

Networking Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 6, 2014

L. Ginsberg
S. Previdi
Y. Yang
Cisco Systems
June 4, 2014

IS-IS Flooding Scope LSPs
draft-ietf-isis-fs-lsp-02.txt

Abstract

Intermediate System To Intermediate System (IS-IS) provides efficient and reliable flooding of information to its peers. However the current flooding scopes are limited to either area wide scope or domain wide scope. There are existing use cases where support of other flooding scopes are desirable. This document defines new Protocol Data Units (PDUs) which provide support for new flooding scopes as well as additional space for advertising information targeted for the currently supported flooding scopes. This document also defines extended TLVs and sub-TLVs which are encoded using 16 bit fields for type and length.

The protocol extensions defined in this document are not backwards compatible with existing implementations and so must be deployed with care.

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 December 6, 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.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	3
2. Extended TLVs	4
2.1. Use of Extended TLVs and Extended sub-TLVs	5
2.2. Use of Standard Code Points in Extended TLVs and Extended sub-TLVs	5
3. Definition of New PDUs	6
3.1. Flooding Scoped LSP Format	6
3.2. Flooding Scoped CSNP Format	9
3.3. Flooding Scope PSNP Format	10
4. Flooding Scope Update Process Operation	12
4.1. Scope Types	12
4.2. Operation on Point-to-Point Circuits	12
4.3. Operation on Broadcast Circuits	13
4.4. Use of Authentication	13
4.5. Priority Flooding	13
5. Deployment Considerations	14
6. Graceful Restart Interactions	14

7. Multi-instance Interactions	14
8. Circuit Scoped Flooding	14
9. Extending LSP Set Capacity	15
10. Domain Scoped Flooding	16
11. Announcing Support for Flooding Scopes	17
12. IANA Considerations	18
13. Security Considerations	19
14. Acknowledgements	20
15. References	20
15.1. Normative References	20
15.2. Informational References	21
Appendix A. Change History	21
Authors' Addresses	21

1. Introduction

The Update Process as defined by [IS-IS] provides reliable and efficient flooding of information to all routers in a given flooding scope. Currently the protocol supports two flooding scopes and associated Protocol Data Units (PDUs). Level 1 (L1) Link State PDUs (LSPs) are flooded to all routers in an area. Level 2 (L2) LSPs are flooded to all routers in the Level 2 sub-domain. The basic operation of the Update Process can be applied to any subset of the routers in a given topology so long as that topology is not partitioned. It is therefore possible to introduce new PDUs in support of other flooding scopes and utilize the same Update Process machinery to provide the same reliability and efficiency which the Update Process currently provides for L1 and L2 scopes. This document defines these new PDUs and the modified Update Process rules which are to be used in supporting new flooding scopes.

New deployment cases have introduced the need for reliable and efficient circuit scoped flooding. For example, Appointed Forwarder information as defined in [RFC7176] needs to be flooded reliably and efficiently to all R Bridges on a broadcast circuit. Currently, only Intermediate System to Intermediate System Hellos (IIHs) have the matching scope - but IIHs are unreliable i.e. individual IIHs may be lost without affecting correct operation of the protocol. To provide reliability in cases where the set of information to be flooded exceeds the carrying capacity of a single PDU requires sending the information periodically even when no changes in the content have occurred. When the information content is large this is inefficient and still does not provide a guarantee of reliability. This document defines circuit scoped flooding in order to provide a solution for such cases.

Another existing limitation of [IS-IS] is the carrying capacity of an LSP set. It has been noted in [RFC5311] that the set of LSPs that

may be originated by a system at each level is limited to 256 LSPs and the maximum size of each LSP is limited by the minimum Maximum Transmission Unit (MTU) of any link used to flood LSPs. [RFC5311] has defined a backwards compatible protocol extension which can be used to overcome this limitation if needed. While the [RFC5311] solution is viable, in order to be interoperable with routers which do not support the extension it imposes some restrictions on what can/cannot be advertised in the Extended LSPs and requires allocation of multiple unique system IDs to a given router. A more flexible and less constraining solution is possible if interoperability with legacy routers is not a requirement. As the introduction of new PDUs required to support new flooding scopes is by definition not interoperable with legacy routers, it is possible to simultaneously introduce an alternative solution to the limited LSP set carrying capacity of Level 1 and Level 2 LSPs as part of the extensions defined in this document. This capability is also defined in this document.

Standard IS-IS TLVs (Type/Length/Value) are encoded using an eight bit type and an 8 bit length. In cases where the set of information about a single object exceeds 255 octets multiple TLVs are required to encode all of the relevant information. This document introduces extended TLVs and extended sub-TLVs which use a 16 bit type field and a 16 bit length field.

The PDU type field in the common header for all IS-IS PDUs is a 5 bit field. The possible PDU types supported by the protocol are therefore limited to a maximum of 32. In order to minimize the need to introduce additional PDU types in the future, the new PDUs introduced in this document are defined so as to allow multiple flooding scopes to be associated with the same PDU type. This means if new flooding scopes are required in the future the same PDU type can be used.

2. Extended TLVs

Standard TLVs as defined in [IS-IS] as well as standard sub-TLVs (first introduced in [RFC5305]) have an eight bit type field and an eight bit length field. This constrains the information included in a single TLV or sub-TLV to 255 octets. With the increasing use of sub-TLVs it becomes more likely that the amount of information about a single object which needs to be advertised may exceed 255 octets. In such cases the information is encoded in multiple TLVs. This leads to less efficient encoding since the information which uniquely identifies the object must be repeated in each TLV and requires additional implementation complexity when receiving the information to ensure that all information about the object is correctly collected from the multiple TLVs.

This document introduces extended TLVs and extended sub-TLVs. These are encoded using a 16 bit type field and a 16 bit length field.

2.1. Use of Extended TLVs and Extended sub-TLVs

The following restrictions apply to the use of extended TLVs and extended sub-TLVs:

- o Extended TLVs and extended sub-TLVs are permitted only in Flooding Scoped PDUs which have a flooding scope designated for their use (defined later in this document)
- o A given flooding scope supports the use of either standard TLVs and standard sub-TLVs or the use of extended TLVs and extended sub-TLVs but not both
- o Extended TLVs and extended sub-TLVs MUST be used together i.e., using Standard sub-TLVs within an Extended TLV or using Extended sub-TLVs within a Standard TLV is invalid
- o If additional levels of TLVs (e.g., sub-sub-TLVs) are introduced in the future then the size of the type/length fields in these new sub-types MUST match the size used in the parent
- o The 16 bit type and length fields are encoded in network byte order
- o Use of extended TLVs and extended sub-TLVs does not alter in any way the maximum size of PDUs which may sent or received

2.2. Use of Standard Code Points in Extended TLVs and Extended sub-TLVs

Standard TLV and standard sub-TLV code points as defined in the IANA IS-IS TLV Codepoints Registry MAY be used in extended TLVs and extended sub-TLVs. Encoding is as specified for each of the standard TLVs and standard sub-TLVs with the following differences:

- o The eight bit type is encoded as an unsigned 16 bit integer where the 8 MSBs are all 0
- o The eight bit length field is replaced by the 16 bit length field
- o The length MAY take on values greater than 255

3. Definition of New PDUs

In support of new flooding scopes the following new PDUs are required:

- o Flooding Scoped LSPs (FS-LSPs)
- o Flooding Scoped Complete Sequence Number PDUs (FS-CSNPs)
- o Flooding Scoped Partial Sequence Number PDUs (FS-PSNPs)

Each of these PDUs is intentionally defined with a header as similar in format as possible to the corresponding PDU types currently defined in [IS-IS]. Although it might have been possible to eliminate or redefine PDU header fields in a new way the existing formats are retained in order to allow maximum reuse of existing PDU processing logic in an implementation.

