4.4. Procedures	6
5. Compatibility	6
6. IANA Considerations	7
7. Security Considerations	7
8. References	7
8.1. Normative References	7
8.2. Informative References	7
Authors' Addresses	8

1. Introduction

As the virtual network operators develop, it is desirable to provide better network virtualization solutions to facilitate the service provision. [I-D.li-mpls-network-virtualization-framework] proposes a new framework to implement MPLS virtual network based on the architecture of central controlled IGP. It is to allocate MPLS global label for the virtual network topologies, network nodes and links by an IGP controller to IGP clients. Thus MPLS global labels becomes the unique identifications in the underlying networks to compose the virtual networks.

This document defines the corresponding IS-IS protocol extensions and procedures to support MPLS virtual network topology. The other document will define the corresponding IS-IS protocol extensions and procedures to support MPLS virtualized network nodes and links.

2. Terminology

Underlying Network: It is the network which the virtual network is built based on. The underlying network can be the physical network or the virtual network.

MPLS Virtual Network: The virtual network is built based on the underlying network. It is composed by virtual nodes and virtual links which are identified by MPLS global label. In this document, the concept of virtual network is the same as that of MPLS virtual network.

MPLS Virtual Network Topology: It is the topology of the MPLS virtual network. It can be identified multi-topology ID of corresponding virtual network. MPLS global label is allocated to represent the virtual network topology.

3. Overview

[I-D.li-rtgwg-cc-igp-arch] defines the central controlled architecture for IGP. In [I-D.li-mpls-network-virtualization-framework], a new framework is defined to implement MPLS virtual network based on central controlled IGP. In the MPLS virtual network, the virtual network topology can be identified by the Multi-Topology ID. The global label for the virtual network topology is allocated by the IGP controller and the label binding between the Multi-Topology ID and the Global Label are flooded from the IGP controller to IGP clients. When IGP clients receives the label binding, it can install the MPLS forwarding entry to map the incoming label to the forwarding instance corresponding to the Multi-Topology.

3.1. Application for MRT FRR

MRT FRR [I-D.ietf-rtgwg-mrt-frr-architecture] has been proposed to provide 100% coverage of FRR deployment in the network. There are two forwarding mechanisms defined in : IP tunnel forwarding mechanism and LDP Multi-Topology mechanism. IP tunnel forwarding mechanism need to set up IP tunnels which must introduce extra IP addresses. It is difficult to satisfy the scalability requirement for deployment. LDP Multi-Topology is a scalable way to implement the MRT FRR forwarding. But for the pure IP network it has to introduce the new MPLS protocol. Moreover, it may use more labels which are allocated for prefixes in three topologies: the default topology, the blue topology and the red topology.

When the MPLS virtual network mechanism is introduced, the MRT FRR forwarding in the IP network can be simplified greatly. Two MPLS global label can be allocated to identify the blue topology and the red topology. According to MRT calculation, the forwarding instances

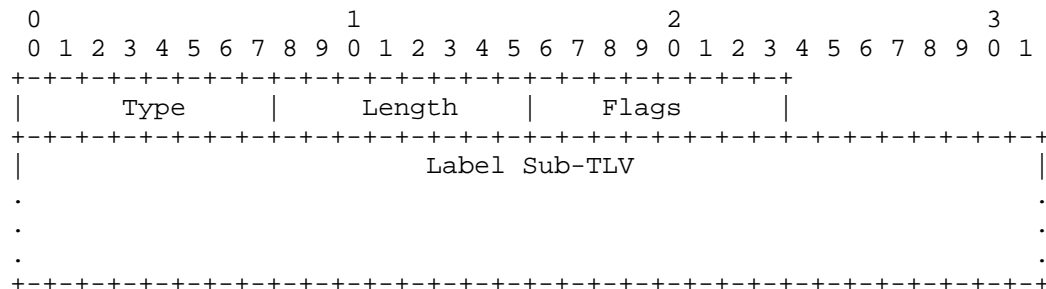
for the default topology, the red topology and the blue topology can be installed. For a specific IP prefix, the forwarding entry will consist of the primary path in the default topology and the secondary path in the red topology or the blue topology. IP packets will forward hop by hop according to the FIB of the default topology. When failure happens, they will be forwarded in the blue topology or the red topology which contains the secondary path. When forwarding, the global label for the Multi-Topology is encapsulated for the IP packets. When the next hop node receives the packets, it will decapsulate the label and map to the corresponding forwarding instance of the red topology or the blue topology according to the MPLS forwarding entry. When determine the outgoing interface and the next hop after looking up the Multi-Topology FIB according to the destination IP address, the packet will encapsulate the global label again which represent the red topology or the blue topology and will be forwarded to the next hop. This forwarding process will be done by each node until it reaches the destination.

In the MRT FRR process, there are only two MPLS forwarding entries to map the label to the red topology and the blue topology. Moreover, it is done by IGP extensions instead of introducing LDP, which can also simplify the network operation and management.

4. IS-IS Extensions

4.1. IS-IS Label Mapping TLV

A new IS-IS TLV, call as Label Mapping TLV, is introduced to allocate MPLS label. The IS-IS Label Mapping TLV format is shown in the following figure. The type of IS-IS Label Mapping TLV is to be defined by IANA. The flags in the TLV are to be defined. There are a series of sub-TLV in the TLV which length can be up to 252 octets. There MUST be at least one Label Sub-TLV and one FEC sub-TLV in the TLV. The Label sub-TLV contains the label value allocated. The FEC sub-TLV contains the Forwarding Equivalent Class for which the label is allocated. In this document, one FEC sub-TLV is defined, called as MPLS Multi-Topology sub-TLV.



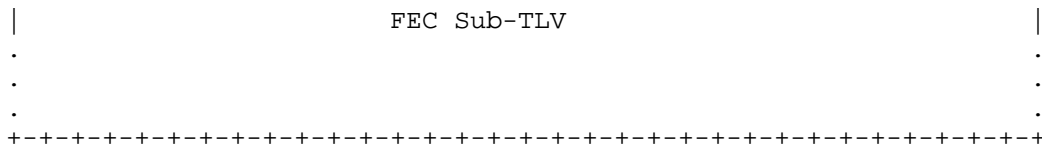


Figure 1: IS-IS Label Mapping TLV format

4.2. Label Sub-TLV

As defined in [I-D.li-mpls-global-label-framework], there are two types of methods to allocated global label: the first one is to use the existing MPLS label range, that is, from 16 to 2^{20} ; the second one is to use the expanded MPLS label range which can be more than 2^{20} . Taking into account the future expansion, the Label Sub-TLV has following format:

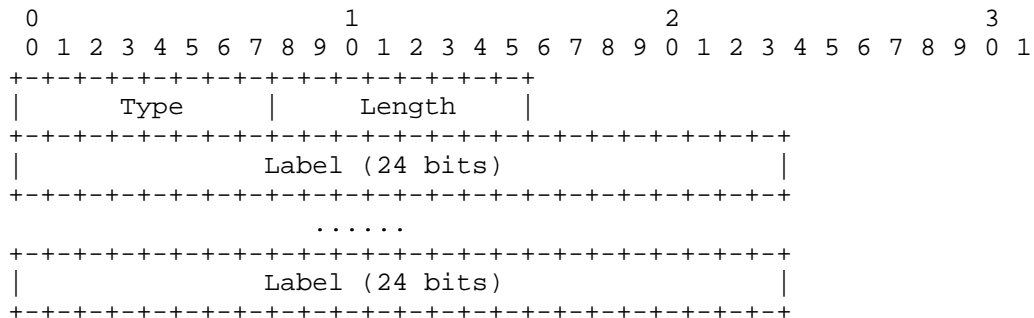


Figure 2: Label Sub-TLV Format

- o Type: 1 octet of sub-TLV type. It is to be allocated by IANA.
- o Length: 1 octet of length of the value field of the sub-TLV. It can be up to 252 octets.
- o MPLS Label Field: variable length. It consists of one or more labels. Each label is encoded as 3 octets, where the high-order 20 bits contain the label value, and the low order bit contains "Bottom of Stack".

The label stack in the Label sub-TLV can construct a big label range which can exceed 2^{20} . If there is only one label field, it is consistent with the existing MPLS label range from 16 to 2^{20} .

4.3. MPLS Multi-Topology Sub-TLV

When implement virtual topology, the global label is allocated for the Multi-topology ID. MPLS Multi-Topology ID can be seen as the FEC for the label mapping. The MPLS Multi-Topology sub-TLV format has following format:

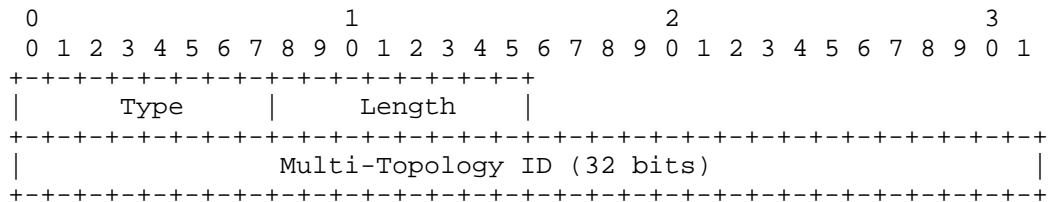


Figure 3: Multi-Topology Sub- TLV format

- o Type: 1 octet of sub-TLV type. It is to be allocated by IANA.
- o Length: 1 octet of length of the value field of the sub-TLV. It is 4.
- o Multi-Topology ID: 4 octets. It contains the MPLS Multi-Topology ID for which the global label is allocated to implement virtual topology.

4.4. Procedures

When the IGP controller needs to implement the MPLS virtual network topology, the IGP controller MUST originate a new LSP comprising the Label Mapping TLV for the MPLS virtual network topology. The Label Mapping TLV MUST contain one or more pairs of the Label sub-TLV and the Multi-Topology ID sub-TLV. If the length of these sub-TLVs can exceeds 252 octets, there SHOULD be multiple Label Mapping TLVs in IS-IS LSP.

When receiving the Label Mapping to implement the virtual network topology, the IGP clients SHOULD get the global label and the corresponding multi-topology from the sub-TLVs in the label mapping TLV and install MPLS forwarding entry accordingly.

5. Compatibility

Routers that do not support these MPLS Virtualization extensions SHOULD silently ignore the TLV and the sub-TLVs defined in this document.

6. IANA Considerations

This document request to allocate a type value for the Label Mapping TLV, a type value for the Label sub-TLV and a type value for the MPLS Multi-Topology sub-TLV.

7. Security Considerations

TBD.

8. References

8.1. Normative References

[I-D.li-mpls-global-label-framework]

Li, Z., Zhao, Q., and T. Yang, "A Framework of MPLS Global Label", draft-li-mpls-global-label-framework-00 (work in progress), July 2013.