Note that in the case of all FS PDUs, the Maximum Area Addresses field in the header of the corresponding standard PDU has been replaced with a Scope field. The maximum area addresses checks specified in [IS-IS] are therefore not performed on FS PDUs.

3.1. Flooding Scoped LSP Format

An FS-LSP has the following format:

	No. of octets
+-----+	
Intradomain Routeing Protocol Discriminator	1
+-----+	
Length Indicator	1
+-----+	
Version/Protocol ID Extension	1
+-----+	
ID Length	1
+-----+	
R R R PDU Type	1
+-----+	
Version	1
+-----+	
Reserved	1
+-----+	
P Scope	1
+-----+	
PDU Length	2

+-----+		
Remaining Lifetime		2
+-----+		
FS LSP ID		ID Length + 2
+-----+		
Sequence Number		4
+-----+		
Checksum		2
+-----+		
Reserved LSPDBOL IS Type		1
+-----+		
: Variable Length Fields :		Variable
+-----+		

Intradomain Routing Protocol Discriminator - 0x83
(as defined in [IS-IS])

Length Indicator - Length of the Fixed Header in octets

Version/Protocol ID Extension - 1

ID Length - As defined in [IS-IS]

PDU Type - 10 (Subject to assignment by IANA) Format as defined in [IS-IS]

Version - 1

Reserved - transmitted as zero, ignored on receipt

Scope - Bits 1-7 define the flooding scope.
The value 0 is reserved and MUST NOT be used. Received FS-LSPs with a scope of 0 MUST be ignored and MUST NOT be flooded.
P - Bit 8 - Priority Bit. If set to 1 this LSP SHOULD be flooded at high priority.
Scopes (1 - 63) are reserved for use with standard TLVs and standard sub-TLVs.
Scopes (64 - 127) are reserved for use with extended TLV and extended sub-TLVs.

PDU Length - Entire Length of this PDU, in octets, including the header.

Remaining Lifetime - Number of seconds before this FS-LSP is considered expired.

FS LSP ID - the system ID of the source of the FS-LSP. One of

the following two formats is used:

FS LSP ID Standard Format

+-----+	
Source ID	ID Length
+-----+	
Pseudonode ID	1
+-----+	
FS LSP Number	1
+-----+	

FS LSP ID Extended Format

+-----+	
Source ID	ID Length
+-----+	
Extended FS LSP Number	2
+-----+	

Which format is used is specific to the Scope and MUST be defined when the specific flooding scope is defined.

Sequence Number - sequence number of this FS-LSP

Checksum - Checksum of contents of FS-LSP from Source ID to end.
Checksum is computed as defined in [IS-IS].

Reserved/LSPDBOL/IS Type

Bits 4-8 are reserved, which means they are transmitted as 0 and ignored on receipt.

LSPDBOL - Bit 3 - A value of 0 indicates no FS-LSP Database Overload and a value of 1 indicates that the FS-LSP Database is overloaded. The overload condition is specific to FS-LSPs with the scope specified in the scope field.

IS Type - Bits 1 and 2. The type of Intermediate System as defined in [IS-IS].

Variable Length Fields which are allowed in an FS-LSP are specific to

the defined scope.

3.2. Flooding Scoped CSNP Format

An FS-CSNP has the following format:

	No. of octets
+-----+ Intradomain Routeing Protocol Discriminator	1
+-----+ Length Indicator	1
+-----+ Version/Protocol ID Extension	1
+-----+ ID Length	1
+-----+ R R R PDU Type	1
+-----+ Version	1
+-----+ Reserved	1
+-----+ R Scope	1
+-----+ PDU Length	2
+-----+ Source ID	ID Length + 1
+-----+ Start FS-LSP ID	ID Length + 2
+-----+ End FS-LSP ID	ID Length + 2
+-----+ : Variable Length Fields :	Variable
+-----+	

Intradomain Routeing Protocol Discriminator - 0x83
(as defined in [IS-IS])

Length Indicator - Length of the Fixed Header in octets

Version/Protocol ID Extension - 1

ID Length - As defined in [IS-IS]

PDU Type - 11 (Subject to assignment by IANA) Format as defined in [IS-IS]

Version - 1

Reserved - transmitted as zero, ignored on receipt

Scope - Bits 1-7 define the flooding scope.

The value 0 is reserved and MUST NOT be used. Received FS-CSNPs with a scope of 0 MUST be ignored.

Bit 8 is Reserved which means it is transmitted as 0 and ignored on receipt.

Scopes (1 - 63) are reserved for use with standard TLVs and standard sub-TLVs.

Scopes (64 - 127) are reserved for use with extended TLV and extended sub-TLVs.

PDU Length - Entire Length of this PDU, in octets, including the header.

Source ID - the system ID of the Intermediate System (with zero Circuit ID) generating this Sequence Numbers PDU

Start FS-LSP ID - The FS-LSP ID of the first FS-LSP with the specified scope in the range covered by this FS-CSNP.

End FS-LSP ID - The FS-LSP ID of the last FS-LSP with the specified scope in the range covered by this FS-CSNP.

Variable Length Fields which are allowed in an FS-CSNP are limited to those TLVs which are supported by standard CSNP.

3.3. Flooding Scope PSNP Format

An FS-PSNP has the following format:

	No. of octets
+-----+	
Intradomain Routeing	1
Protocol Discriminator	
+-----+	
Length Indicator	1
+-----+	
Version/Protocol ID	1
Extension	
+-----+	

ID Length		1
+-----+		
R R R PDU Type		1
+-----+		
Version		1
+-----+		
Reserved		1
+-----+		
U Scope		1
+-----+		
PDU Length		2
+-----+		
Source ID		ID Length + 1
+-----+		
: Variable Length Fields :		Variable
+-----+		

Intradomain Routing Protocol Discriminator - 0x83
(as defined in [IS-IS])

Length Indicator - Length of the Fixed Header in octets

Version/Protocol ID Extension - 1

ID Length - As defined in [IS-IS]

PDU Type - 12 (Subject to assignment by IANA) Format
as defined in [IS-IS]

Version - 1

Reserved - transmitted as zero, ignored on receipt

Scope - Bits 1-7 define the flooding scope.

The value 0 is reserved and MUST NOT be used. Received
FS-PSNPs with a scope of 0 MUST be ignored.

U - Bit 8 - A value of 0 indicates that the specified
flooding scope is supported. A value of 1 indicates
that the specified flooding scope is unsupported. When
U = 1, variable length fields other than authentication
MUST NOT be included in the PDU.

Scopes (1 - 63) are reserved for use with standard TLVs and
standard sub-TLVs.

Scopes (64 - 127) are reserved for use with extended TLV and
extended sub-TLVs.