[I-D.li-mpls-network-virtualization-framework]

Li, Z. and M. Li, "Framework of Network Virtualization Based on MPLS Global Label", draft-li-mpls-network-virtualization-framework-00 (work in progress), October 2013.

[I-D.li-rtgwg-cc-igp-arch]

Li, Z., Chen, H., and G. Yan, "An Architecture of Central Controlled Interior Gateway Protocol (IGP)", draft-li-rtgwg-cc-igp-arch-00 (work in progress), October 2013.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[I-D.gredler-isis-label-advertisement]

Gredler, H., Amante, S., Scholl, T., and L. Jalil, "Advertising MPLS labels in IS-IS", draft-gredler-isis-label-advertisement-03 (work in progress), May 2013.

[I-D.ietf-rtgwg-mrt-frr-architecture]

Atlas, A., Kebler, R., Envedi, G., Csaszar, A., Tantsura, J., Konstantynowicz, M., and R. White, "An Architecture for IP/LDP Fast-Reroute Using Maximally Redundant Trees", draft-ietf-rtgwg-mrt-frr-architecture-03 (work in progress), July 2013.

[I-D.previdi-isis-segment-routing-extensions]

Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., and
S. Litkowski, "IS-IS Extensions for Segment Routing",
draft-previdi-isis-segment-routing-extensions-02 (work in
progress), July 2013.

Authors' Addresses

Zhenbin Li
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: lizhenbin@huawei.com

Quintin Zhao
Huawei Technologies
Boston, MA
USA

Email: quintin.zhao@huawei.com

Network Working Group
Internet Draft
Intended status: Standards Track
Expires: August 18, 2014

B. Liu
Huawei Technologies
Bruno Decraene
Orange
I. Farrer
Deutsche Telekom AG
M. Abrahamsson
T-System
February 14, 2014

ISIS Auto-Configuration
draft-liu-isis-auto-conf-01.txt

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 <http://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 August 18, 2014.

Copyright Notice0

Copyright (c) 2012 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 (<http://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.

Abstract

This document describes mechanisms for IS-IS to be self-configuring. Such mechanisms could reduce the management burden to configure a network. One obvious environment that could benefit from these mechanisms is IPv6 home network where plug-and-play would be expected. Besides home network, some simple enterprise/ISP networks might also benefit from the self-configuring mechanisms.

Table of Contents

1. Introduction	3
2. Design Scope	3
3. Protocol Specification	4
3.1. IS-IS Default Configuration	4
3.2. IS-IS NET Generation	4
3.3. IS-IS NET Duplication Detection and Resolution.....	5
3.3.1. Router-Hardware-Fingerprint TLV	5
3.3.2. NET Duplication Detection and Resolution Procedures.	5
3.4. Authentication TLV	6
3.5. Wide Metric	7
3.6. Adjacency Formation Consideration	7
4. Co-existence with Other IGP Auto-configuration	7
5. Security Considerations	7
6. IANA Considerations	8
7. Acknowledgments	8
8. References	8
8.1. Normative References	8
8.2. Informative References	8

1. Introduction

This memo describes mechanisms for IS-IS [RFC1195][RFC5308] to be auto-configuring. Such mechanisms could reduce the management burden to configure a network. One example is home network where plug-and-play would be expected. Besides home network, some simple enterprise/ISP networks might also potentially benefit from the auto-configuring mechanisms.

In addition, this memo defines how such un-configured routers should behave, also limits the risk on existing network using IS-IS (Section 3.4 & 3.5).

IS-IS auto-configuration mainly contains the following aspects:

1. IS-IS Default Configuration
2. IS-IS NET self-generation
3. NET duplication detection and resolution
4. Authentication and Wide Metric TLV

2. Design Scope

The auto-configuring mechanisms are not specifically designed based on IPv4 or IPv6.

The auto-configuring mechanisms enabled interfaces are assumed to have a 48-bit MAC address.

The main targeted application scenarios are supposed to be home networks or small enterprise networks .etc. where plug-n-play is expected and complex topology/hierarchy is not needed. Sophisticate requirements from service provider networks are out of scope.

So this document does not provide a complete configuration-free alternative to the IS-IS protocol. The following features of IS-IS are NOT supported by this document:

- o Auto-configuring multiple IS-IS processes. The auto-configuration mechanisms only support configuring a single process.
- o Route between multiple IS-IS areas. The auto-configuration mechanisms only support routers that are within a single area.

- o Auto-configuring multiple operation levels. The auto-configuration mechanisms only support level-1 operation mode.
- o This document does not consider interoperability with other routing protocols.

3. Protocol Specification

3.1. IS-IS Default Configuration

- o IS-IS SHOULD be enabled on all interfaces in a router that requires the IS-IS auto-configuration as default. For some specific situations, interface MAY be excluded if it is clear that running IS-IS on the interface is not required.
- o IS-IS interfaces MUST be auto-configured to an interface type corresponding to their layer-2 capability. For example, Ethernet interfaces will be auto-configured as broadcast networks and Point-to-Point Protocol (PPP) interfaces will be auto-configured as Point-to-Point interfaces.
- o IS-IS auto-configuration interfaces MUST be configured with level-1.

3.2. IS-IS NET Generation

In IS-IS, a router (known as an IS) is identified by an Network Entity Title (NET) which is the address of a Network Service Access Point (NSAP) and represented with an IS-IS specific address format. The NSAP is a logical entity which represents an instance of the IS-IS protocol running on an IS.

The NET consists of three parts. The auto-generation mechanisms of each part are described as the following:

Area address: This field is 1 to 13 octets in length. In IS-IS auto-configuring, this field MUST be 0 in 13 octets length.