PDU Length - Entire Length of this PDU, in octets, including
the header.

Source ID - the system ID of the Intermediate System
(with zero Circuit ID) generating this Sequence Numbers PDU

Variable Length Fields which are allowed in an FS-PSNP are
limited to those TLVs which are supported by standard PSNPs.

4. Flooding Scope Update Process Operation

The Update Process as defined in [IS-IS] maintains a Link State Database (LSDB) for each level supported. Each level specific LSDB contains the full set of LSPs generated by all routers operating in that level specific scope. The introduction of FS-LSPs creates additional LSDBs (FS-LSDBs) for each additional scope supported. The set of FS-LSPs in each FS-LSDB consists of all FS-LSPs generated by all routers operating in that scope. There is therefore an additional instance of the Update Process for each supported flooding scope.

Operation of the scope specific Update Process follows the Update Process specification in [IS-IS]. The circuit(s) on which FS-LSPs are flooded are limited to those circuits which are participating in the given scope. Similarly the sending/receiving of FS-CSNPs and FS-PSNPs is limited to the circuits participating in the given scope.

Consistent support of a given flooding scope on a circuit by all routers operating on that circuit is required.

4.1. Scope Types

A flooding scope may be limited to a single circuit (circuit scope). Circuit scopes may be further limited by level (L1 circuit scope/L2 circuit scope).

A flooding scope may be limited to all circuits enabled for L1 routing (area scope).

A flooding scope may be limited to all circuits enabled for L2 routing (L2 sub-domain scope).

Additional scopes may be defined which include all circuits enabled for either L1 or L2 routing (domain-wide scope).

4.2. Operation on Point-to-Point Circuits

When a new adjacency is formed, synchronization of all FS-LSDBs supported on that circuit is required. Therefore FS-CSNPs for all supported scopes MUST be sent when a new adjacency reaches the UP

state. Send Receive Message (SRM) bit MUST be set for all FS-LSPs associated with the scopes supported on that circuit. Receipt of an FS-PSNP with the U bit equal to 1 indicates that the neighbor does not support that scope (although it does support FS PDUs). This MUST cause SRM bit to be cleared for all FS-LSPs with the matching scope which are currently marked for flooding on that circuit.

4.3. Operation on Broadcast Circuits

FS PDUs are sent to the same destination address(es) as standard PDUs for the given protocol instance. For specification of the defined destination addresses consult [IS-IS], [IEEEaq], [RFC6822], and [RFC6325].

The Designated Intermediate System (DIS) for a broadcast circuit has the responsibility to generate periodic scope specific FS-CSNPs for all supported scopes. A scope specific DIS is NOT elected as all routers on a circuit MUST support a consistent set of flooding scopes.

It is possible that a scope may be defined which is not level specific. In such a case the DIS for each level enabled on a broadcast circuit MUST independently send FS PDUs for that scope to the appropriate level specific destination address. This may result in redundant flooding of FS-LSPs for that scope.

4.4. Use of Authentication

Authentication TLVs MAY be included in FS PDUs. When authentication is in use, the scope is first used to select the authentication configuration that is applicable. The authentication check is then performed as normal. Although scope specific authentication MAY be used, sharing of authentication among multiple scopes and/or with the standard LSP/CSNP/PSNP PDUs is considered sufficient.

4.5. Priority Flooding

When the FS LSP ID Extended Format is used the set of LSPs generated by an IS may be quite large. It may be useful to identify those LSPs in the set which contain information of higher priority. Such LSPs will have the P bit set to 1 in the Scope field in the LSP header. Such LSPs SHOULD be flooded at a higher priority than LSPs with the P bit set to 0. This is a suggested behavior on the part of the originator of the LSP. When an LSP is purged the original state of the P bit MUST be preserved.

5. Deployment Considerations

Introduction of new PDU types is incompatible with legacy implementations. Legacy implementations do not support the FS specific Update process(es) and therefore flooding of the FS-LSPs throughout the defined scope is unreliable when not all routers in the defined scope support FS PDUs. Further, legacy implementations will likely treat the reception of an FS PDUs as an error. Even when all routers in a given scope support FS PDUs, if not all routers in the flooding domain for a given scope support that scope, then flooding of the FS-LSPs may be compromised. Therefore all routers in the flooding domain for a given scope SHOULD support both FS PDUs and the specified scope before use of that scope can be enabled.

The U bit in FS-PSNPs provides a means to suppress retransmissions of unsupported scopes. Routers which support FS PDUs SHOULD support the sending of PSNPs with the U bit equal to 1 when an FS-LSP is received with a scope which is unsupported. Routers which support FS PDUs SHOULD trigger management notifications when FS PDUs are received for unsupported scopes and when PSNPs with the Ubit equal to 1 are received.

6. Graceful Restart Interactions

[RFC5306] defines protocol extensions in support of graceful restart of a routing instance. Synchronization of all supported FS-LSDBs is required in order for database synchronization to be complete. This involves the use of additional T2 timers. Receipt of a PSNP with the U bit equal to 1 will cause FS-LSDB synchronization with that neighbor to be considered complete for that scope. See [RFC5306] for further details.

7. Multi-instance Interactions

In cases where FS-PDUs are associated with a non-zero instance the use of IID-TLVs in FS-PDUs follows the rules for use in LSPs, CSNPs, PSNPs as defined in [RFC6822].

8. Circuit Scoped Flooding

This document defines four circuit scoped flooding identifiers:

- o Level 1 circuit scope (L1CS) - this uses standard TLVs and standard sub-TLVs
- o Level 2 circuit scope (L2CS) - this uses standard TLVs and standard sub-TLVs

- o Extended Level 1 circuit scope (E-L1CS) - this uses extended TLVs and extended sub-TLVs
- o Extended Level 2 circuit scope (E-L2CS) - this uses extended TLVs and extended sub-TLVs

FS-LSPs with the scope field set to one of these values contain information specific to the circuit on which they are flooded. When received, such FS-LSPs MUST NOT be flooded on any other circuit. The FS LSP ID Extended format is used in these PDUs. The FS-LSDB associated with circuit scoped FS-LSPs consists of the set of FS-LSPs which both have matching circuit scope and are transmitted (locally generated) or received on a specific circuit.

The set of TLVs which may be included in such FS-LSPs is specific to the given use case and is outside the scope of this document.

9. Extending LSP Set Capacity

The need for additional space in the set of LSPs generated by a single IS has been articulated in [RFC5311]. When legacy interoperability is not a requirement, the use of FS-LSPs meets that need without requiring the assignment of alias system-ids to a single IS. Four flooding scopes are defined for this purpose:

- o Level 1 Scope (L1FS) - this uses standard TLVs and standard sub-TLVs
- o Level 2 Scope (L2FS) - this uses standard TLVs and standard sub-TLVs
- o Extended Level 1 Scope (E-L1FS) - this uses extended TLVs and extended sub-TLVs
- o Extended Level 2 Scope (E-L2FS) - this uses extended TLVs and extended sub-TLVs

L1FS and E-L1FS LSPs are flooded on all L1 circuits. L2FS and E-L2FS LSPs are flooded on all L2 circuits.

The FS LSP ID Extended format is used in these PDUs. This provides 64K of additional LSPs which may be generated by a single system at each level.

LxFS LSPs are used by the level specific Decision Process (defined in [IS-IS]) in the same manner as standard LSPs (i.e. as additional information sourced by the same IS) subject to the following restrictions:

- o A valid version of standard LSP #0 from the same IS at the corresponding Level MUST be present in the LSDB in order for the LxFS set to be usable
- o Information in an LxFS LSP (e.g. IS-Neighbor information) which supports using the originating IS as a transit node MUST NOT be used when the Overload bit is set in the corresponding standard LSP #0
- o TLVs which are restricted to standard LSP #0 MUST NOT appear in LxFS LSPs.

There are no further restrictions as to what TLVs may be advertised in FS-LSPs.

10. Domain Scoped Flooding

Existing support for flooding information domain wide (i.e. to L1 routers in all areas as well as to routers in the Level 2 sub-domain) requires the use of leaking procedures between levels. For further details see [RFC4971]. This is sufficient when the data being flooded domain-wide consists of individual TLVs. If it is desired to retain the identity of the originating IS for the complete contents of a PDU, then support for flooding the unchanged PDU is desirable. This document therefore defines two flooding scopes in support of domain-wide flooding. FS-LSPs with this scope MUST be flooded on all circuits regardless of what level(s) are supported on that circuit.

- o Domain Scope (DSFS) - this uses standard TLVs and standard sub-TLVs
- o Extended Domain Scope (E-DSFS) - this uses extended TLVs and extended sub-TLVs

The FS LSP ID Extended format is used in these PDUs.

Use of information in FS-LSPs for a given scope depends on determining the reachability to the IS originating the FS-LSP. This presents challenges for FS-LSPs with domain-scopes because no single IS has the full view of the topology across all areas. It is therefore necessary for the originator of domain scoped DSFS and E-DSFS LSPs to advertise an identifier which will allow an IS who receives such an FS-LSP to determine whether the source of the FS-LSP is currently reachable. The identifier required depends on what "address-families" are being advertised.

When IS-IS is deployed in support of Layer 3 routing for IPv4 and/or IPv6 then FS-LSP #0 with domain-wide scope MUST include at least one of the following TLVs:

- o IPv4 Traffic Engineering Router ID (TLV 134)
- o IPv6 Traffic Engineering Router ID (TLV 140)

When IS-IS is deployed in support of Layer 2 routing, current standards (e.g. [RFC6325]) only support a single area. Therefore domain-wide scope is not yet applicable. When the Layer 2 standards are updated to include multi-area support the identifiers which can be used to support inter-area reachability will be defined - at which point the use of domain-wide scope for Layer 2 can be fully defined.

11. Announcing Support for Flooding Scopes

Announcements of support for flooding scope may be useful in validating that full support has been deployed and/or in isolating the reasons for incomplete flooding of FS-LSPs for a given scope.

ISs supporting FS-PDUs MAY announce supported scopes in IIH PDUs. To do so a new TLV is defined.

Scoped Flooding Support

Type: 243 (suggested - to be assigned by IANA)

Length: 1 - 127

Value

	No of octets
+-----+ R Supported Scope	1
+-----+	
: :	
+-----+	
R Supported Scope	1
+-----+	

A list of the circuit scopes supported on this circuit and other non-circuit flooding scopes supported.
R bit MUST be 0 and is ignored on receipt.

In a Point-Point IIH L1, L2, domain-wide, and all circuit scopes MAY be advertised.

In Level 1 LAN IIHs L1, domain-wide, and L1 circuit scopes MAY be advertised. L2 scopes and L2 circuit scopes MUST NOT be advertised.

In Level 2 LAN IIHs L2, domain-wide, and L2 circuit scopes MAY be advertised. L1 scopes and L1 circuit scopes MUST NOT be advertised.

Information in this TLV MUST NOT be considered in adjacency formation.

Whether information in this TLV is used to determine when FS-LSPs associated with a locally supported scope are flooded is an implementation choice.

12. IANA Considerations

This document requires the definition of three new PDU types that need to be reflected in the ISIS PDU registry. Values below are suggested values subject to assignment by IANA.

Value	Description
-----	-----
10	FS-LSP
11	FS-CSNP
12	FS-PSNP

This document requires that a new IANA registry be created to control the assignment of scope identifiers in FS-PDUs. The registration

procedure is "Expert Review" as defined in [RFC5226]. Suggested registry name is "LSP Flooding Scoped Identifier Registry". A scope identifier is a number from 1-127 inclusive. Values 1 - 63 are reserved for PDUs which use standard TLVs and standard sub-TLVs. Values 64 - 127 are reserved for PDUs which use extended TLVs and extended sub-TLVs. The list of hello PDUs in which support for a given scope MAY be announced (using Scope Flooding Support TLV) is specified for each defined scope.

The following scope identifiers are defined by this document. Values are suggested values subject to assignment by IANA.

Value	Description	FS LSP ID Format/ TLV Format	IIH Announce P2P L1LAN L2LAN
1	Level 1 Circuit Flooding Scope	Extended/Standard	Y Y N
2	Level 2 Circuit Flooding Scope	Extended/Standard	Y N Y
3	Level 1 Flooding Scope	Extended/Standard	Y Y N
4	Level 2 Flooding Scope	Extended/Standard	Y N Y
5	Domain-wide Flooding Scope	Extended/Standard	Y Y Y
(6-63) Unassigned			
64	Level 1 Circuit Flooding Scope	Extended/Extended	Y Y N
65	Level 2 Circuit Flooding Scope	Extended/Extended	Y N Y
66	Level 1 Flooding Scope	Extended/Extended	Y Y N
67	Level 2 Flooding Scope	Extended/Extended	Y N Y
68	Domain-wide Flooding Scope	Extended/Extended	Y Y Y
(69-127) Unassigned			

This document requires the definition of a new IS-IS TLV to be reflected in the "IS-IS TLV Codepoints" registry:

Type	Description	IIH	LSP	SNP	Purge
243	Circuit Scoped Flooding Support	Y	N	N	N

The IANA TLV codepoints registry is extended to allow definition of codepoints less than or equal to 65535. Codepoints greater than 255 can only be used in PDUs designated to support extended TLVs.

13. Security Considerations

Security concerns for IS-IS are addressed in [IS-IS], [RFC5304], and [RFC5310].

The new PDUs introduced are subject to the same security issues associated with their standard LSP/CSNP/PSNP counterparts. To the extent that additional PDUs represent additional load for routers in the network this increases the opportunity for denial of service attacks.

14. Acknowledgements

The authors wish to thank Ayan Banerjee, Donald Eastlake, Hannes Gredler, and Mike Shand for their comments.

15. References

15.1. Normative References

- [IEEE802.1aq] "Standard for Local and metropolitan area networks: Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks - Amendment 20: Shortest Path Bridging", IEEE Std 802.1aq-2012, 29 June 2012.", 2012.
- [IS-IS] "Intermediate system to Intermediate system intra-domain routing 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.
- [RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [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.
- [RFC5306] Shand, M. and L. Ginsberg, "Restart Signaling for IS-IS", RFC 5306, 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.
- [RFC6822] Previdi, S., Ginsberg, L., Shand, M., Roy, A., and D. Ward, "IS-IS Multi-Instance", RFC 6822, December 2012.