System ID: This field follows the area address field, and is 6 octets in length. As specified in IS-IS protocol, this field must be unique among all level-1 routers in the same area when the IS operates at Level 1. In IS-IS auto-configuring, this field SHOULD be the MAC address of one IS-IS enabled interface.

NSEL: This field is the N-selector, and is 1 octet in length. In IS-IS auto-configuring, it must be set to "00".

3.3. IS-IS NET Duplication Detection and Resolution

As described in Section 3, in IS-IS auto-configuring the NETs are distinguished by the System ID field in which it is a MAC address. So for IS-IS neighbors' NET duplication, it is equal to MAC address duplication in a LAN, which means a serious problem that devices need to be changed. So the NET duplication detection and resolution mechanism is actually used between non-neighbors which are within the same IS-IS area.

The rational of IS-IS NET duplication detection and resolution is described as the following.

3.3.1. Router-Hardware-Fingerprint TLV

The Router-Hardware-Fingerprint TLV is defined in [OSPFv3AC]. This document re-uses it to achieve NET duplication detection.

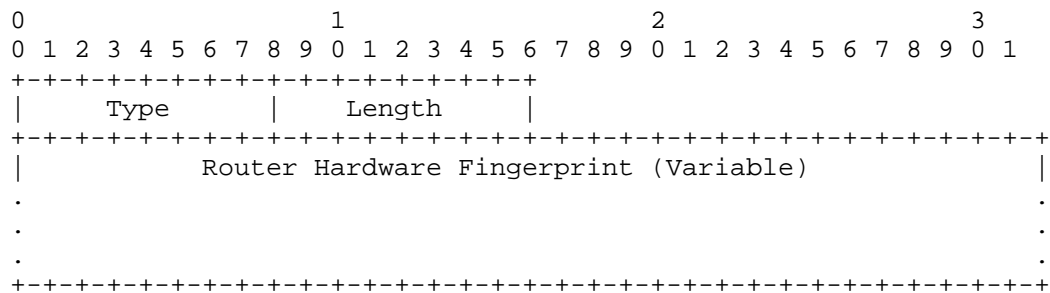


Figure 1 Router-Hardware-Fingerprint TLV Format

As defined in [OSPFv3AC], the contents of the hardware fingerprint should be some combination of CPU ID, or serial number(s) that provides an extremely high probability of uniqueness. It MUST be based on hardware attributes that will not change across hard and soft restarts. The length of the Router-Hardware-Fingerprint is variable but must be 32 octets or greater.

Note that, since the TLV is to detect MAC address based NET duplication, the TLV content MUST NOT only use MAC address. MAC address plus other information are also not recommended to use.

3.3.2. NET Duplication Detection and Resolution Procedures

1) Flood the Router-Hardware-Fingerprint TLVs

When an IS-IS auto-configuration router gets online, it MUST include the Router-Hardware-Fingerprint TLV in the first originated level-1

LSP. Then all the routers in the area could receive the information in the TLV.

2) Compare the received Router-Hardware-Fingerprint TLVs

An IS-IS auto-configuring router MUST compare a received self-originated LSP's Router-Hardware-Fingerprint TLV against its own one. If they are equal, it means the LSP was indeed originated by the router itself; if they are not equal, it means some other router has the same NET originated the LSP, thus there is a NET duplication.

3) Duplication resolution

When NET duplication occurs, the router with the numerically smaller router hardware fingerprint MUST generate a new NET.

4) Purge the LSPs containing duplicated NET

Before flooding the new NET, the LSP with the prior duplicate NET MUST be purged. And any IS-IS neighbor adjacencies MUST be reestablished.

5) Re-join the network with the new NET

After purging the LSPs with the duplicated NET, the router re-join the IS-IS auto-configuration network with the newly generated NET.

3.4. Authentication TLV

Every IS-IS auto-configuration message MUST include an authentication TLV (TLV 10, [RFC5304]) with the Type 1 authentication mode ("Cleartext Password") in order to avoid the auto-conf router to accidentally join an existing IS-IS network which is not intended to be auto-configured.

This feature is necessary because a low end CPE joining an existing IS-IS network might seriously break it or cause unnecessary management confusion.

The cleartext password is specified as "isis-autoconf". Routers that implement IS-IS auto-configuration MUST use this password as default, so that different routers could authenticate each other with no human intervene as default. And routers MUST be able to set manual password by the users.

3.5. Wide Metric

IS-IS auto-configuration routers SHOULD support wide metric (TLV 22, [RFC5305]). It is recommended that IS-IS auto-configuration routers use a high metric value (e.g. 1000000) as default in order to typically prefer the manually configured adjacencies rather than the auto-conf ones.

3.6. Adjacency Formation Consideration

ISIS does not require strict hold timers matching to form adjacency. But a reasonable range might be needed. Whether we need to specify a best practice timers in ISIS-AC is an open question.[TBD].

4. Co-existence with Other IGP Auto-configuration

If a router supports multiple IGP auto-configuration mechanisms (e.g. both IS-IS auto-configuration and OSPF auto-configuration), then in practice it is a problem that there should be a mechanism to decide which IGP to be used, or even both.

However, it is not proper to specify choice/interaction of multiple IGPs in any standalone IGP auto-configuration protocols. It should be done in the CPE level. Currently, there is some relevant work emerging, for example, the suggestion from [HOMENET-HNCP] is to have the proposed HNCP (Home Network Control Protocol) choose what IGP should be used.

5. Security Considerations

Unwanted routers could easily join in an existing IS-IS auto-configuration network by setting the authentication password as "isis-autoconf" default value or sniff the cleartext password online. However, this is a common security risk shared by other IS-IS networks that don't set proper authentication mechanisms. For wired deployment, the wired line itself could be considered as an implicit authentication that normally unwanted routers are not able to connect to the wire line; for wireless deployment, the authentication could be achieved at the lower wireless link layer.

Malicious router could modify the SystemID field to cause NET duplication detection and resolution, thus cause the routing system to vibrate.

6. IANA Considerations

The Router Hardware Fingerprint TLV type code needs an assignment by IANA.

7. Acknowledgments

Many useful comments and contributions were made by Sheng Jiang.

This document was inspired by [OSPFv3AC].

8. References

8.1. Normative References

- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.
- [RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, October 2008.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.

8.2. Informative References

- [OSPFv3AC] Lindem, A., and J. Arkko, "OSPFv3 Auto-Configuration", Work in Progress, October 2013
- [HOMENET-HNCP] Stenberg, M., and S. Barth, "Home Networking Control Protocol", Work in Progress, February 05

Authors' Addresses

Bing Liu
Huawei Technologies Co., Ltd
Q14, Huawei Campus
No.156 Beiqing Rd.
Hai-Dian District, Beijing 100095
P.R. China

Email: leo.liubing@huawei.com

Bruno Decraene
Orange
Issy-les-Moulineaux
FR

Email: bruno.decraene@orange.com

Ian Farrer
Deutsche Telekom AG
Bonn,
Germany

Email: ian.farrer@telekom.de

Mikael Abrahamsson
T-Systems
Stockholm,
Sweden

Email: mikael.abrahamsson@t-systems.se

IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: August 18, 2014

P. Sarkar, Ed.
H. Gredler
S. Hegde
H. Raghuvier
Juniper Networks, Inc.
S. Litkowski
B. Decraene
Orange
February 14, 2014

Advertising Per-node Admin Tags in IS-IS
draft-psarkar-isis-node-admin-tag-01

Abstract

This document describes an extension to IS-IS protocol [ISO10589], [RFC1195] to add an optional operational capability, that allows tagging and grouping of the nodes in an IS-IS domain. This allows simple management and easy control over route and path selection, based on local configured policies.

This document describes the protocol extensions to disseminate per-node admin-tags in IS-IS protocols.

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 <http://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 August 18, 2014.

Copyright Notice

Copyright (c) 2014 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 (<http://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	3
2. Administrative Tag	3
3. TLV format	3
3.1. Per-node Admin Tag sub-TLV	3
3.2. Ordering of tags	4
4. Applications	4
5. Security Considerations	5
6. IANA Considerations	5
7. Acknowledgments	5
8. References	5
8.1. Normative References	5
8.2. Informative References	5
Authors' Addresses	6

1. Introduction

This document provides mechanisms to advertise per-node administrative tags in the IS-IS Link State PDU [RFC1195]. In certain path-selection applications like for example in traffic-engineering or LFA [RFC5286] selection there is a need to tag the nodes based on their roles in the network and have policies to prefer or prune a certain group of nodes.

2. Administrative Tag

For the purpose of advertising per-node administrative tags within IS-IS, a new sub-TLV to the IS-IS Router Capability TLV-242 that is defined in [RFC4971] is proposed. Because path selection is a functional set which applies both to TE and non-TE applications the same has not added as a new sub-TLV in the Traffic Engineering TLVs [RFC5305].

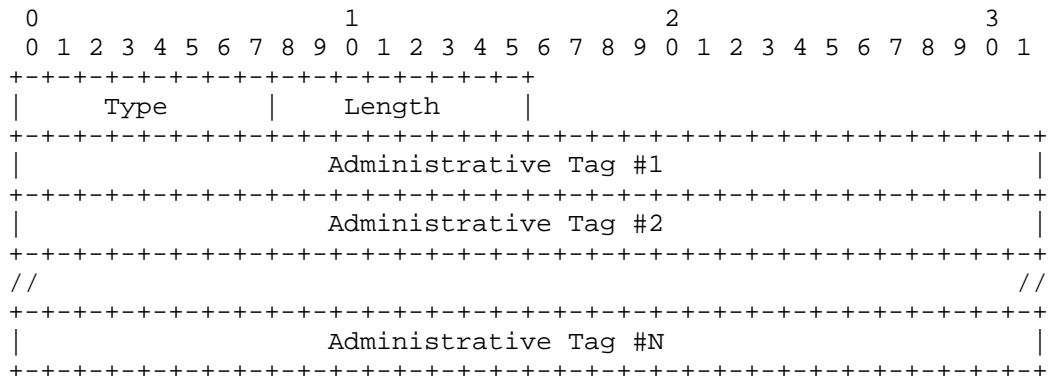
An administrative Tag is a 32-bit integer value that can be used to identify a group of nodes in the IS-IS domain. The new sub-TLV specifies one or more administrative tag values. An IS-IS node advertises the set of groups it is part of in the specific IS-IS level. As an example, all PE-nodes may be configured with certain tag value, whereas all P-nodes are configured with a different tag value in.

The new sub-TLV defined will be carried inside the IS-IS Router Capability TLV-242 (defined in [RFC4971]) in the Link State PDUs originated by the router. Link State PDUs [ISO10589] has level-wise (i.e. L1 or L2) flooding scope. Choosing the flooding scope to flood the group tags are defined by the policies and is a local matter. Once a group tag is decided in a specific level the same will be inserted in the administrative tag sub-TLV in the Link State PDU for the same level. Implementations should allow configuring both a 'global' and 'per-level' admin tag. In the absence of a specific admin tag configuration for a specific level the global admin tag should be copied in to the LSP PDU for the same level.