15.2. Informational References

- [RFC5311] McPherson, D., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, February 2009.
- [RFC6325] Perlman, R., Eastlake, D., Dutt, D., Gai, S., and A. Ghanwani, "Routing Bridges (RBridges): Base Protocol Specification", RFC 6325, July 2011.
- [RFC7176] Eastlake, D., Senevirathne, T., Ghanwani, A., Dutt, D., and A. Banerjee, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 7176, May 2014.

Appendix A. Change History

Changes from 01 to 02 version

- o Updated Section 11 to state what scopes MUST NOT be announced in a given IIH PDU
- o Updated IANA section for new "LSP Flooding Scoped Identifier Registry" to include the hello PDUs in which a given scope may be announced.

Authors' Addresses

Les Ginsberg
Cisco Systems
510 McCarthy Blvd.
Milpitas, CA 95035
USA

Email: ginsberg@cisco.com

Stefano Previdi
Cisco Systems
Via Del Serafico 200
Rome 0144
Italy

Email: sprevidi@cisco.com

Yi Yang
Cisco Systems
7100-9 Kit Creek Road
Research Triangle Park, North Carolina 27709-4987
USA

Email: yiya@cisco.com

Networking Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 30, 2014

L. Ginsberg
S. Mirtorabi
S. Previdi
A. Roy
Cisco Systems
June 28, 2014

IS-IS Support for Unidirectional Links
draft-ietf-isis-udl-02.txt

Abstract

This document defines support for the operation of IS-IS over Unidirectional Links without the use of tunnels or encapsulation of IS-IS Protocol Data Units. Adjacency establishment when the return path from the router at the receive end of a unidirectional link to the router at the transmit end of the unidirectional link is via another unidirectional link is supported. The extensions defined here are backwards compatible - only the routers directly connected to a unidirectional link need to be upgraded.

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 December 30, 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.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	3
2. Encoding Extensions	3
2.1. UDL LSPs and the UDL-TLV	4
2.2. UDL Intermediate System Neighbors sub-TLV	4
2.2.1. UDL Point-to-Point Intermediate System Neighbor Sub-TLV	5
2.2.2. UDL LAN Intermediate System Neighbor Sub-TLV	5
2.2.3. Sub-TLVs Associated w an IS Neighbor	6
2.3. UDL Manual Area Addresses sub-TLV	9
3. Adjacency Establishment	10
3.1. Adjacency Establishment in Point-to-Point Mode	10
3.2. Adjacency Establishment in Broadcast Mode	11
3.3. UDL link metric configuration	12
4. Adjacency Maintenance	12
4.1. Adjacency Maintenance by IS-T	12
4.2. Adjacency Maintenance by IS-R	13
4.3. Use of BFD	14
4.4. Graceful Restart Support	14
5. Operation of the Update Process on a UDL	14

6. Support for UDL on the Return Path	15
7. IANA Considerations	16
8. Security Considerations	17
9. Acknowledgements	17
10. References	17
10.1. Normative References	17
10.2. Informational References	18
Authors' Addresses	19

1. Introduction

Operation of IS-IS depends upon two-way connectivity. Adjacencies are formed by exchanging hellos on a link, flooding of the link state database is made reliable by exchanges between neighbors on a link, etc. However, there are deployments where operation of the protocol is desired over links which are unidirectional i.e., one end of the link can only send Protocol Data Units (PDUs) and one end of the link can only receive PDUs. Traditional methods of supporting Unidirectional Links (UDLs) have involved establishing a tunnel from the Intermediate System (IS) at the receive end of the UDL to the IS at the transmit end of the UDL, encapsulating/decapsulating the IS-IS PDUs as they enter/exit the tunnel, and associating the PDUs received via the tunnel with the UDL at the transmit end. This typically requires static configuration and may introduce Maximum Transmission Unit (MTU) issues due to the required encapsulation.

This specification defines extensions to the protocol which support correct and reliable operation of IS-IS over UDLs without the need for tunnels or any form of encapsulation.

2. Encoding Extensions

Although the IS at the transmit end of a UDL link (IS-T) can send IS-IS PDUs normally on the link, the IS at the receive end of a UDL link (IS-R) requires assistance from other ISs in the network to pass the information it would normally send directly to IS-T. The Update Process as defined in [IS-IS] allows information generated by one IS in the network to be reliably flooded to all other ISs in the network using Link State PDUs (LSPs). The extensions defined here utilize LSPs to allow IS-R to send information normally sent in hellos (IIHs) or sequence number PDUs (SNPs) to IS-T in LSPs. As LSPs are flooded to all ISs in an area/sub-domain, care is taken to minimize the LSP churn necessary to support adjacency establishment and maintenance between IS-T and IS-R.

2.1. UDL LSPs and the UDL-TLV

Routers on the receive end of a UDL MUST reserve at least one LSP (for each level supported on the UDL) to advertise the UDL information described below. Such LSPs are referred to as UDL-LSPs although the only distinction between a UDL-LSP and other LSPs is in the TLV information which is present in such an LSP. LSP #0 MUST NOT be used to send UDL information. UDL-LSPs have the following special characteristics:

1. The only TLV which may be advertised in UDL-LSPs is the UDL TLV described below and (optionally) an Authentication TLV and/or Purge Originator Identification TLV [RFC6232]. This requirement is enforced by the originator of the UDL-LSP but is not checked by receiving systems i.e., other TLVs which are included in a UDL-LSP are processed normally. The reason for the restriction is to minimize the number of LSPs which have UDL information content.
2. Routers on the transmit side of a UDL flood UDL-LSPs regardless of the existence of an adjacency in the UP state on that circuit. Flooding of UDL-LSPs on circuits other than a UDL is as specified in [IS-IS] i.e., no special handling.

A new TLV is defined in which UDL specific information appears. All information in a UDL-TLV is encoded in sub-TLVs. UDL sub-TLVs are formatted as specified in [RFC5305]. The format of the UDL-TLV is therefore:

	No. of octets
+-----+ Type (11) (To be assigned by IANA) +-----+	1
+-----+ Length +-----+	1
+-----+ Sub-TLVs : +-----+	3 - 255

2.2. UDL Intermediate System Neighbors sub-TLV

UDL links may operate in Point-to-Point mode or in broadcast mode (assuming the subnetwork is a broadcast subnetwork). There are therefore two types of Intermediate System Neighbors sub-TLVs defined. A UDL-TLV MUST NOT contain more than one Intermediate

System Neighbors sub-TLV. If multiple Intermediate System Neighbors sub-TLVs appear in a UDL-TLV all information in that UDL-TLV MUST be ignored.

2.2.1. UDL Point-to-Point Intermediate System Neighbor Sub-TLV

The UDL Point-to-Point Intermediate System Neighbor Sub-TLV describes an adjacency on a UDL which is operating in Point-to-Point mode i.e. either a Point-to-Point subnetwork or a LAN subnetwork operating in Point-to-Point mode as described in [RFC5309]. The information encoded follows the format for the Point-to-Point Three-Way Adjacency TLV as defined in [RFC5303] but may also include the local LAN address when the underlying subnetwork is a LAN.