3. TLV format

3.1. Per-node Admin Tag sub-TLV

The new Administrative Tag sub-TLV, like other ISIS Capability sub-TLVs, is formatted as Type/Length/Value (TLV) triplets. Figure 1 below shows the format of the new sub-TLV.



Type : TBA

Length: A 8-bit field that indicates the length of the value portion in octets and will be a multiple of 4 octets dependent on the number of tags advertised.

Value: A sequence of multiple 4 octets defining the administrative tags.

Figure 1: IS-IS per-node Administrative Tag sub-TLV

The 'Per-node Admin Tag' sub-TLV may be generated more than once by an originating router. This MAY happen if a node carries more than 63 per-node admin groups and a single sub-TLV does not provide sufficient space. As such occurrence of the 'Per-node Admin Tag' sub-TLV does not cancel previous announcements, but rather is cumulative.

3.2. Ordering of tags

The semantics of the tag order are implementation-dependent. There is no implied meaning to the ordering of the tags that indicates a certain operation or set of operations that need to be performed based on the ordering.

Each tag SHOULD be treated as an independent identifier that MAY be used in policy to perform a policy action. Whether or not tag A precedes or succeeds tag B SHOULD not change the meaning of the tag set.

4. Applications

Increased deployment of Loop Free Alternates (LFA) [RFC5286] has

exposed some limitations. A recent draft on Operation management of Loop Free Alternates [I-D.ietf-rtgwg-lfa-manageability] proposes refinements to address those limitations. One of the proposed refinements is to be able to group the nodes in IGP domain with administrative tags and engineer the alternate paths based on configured policies.

The proposal in this document helps provide the capability to advertise group tags within IS-IS protocol and perform policy based LFA selection. The policies configured on each node can then make use of these tags to prefer or prune LFAs via certain group of nodes.

5. Security Considerations

This document does not introduce any further security issues other than those discussed in [ISO10589] and [RFC1195].

6. IANA Considerations

IANA maintains the registry for the Router Capability sub-TLVs. IS-IS Administrative Tags will require one new type code for the sub-TLV defined in this document.

7. Acknowledgments

8. References

8.1. Normative References

[ISO10589]
"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), ISO/IEC 10589:2002, Second Edition.", Nov 2002.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[I-D.ietf-rtgwg-lfa-manageability]
Litkowski, S., Decraene, B., Filsfils, C., and K. Raza,
"Operational management of Loop Free Alternates",

draft-ietf-rtgwg-lfa-manageability-00 (work in progress),
May 2013.

- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.
- [RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.
- [RFC5286] Atlas, A. and A. Zinin, "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, September 2008.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

Authors' Addresses

Pushpasis Sarkar (editor)
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: psarkar@juniper.net

Hannes Gredler
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: hannes@juniper.net

Shraddha Hegde
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: shraddha@juniper.net

Harish Raghuveer
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: hraghuveer@juniper.net

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com

Bruno Decraene
Orange

Email: bruno.decraene@orange.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 28, 2014

X. Xu
M. Chen
Huawei
December 25, 2013

Advertising Global Labels or SIDs Using IS-IS
draft-xu-isis-global-label-sid-adv-00

Abstract

Segment Routing (SR) is a new MPLS paradigm in which each SR-capable router is required to advertise global MPLS labels or Segment IDs (SID) for its attached prefixes by using link-state IGPs, e.g., IS-IS. One major challenge associated with such global MPLS label or SID advertisement mechanism is how to avoid a given global MPLS label or SID from being allocated by different routers to different prefixes. Although such global label or SID allocation collision problem can be addressed through manual allocation, it is error-prone and nonautomatic therefore may not be suitable in large-scale SR network environments. This document proposes an alternative approach for allocating and advertising global MPLS labels or SIDs via IS-IS so as to eliminate the potential risk of label allocation collision.

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 <http://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 June 28, 2014.

Copyright Notice

Copyright (c) 2013 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 (<http://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
1.1. Requirements Language	3
2. Terminology	3
3. Advertising Label Bindings for Prefixes using IS-IS	3
4. Advertising SID Bindings for Prefixes using IS-IS	4
5. Requesting Label Bindings for Prefixes using IS-IS	5
6. Requesting SID Bindings for Prefixes using IS-IS	5
7. Mapping Server Redundancy	6
8. Acknowledgements	7
9. IANA Considerations	7
10. Security Considerations	7
11. References	7
11.1. Normative References	7
11.2. Informative References	7
Authors' Addresses	8

1. Introduction

Segment Routing (SR) [I-D.filsfils-rtgwg-segment-routing] is a new MPLS paradigm in which each SR-capable router is required to advertise global MPLS labels or Segment IDs (SID) for its attached prefixes by using link-state IGPs, e.g., IS-IS [I-D.previdi-isis-segment-routing-extensions]. One major challenge associated with such global MPLS label or SID advertisement mechanism is how to avoid a given global MPLS label or SID from being allocated by different routers to different prefixes. Although such global label or SID allocation collision problem can be addressed through manual allocation, it is error-prone and nonautomatic therefore may not be suitable in large-scale SR network environments.

This document proposes an alternative approach for allocating and advertising global MPLS labels or SIDs via IS-IS so as to eliminate the potential risk of label allocation collision. The basic idea of this approach is to allow a particular IGP router to allocate global MPLS labels or SIDs for those prefixes attached to each SR-capable router and meanwhile advertise the corresponding label or SID

bindings in the IGP domain scope. That particular IGP router is therefore referred to as a mapping server. As for how the mapping server knows which prefixes need to be allocated with global labels or SIDs, it can be achieved either by configuration on the mapping server or by advertisement from SR-capable routers. In the multi-level scenario where route summarization between levels is enabled, the IP longest-match algorithm SHOULD be used by SR-capable routers when processing label or SID bindings advertised by the mapping server, just as the mechanism defined in [RFC5283].

1.1. 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].

2. Terminology

This memo makes use of the terms defined in [I-D.filsfils-rtgwg-segment-routing] and [RFC4971].

3. Advertising Label Bindings for Prefixes using IS-IS

A mapping server could use one or more of the following TLVs to advertise global labels for those prefixes which need to be allocated with global labels.

- o TLV-135 (IPv4) [RFC5305]
- o TLV-235 (MT-IPv4) [RFC5120]
- o TLV-236 (IPv6) [RFC5308]
- o TLV-237 (MT-IPv6) [RFC5120]

A Label Binding Sub-TLV (TBD) as shown below is associated with a prefix which is contained in one of the above TLVs:

```

      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=TBD   |  Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|P|  Reserved |             MPLS Label (20 bit)             |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Type: TBD

Length: 4

P-Flag: if set, the penultimate hop router MUST perform PHP action on the allocated MPLS label. For a given prefix, the P-Flag in the Label Binding Sub-TLV MUST be set to the same value as that of the P-Flag in the Label Request Sub-TLV if a label request message (See Section 5 of this document) for that prefix is received by the mapping server.

MPLS Label: a global label for the prefix which is carried in the TLV containing this sub-TLV.

Since the mapping server uses these TLVs for label binding advertisement purpose other than building the normal IP routing table, the Metric field MUST be set to a value larger than MAX_PATH_METRIC (i.e., 0xFE000000) according to the following specification as defined in [RFC5305] "...If a prefix is advertised with a metric larger than MAX_PATH_METRIC (0xFE000000, see paragraph 3.0), this prefix MUST NOT be considered during the normal SPF computation. This allows advertisement of a prefix for purposes other than building the normal IP routing table...". In addition, when propagating those TLVs across levels, the Label Binding Sub-TLVs contained in them MUST be preserved.

4. Advertising SID Bindings for Prefixes using IS-IS

A mapping server could use one or more of the Extended IP Reachability TLVs (i.e., TLV-135, TLV-235, TLV-236 and TLV-237) to advertise SIDs for those prefixes which need to be allocated with SIDs.

A SID Binding Sub-TLV (TBD) as shown below is associated with a prefix which is contained in one of the above TLVs:

```

      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=TBD   |  Length   |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     SID                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Type: TBD

Length: 4

SID: a SID for the prefix which is carried in the TLV containing this sub-TLV.

Since the mapping server uses these TLVs for label binding advertisement purpose other than building the normal IP routing table, the Metric field MUST be set to a value larger than MAX_PATH_METRIC (i.e., 0xFE000000). In addition, when propagating those TLVs across levels, the SID Binding Sub-TLVs contained in them MUST be preserved.

5. Requesting Label Bindings for Prefixes using IS-IS

When advertising IP reachability information by using one of the Extended IP Reachability TLVs (i.e., TLV-135, TLV-235, TLV-236 and TLV-237), SR-capable IS-IS routers SHOULD mark those attached prefixes which need to be allocated with global labels by associating each of these prefixes with a Label Request sub-TLV (type code=TBD) as shown below. In addition, when propagating those TLVs across levels, the Label Request Sub-TLVs contained in them MUST be preserved.

```

      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=TBD   |  Length   |                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|P|                                     Reserved                    |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Type: TBD

Length: 4

P-Flag: if set, the penultimate hop router MUST perform PHP action on the required label.

In the multi-level scenario where route summarization between levels is required, separate Extended IP Reachability TLVs other than those for IP reachability advertisement purpose SHOULD be used for label binding request purpose. Since these separate TLVs are not used for the purpose of building the normal IP routing table, the Metric field MUST be set to a value larger than MAX_PATH_METRIC (i.e., 0xFE000000). In addition, when propagating those TLVs across levels, the Label Request Sub-TLVs contained in them MUST be preserved.

6. Requesting SID Bindings for Prefixes using IS-IS

When advertising IP reachability information by using one of the Extended IP Reachability TLVs (i.e., TLV-135, TLV-235, TLV-236 and TLV-237), SR-capable IS-IS routers SHOULD mark those attached prefixes which need to be allocated with SIDs by associating each of

these prefixes with a SID Request sub-TLV (Type Code=TBD and Length=0)

In the multi-level scenario where route summarization between levels is required, separate Extended IP Reachability TLVs other than those for IP reachability advertisement purpose SHOULD be used for SID binding request purpose. Since these separate TLVs are not used for the purpose of building the normal IP routing table, the Metric field MUST be set to a value larger than MAX_PATH_METRIC (i.e., 0xFE000000). In addition, when propagating those TLVs across levels, the SID Request Sub-TLVs contained in them MUST be preserved.

7. Mapping Server Redundancy

For redundancy purpose, more than one router could be configured as candidates for mapping servers. Each candidate for mapping servers SHOULD advertise its capability of being a mapping servers by using IS-IS Router Capability TLV. The one with the highest priority SHOULD be elected as the primary mapping server which is eligible to allocate and advertise global labels or SIDs for prefixes on behalf of SR-capable routers. The comparison of IS-IS System ID breaks the tie between two or more candidates with the same highest priority. Meanwhile, the one with the second highest priority SHOULD be elected as a backup mapping server. This backup mapping server SHOULD advertise the same label bindings as those advertised by the primary mapping server. In this way, the unnecessary changes to the data plane (i.e., MPLS forwarding table) of SR-capable routers can be avoided in the event of mapping server failover.

Each candidate mapping server SHOULD advertise its capability of being a mapping server and the corresponding priority for mapping server election by attaching a Mapping Server Capability Sub-TLV (type code=TBD) shown as below to an IS-IS Router Capability TLV [RFC4971] with the S flag set (with domain-wide flooding scope).