	No. of octets
+-----+ Type (240) (To be assigned by IANA) +-----+	1
+-----+ Length (9 + ID Length) to (15 + ID Length) +-----+	1
+-----+ Adjacency 3-way state +-----+	1
+-----+ Extended Local Circuit ID +-----+	4
+-----+ Neighbor System ID +-----+	ID Length
+-----+ Neighbor Extended Local Circuit ID +-----+	4
+-----+ Local LAN Address +-----+	6

2.2.2. UDL LAN Intermediate System Neighbor Sub-TLV

The UDL LAN Intermediate System Neighbor sub-TLV describes an adjacency on a UDL operating in broadcast mode on a LAN subnetwork.

	No. of octets
+-----+	
Type (6)	1
(To be assigned by IANA)	
+-----+	
Length (7 + ID Length)	1
+-----+	
Neighbor LAN ID	ID Length + 1
+-----+	
Local LAN Address	6
+-----+	

2.2.3. Sub-TLVs Associated w an IS Neighbor

A number of sub-TLVs require the presence of a UDL IS-Neighbor sub-TLV (either Point-to-Point or LAN) in the UDL-TLV in order to provide appropriate context for the information being advertised. These sub-TLVs are described in the sub-sections below.

2.2.3.1. UDL LSP Range sub-TLV

The content of this sub-TLV describes a range of LSPs for which the originating router requires an update. Only the neighbor specified in the associated UDL IS-Neighbor sub-TLV processes the LSP range mentioned in this sub-TLV.

	No. of octets
+-----+	
Type (8)	1
(To be assigned by IANA)	
+-----+	
Length (ID Length + 2)* 2	1
+-----+	
Start LSP ID	ID Length + 2
+-----+	
End LSP ID	ID Length + 2
+-----+	

2.2.3.2. UDL LSP Entry sub-TLV

The content of this sub-TLV describes LSPs for which the originating router requires an update. Only the neighbor specified in the associated UDL IS-Neighbor sub-TLV processes the LSP entries specified in this sub-TLV.

	No. of octets
+-----+	
Type (9)	1
(To be assigned by IANA)	
+-----+	
Length (10 + ID Length)*N	1
+-----+	
: LSP Entries	:
+-----+	

Each LSP Entry has the following format:

+-----+	
Remaining Lifetime	2
+-----+	
LSP ID	ID Length + 2
+-----+	
LSP Sequence Number	4
+-----+	
Checksum	2
+-----+	

2.2.3.3. Protocols Supported sub-TLV

This sub-TLV specifies the set of Network Layer Protocol Identifiers (NLPIDs) that the originating system is capable of forwarding as defined in [RFC1195].

	No. of octets
+-----+	
Type (129)	1
(To be assigned by IANA)	
+-----+	
Length	Number of NLPIDs
+-----+	
: NLPIDs	: 1 octet/NLPID
+-----+	

2.2.3.4. IP Address sub-TLV

This sub-TLV specifies the set of IP addresses configured on the interface as defined in [RFC1195].

		No. of octets
+-----+		
Type (132)		1
(To be assigned by IANA)		
+-----+		
Length		4 * # of addresses
+-----+		
: IP Address(es)	:	4 octets/address
+-----+		

2.2.3.5. Multi-Topology sub-TLV

This sub-TLV specifies the set of topology identifiers supported as defined in [RFC5120].

		No. of octets
+-----+		
Type (229)		1
(To be assigned by IANA)		
+-----+		
Length		2 * # of MTIDs
+-----+		
: MTIDs	:	2 octets/MTID
+-----+		

NOTE: All flag bits defined in [RFC5120] MUST be transmitted as 0 and ignored on receipt.

2.2.3.6. IPv6 Interface Address sub-TLV

This sub-TLV specifies the set of IPv6 addresses assigned on the local interface as defined in [RFC5308]. Addresses MUST be link local addresses.

		No. of octets
+-----+		
Type (232)		1
(To be assigned by IANA)		
+-----+		
Length		16 * # of IPv6 addresses
+-----+		
: IPv6 Addresses	:	16 octets/Address
+-----+		

2.2.3.7. IPv6 Global Interface Address sub-TLV

This sub-TLV specifies the set of global IPv6 addresses assigned on the local interface as defined in [RFC6119]. . Addresses MUST be global or unique local addresses.

No. of octets	
+-----+	
Type (233)	1
(To be assigned by IANA)	
+-----+	
Length	16 * # of IPv6 addresses
+-----+	
: IPv6 Addresses	: 16 octets/Address
+-----+	

2.3. UDL Manual Area Addresses sub-TLV

This sub-TLV specifies the set of manualAreaAddresses of the originating system. No other sub-TLVs are allowed in a UDL-TLV which has this sub-TLV. Any other sub-TLVs in such a UDL-TLV are ignored on receipt.

No. of octets	
+-----+	
Type (1)	1
(To be assigned by IANA)	
+-----+	
Length	1
+-----+	
: Area Address(es)	:
+-----+	

Each Area Address has the following format:

+-----+	
Address Length	1
+-----+	
Area Address	Address Length
+-----+	

3. Adjacency Establishment

An adjacency over a UDL link may be established over a link operating in Point-to-Point mode (including a LAN subnetwork configured to operate in Point-to-Point mode) or a link operating in broadcast mode. Operation in either mode is identical except for some differences in the manner of adjacency establishment as specified in the following sub-sections.

IS-T utilizes the set of manualAreaAddresses advertised by IS-R in a UDL Manual Area Address sub-TLV in combination with the UDL Intermediate System Neighbor sub-TLV(s) to IS-T advertised by IS-R to determine the level(s) associated with any adjacency to IS-R.

3.1. Adjacency Establishment in Point-to-Point Mode

Adjacency establishment makes use of Three Way Handshake as defined in [RFC5303] when operating in Point-to-Point mode. When operating over a LAN subnetwork, the use of point-to-point operation over LAN as defined in [RFC5309] is also used.

IS-T initiates adjacency establishment by sending Point-to-Point IIHs over the UDL as normal i.e., including Three-Way Handshake TLV. Note that the local circuit ID specified by IS-T need only be unique among the set of Point-to-Point UDL links supported by IS-T on which IS-T is at the transmit end.

Upon receipt of a Point-to-Point IIH IS-R creates an adjacency in the INIT state with IS-T and advertises the existence of the adjacency in its UDL-LSP(s) utilizing the UDL Point-to-Point Intermediate System Neighbor sub-TLV. The Local LAN address is included if the link is a LAN subnetwork operating in Point-to-Point mode. UDL-LSPs of the appropriate level(s) are generated according to the type of the adjacency with IS-T.

When IS-T receives the UDL-LSP(s) generated by IS-R containing the UDL Point-to-Point Intermediate System Neighbor sub-TLV it validates the 3 way information and, if valid, transitions its adjacency to UP state. In subsequent Point-to-Point IIHs IS-T includes IS-R's circuit ID information as indicated in the UDL Point-to-Point IS Neighbor sub-TLV in its 3 way handshake TLV. A complete set of CSNPs is sent to IS-R for the level(s) appropriate for the type of adjacency. LSPs which are updated as a result of the existence of the adjacency to IS-R are sent to IS-R, but IS-T does NOT propagate its full LSP Database. This is done to minimize the amount of redundant flooding.

IS-R uses normal adjacency bring up rules based on the 3 way handshake information it receives in Point-to-Point IIHs from IS-T and advertises its IS neighbor to IS-T in the usual manner i.e. in an LSP other than a UDL-LSP. Following transition of the adjacency to IS-T to the UP state IS-R MAY request IS-T to flood its complete LSP Database by sending an LSP Range sub-TLV to IS-T in a UDL-LSP.

3.2. Adjacency Establishment in Broadcast Mode

IS-T initiates adjacency establishment by sending LAN IIHs of the appropriate level(s) over the UDL as normal. IS-T specifies itself in the LAN ID field of the IIH, including a non-zero circuit ID. Note that the local circuit ID specified by IS-T need only be unique among the set of LAN UDL links supported by IS-T on which IS-T is at the transmit end. This is because pseudo-node LSPs will never be generated for a UDL. Operation in broadcast mode supports a UDL with a single IS-T and multiple IS-Rs.

Upon receipt of a LAN IIH PDU IS-R creates an adjacency in the INIT state with IS-T and advertises the existence of the adjacency in its UDL-LSP(s) utilizing the UDL LAN Intermediate System Neighbor sub-TLV. UDL-LSPs of the appropriate level(s) are generated according to the levels supported by IS-R and IS-T.

When IS-T receives the UDL-LSP(s) generated by IS-R containing the UDL LAN Intermediate System Neighbor sub-TLV(s) it validates the LANID and, if valid, transitions its adjacency to UP state. In subsequent LAN IIH PDUs, IS-T includes IS-R's LAN Address as indicated in the UDL LAN IS Neighbor info. A complete set of CSNPs for the appropriate level is sent over the circuit. LSPs which are updated as a result of the existence of the adjacency to IS-R are sent to IS-R, but IS-T does NOT propagate its full LSP Database. This is done to minimize the amount of redundant flooding.

IS-R uses normal adjacency bring up rules based on the IS Neighbor LAN Address information it receives in LAN IIH PDUs from IS-T and advertises its IS neighbor to IS-T in an LSP other than a UDL-LSP. Note that there is no pseudo-node on a UDL LAN circuit - therefore both IS-T and IS-R MUST advertise an IS Neighbor TLV to each other, not to a pseudo-node. This is identical to what is done on a Point-to-Point subnetwork. Following transition of the adjacency to IS-T to the UP state IS-R MAY request IS-T to flood its complete LSP Database by sending an LSP Range sub-TLV to IS-T in a UDL-LSP.

3.3. UDL link metric configuration

What metrics are configured on a UDL depend upon the intended use of the UDL. If the UDL is to be used for unicast forwarding, then IS-T should be configured with the value appropriate to its intended preference in the network topology and IS-R should be configured with maximum link metric ($2^{24} - 1$) as defined in [RFC5305] (assuming wide metrics are in use). If the UDL is to be used for building a multicast Reverse Path Forwarding tree, then IS-R should be configured with the value appropriate to its intended preference in the network topology and IS-T should be configured with maximum link metric ($2^{24} - 1$). If the link is to be used for both unicast forwarding and multicast, then it is necessary to have two different metric configurations and perform two different SPF calculations. This may be achieved through the use of multi-topology extensions as defined in [RFC5120]. Note that the configured link metrics have no bearing on adjacency establishment - they only affect the building of a Shortest Path Tree (SPT).

4. Adjacency Maintenance

This section defines how adjacencies are maintained once established. Adjacency maintenance is defined without the need to send periodic UDL-LSP updates as this would be a significant burden on the entire network.

4.1. Adjacency Maintenance by IS-T

IS-T sends IIH PDUs as normal on a UDL. As IS-R does NOT send IIH PDUs to IS-T, IS-T maintains the adjacency to IS-R so long as all of the following conditions are TRUE:

- o IS-T has a valid UDL-LSP from IS-R which includes Point-to-Point UDL IS Neighbor information or LAN UDL IS Neighbor information (as appropriate) regarding the adjacency IS-R has with IS-T on the UDL.
- o IS-T can calculate a return path rooted at IS-R to IS-T which does not traverse the UDL on which the adjacency is associated

When either of the above conditions becomes FALSE, IS-T brings down its adjacency to IS-R. Note that the return path calculation is only required when a topology change occurs in the network. It therefore need only be done in conjunction with a normal event driven SPF calculation.

NOTE: Immediately after the adjacency to IS-R has come up, if the only available return path traverses a UDL link on which the

adjacency is still in the process of coming UP, the return path check will fail. This is possible because we bypass normal flooding rules to allow the UDL-LSP to be flooded even when the adjacency is not UP on a UDL link (as described later in this document). If IS-T immediately brings the adjacency to IS-R down in this case, a circular dependency condition arises. To avoid this, if the return path check fails immediately after the adjacency comes up, a timer *Tp* is started. The timer is cancelled when a return path check succeeds. If the timer expires, IS-T brings down the adjacency to IS-R. A recommended value for the timer *Tp* is a small multiple (e.g., "twice") of the estimated time necessary to propagate LSPs across the entire domain.

Although it is unorthodox to bring up an adjacency without confirmed two way connectivity, the extension is well grounded because the receipt of IS-R's UDL-LSP by IS-T is indicative of the existence of a return path even though it cannot yet be confirmed by examination of the LSP database. This unconfirmed two way connectivity is a condition which we do not want to persist indefinitely - hence the use of timer *Tp*.

4.2. Adjacency Maintenance by IS-R

IS-R maintains its adjacency with IS-T based on receipt of IIHs from IS-T as normal. So long as IS-T follows the rules for adjacency maintenance described in the previous section this is sufficient.

Further protection against pathological behavior on the part of IS-T (e.g., failure to perform the return path calculation after a topology change) MAY be implemented by IS-R. When IS-R receives a CSNP from IS-T which contains an SNP entry identifying an LSP which is not in IS-R's Link State Database (LSDB) a timer *Tf* is started for each such LSP. This includes entries which are older than, newer than, or non-existent in IS-R's LSDB. The timer *Tf* is cancelled if:

- o The associated LSP is received by IS-R on any circuit by normal operation of the Update process or
- o A subsequent set of CSNPs received from IS-T does not include the LSP entry

If any timer *Tf* expires IS-R brings down the adjacency with IS-T.

In the absence of pathological behavior by IS-T the *Tf* extension is not required. Its use is therefore optional.

4.3. Use of BFD

A multi-hop BFD session [RFC5883] MAY be established between IS-T and IS-R. This can be used to provide fast failure detection. If used, this would also make the calculation by IS-T of a return path from IS-R to IS-T optional.

Support for [RFC6213] requires that the BFD session come up before the IS-IS adjacency comes up when both neighbors advertise BFD support. In the event that there is a UDL link on the return path from IS-R to IS-T and the adjacency on that link is also in the process of coming up this could introduce a circular dependency between the state of the BFD sessions and the state of the UDL adjacencies. Therefore [RFC6213] is NOT supported on UDLs.

4.4. Graceful Restart Support

Graceful restart as defined in [RFC5306] is NOT supported on UDLs.

In the event IS-R is restarting, signaling of restart state would require IS-R to regenerate UDL-LSPs prior to synchronization of the LSPDB. In the event IS-T is restarting, LSPDB synchronization would require the sending of CSNPs from IS-R to IS-T - which is not supported.