```

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=TBD   |  Length   |  Priority   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Type: TBD

Length: 1

Priority: the priority for mapping server election.

8. Acknowledgements

The authors would like to thank .

9. IANA Considerations

TBD.

10. Security Considerations

This document does not introduce any new security considerations.

11. References

11.1. Normative References

- [I-D.previdi-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., and S. Litkowski, "IS-IS Extensions for Segment Routing", draft-previdi-isis-segment-routing-extensions-04 (work in progress), October 2013.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.
- [RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.
- [RFC5283] Decraene, B., Le Roux, JL., and I. Minei, "LDP Extension for Inter-Area Label Switched Paths (LSPs)", RFC 5283, July 2008.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.

11.2. Informative References

[I-D.filsfils-rtgwg-segment-routing]

Filsfils, C., Previdi, S., Bashandy, A., Decraene, B.,
Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R.,
Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe,
"Segment Routing Architecture", draft-filsfils-rtgwg-
segment-routing-01 (work in progress), October 2013.

Authors' Addresses

Xiaohu Xu
Huawei

Email: xuxiaohu@huawei.com

Mach Chen
Huawei

Email: mach.chen@huawei.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 18, 2014

X. Xu
Huawei
S. Kini
Ericsson
S. Sivabalan
C. Filsfils
Cisco
December 18, 2013

Signaling Entropy Label Capability Using IS-IS
draft-xu-isis-mpls-elc-00

Abstract

Multi Protocol Label Switching (MPLS) has defined a mechanism to load balance traffic flows using Entropy Labels (EL). An ingress LSR cannot insert ELs for packets going into a given tunnel unless an egress LSR has indicated via signaling that it can process ELs on that tunnel. This draft defines a mechanism to signal that capability using IS-IS. This mechanism is useful when the label advertisement is also done via IS-IS.

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 <http://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 June 18, 2014.

Copyright Notice

Copyright (c) 2013 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 (<http://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
1.1. Requirements Language	2
2. Terminology	2
3. Advertising ELC using IS-IS	2
4. Acknowledgements	3
5. IANA Considerations	3
6. Security Considerations	3
7. References	3
7.1. Normative References	3
7.2. Informative References	3
Authors' Addresses	4

1. Introduction

Multi Protocol Label Switching (MPLS) has defined a method in [RFC6790] to load balance traffic flows using Entropy Labels (EL). An ingress LSR cannot insert ELs for packets going into a given tunnel unless an egress LSR has indicated via signaling that it can process ELs on that tunnel. [RFC6790] defines the signaling of this capability (a.k.a Entropy Label Capability - ELC) via signaling protocols. Recently, mechanisms are being defined to signal labels via link state Interior Gateway Protocols (IGP) such as IS-IS [I-D.previdi-isis-segment-routing-extensions]. In such scenario the signaling mechanisms defined in [RFC6790] are inadequate. This draft defines a mechanism to signal the ELC using IS-IS. This mechanism is useful when the label advertisement is also done via IS-IS.

1.1. 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].

2. Terminology

This memo makes use of the terms defined in [RFC6790] and [RFC4971].

3. Advertising ELC using IS-IS

The IS-IS Router CAPABILITY TLV defined in [RFC4971] is used by IS-IS routers to announce their capabilities. A new sub-TLV of this TLV, called ELC sub-TLV is defined to advertise the capability of the router to process the ELs. It is formatted as described in [RFC5305] with a Type code to be assigned by IANA and a Length of zero. The scope of the advertisement depends on the application but it is recommended that it SHOULD be domain-wide.

4. Acknowledgements

The authors would like to thank Yimin Shen and George Swallow for their comments.

5. IANA Considerations

This memo includes a request to IANA to allocate a sub-TLV type within the IS-IS Router Capability TLV.

6. Security Considerations

This document does not introduce any new security considerations.

7. References

7.1. Normative References

- [I-D.previdi-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., and S. Litkowski, "IS-IS Extensions for Segment Routing", draft-previdi-isis-segment-routing-extensions-04 (work in progress), October 2013.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.
- [RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, November 2012.

7.2. Informative References

[I-D.filsfils-rtgwg-segment-routing]

Filsfils, C., Previdi, S., Bashandy, A., Decraene, B.,
Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R.,
Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe,
"Segment Routing Architecture", draft-filsfils-rtgwg-
segment-routing-01 (work in progress), October 2013.

Authors' Addresses

Xiaohu Xu
Huawei

Email: xuxiaohu@huawei.com

Sriganesh Kini
Ericsson

Email: sriganesh.kini@ericsson.com

Siva Sivabalan
Cisco

Email: msiva@cisco.com

Clarence Filsfils
Cisco

Email: cfilsfil@cisco.com