5. Operation of the Update Process on a UDL

For purposes of LSP propagation IS-T views the UDL as if it were a broadcast subnetwork where IS-T is the Designated Intermediate System (DIS). This is true regardless of the mode of operation of the circuit (point-to-point or broadcast). Therefore, IS-T propagates new LSPs on the UDL as they arrive but after sending an LSP on the UDL the SRM flag for that LSP is cleared i.e. no acknowledgement for the LSP is required or expected. IS-T also sends periodic CSNPs on the UDL.

IS-R cannot propagate LSPs to IS-T on the UDL. IS-R also cannot acknowledge LSPs received from IS-T on the UDL. In this respect IS-R operates on the UDL in a manner identical to a non-DIS on a broadcast circuit. If an LSP entry in a CSNP received from IS-T identifies an LSP which is "newer than" an LSP in IS-R's LSDB, IS-R MAY request the LSP from IS-T by sending a UDL-LSP with an LSP entry as described above. Since IS-R's UDL-LSP(s) will be propagated throughout the network even though the information is only of use to IS-Ts, it is recommended that some small delay occur between the receipt of a CSNP from IS-T and the generation of a UDL-LSP with an updated LSP entry by IS-R so as to allow for the possible receipt of the LSP either from IS-T or on another link.

If the number of LSP entries to be requested exceeds the space available in the UDL TLV associated with the adjacency to IS-T, IS-R MUST NOT generate multiple UDL TLVs associated with the same adjacency. Instead it should maintain the state of SSN flags appropriately for the LSP entries that require updates and send additional LSP entries (if necessary) in a subsequent UDL-LSP after the previously requested updates arrive.

Use of the LSP Range sub-TLV by IS-R allows more efficient encoding of a request for multiple LSPs. This could be especially useful following an adjacency UP event on a UDL. As described in Section 3, IS-T does NOT propagate its full LSP database following transition of an adjacency to IS-R to the UP state. This is consistent with IS-T operating in the role of DIS on a broadcast circuit. If IS-R has neighbors on other circuits it is possible that it will have received LSPs from other neighbors. In such a case flooding of the full LSP database by IS-T would be redundant. It is therefore left to the discretion of IS-R to request those portions of the LSP database which are not current. This is consistent with IS-R operating as a non-DIS on a broadcast circuit.

On receipt of a UDL-LSP generated by IS-R, IS-T checks the neighbor information in each UDL-TLV. If the information matches an existing adjacency that IS-T has with IS-R then IS-T sets SRM flag on the UDL for any LSPs in its LSDB which are "newer" than the corresponding entries IS-R sent in LSP Entry sub-TLVs in UDL TLVs. SRM flags are also set on the UDL for LSPs which fall in the ranges specified in LSP Range sub-TLVs in UDL TLVs. UDL-TLVs associated with adjacencies to routers other than IS-T are ignored by IS-T.

6. Support for UDL on the Return Path

If all return paths from IS-R to IS-T traverse a UDL, then in order to bring up the adjacency between IS-T and IS-R at least one of the adjacencies on a return path UDL must already be UP. This is required because IS-T relies on receiving the UDL-LSP(s) generated by IS-R in order to bring up its adjacency. In order to overcome a circular dependency in the case where multiple pairs of UDL neighbors are trying to bring up an adjacency at the same time, an extension to LSP propagation rules is required.

When a new UDL-LSP is received by any IS which has one or more active UDLs on which it is operating as an IS-T, the set of neighbors other than the local system which are advertised in UDL-TLVs in the received UDL-LSP is extracted - call this UDL-LSP-ISN-SET. A return path from the originating IS-R to each neighbor in the UDL-LSP-ISN-SET is calculated. If there is no return path to one or more neighbors in this set periodic propagation of that UDL-LSP on all

UDLs on which the local system acts as IS-T is initiated regardless of the state of an adjacency on that UDL. Periodic transmission of that UDL-LSP continues until a return path to all neighbors in the UDL-LSP-ISN-SET exists. This calculation is redone whenever the UDL-LSP is updated and when a topology change in the network occurs as a result of updates to the LSDB. Note that periodic retransmission is only done on UDLs on which the local system acts as IS-T.

If the network is partitioned the lack of a return path from a given IS-R to a given IS-T may persist. It is therefore recommended that the periodic retransmission employ an exponential backoff timer such that when the partition persists the periodic retransmission period is long enough so as to not represent a significant burden. It is recommended that the periodic retransmission be initially set to the locally configured CSNP interval. Note that periodic retransmission is only performed on UDL links and if an IS-R has previously received the same UDL-LSP it will silently ignore the retransmission since the UDL-LSP will already be in its LSDB. Unnecessary reflooding of the retransmitted UDL-LSP beyond the UDL does not occur.

IS-R MUST accept and propagate UDL-LSPs received on a UDL even when there is no adjacency in the UP state on the UDL circuit. Flooding of UDL-LSPs by IS-R uses normal flooding rules. LSPs received by IS-R on the UDL which do NOT include UDL TLVs are discarded unless the adjacency is UP (normal processing).

This extension allows establishment of an adjacency on a UDL even when the return path transits another UDL which is also in the process of bringing up an adjacency. The periodic nature of the flooding is meant to compensate for the unreliability of the flooding. After the adjacency is UP, IS-R can request LSPs from IS-T by putting LSP entries into UDL-LSPs - but that ability is not available until the adjacency is UP.

7. IANA Considerations

This document requires the definition of a new IS-IS TLV to be reflected in the "IS-IS TLV Codepoints" registry:

Type	Description	IIH LSP SNP Purge			
----	-----	---	---	---	---
11	Unidirectional Link Information	N	Y	N	Y

This document requires that a new IANA registry be created to control the assignment of sub-TLV code points to be advertised within a Unidirectional Link Information TLV. The registration procedure is "Expert Review" as defined in [RFC5226]. The following sub-TLVs are

defined by this document. Values are suggested values subject to assignment by IANA.

Value	Description
-----	-----
1	Manual Area Addresses
6	LAN IS Neighbor
8	LSP Range
9	LSP Entry
129	Protocols Supported
132	IP Interface Address
229	Multi-topology
232	IPv6 Interface Address
233	IPv6 Global Interface Address
240	Point-to-Point IS Neighbor

8. Security Considerations

Security concerns for IS-IS are addressed in [IS-IS], [RFC5304], and [RFC5310].

9. Acknowledgements

The idea of supporting IS-IS on UDLs without using tunnels or encapsulation was originally introduced in the US patent "Support of unidirectional link in IS-IS without IP encapsulation and in presence of unidirectional return path" (patent number: 7,957,380), by Sina Mirtorabi, Abhay Kumar Roy, Lester Ginsberg.

10. References

10.1. Normative References

- [IS-IS] "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.
- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

- [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.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC5303] Katz, D., Saluja, R., and D. Eastlake, "Three-Way Handshake for IS-IS Point-to-Point Adjacencies", RFC 5303, 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.
- [RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, February 2011.

10.2. Informational References

- [RFC5304] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, October 2008.
- [RFC5306] Shand, M. and L. Ginsberg, "Restart Signaling for IS-IS", RFC 5306, October 2008.
- [RFC5309] Shen, N. and A. Zinin, "Point-to-Point Operation over LAN in Link State Routing Protocols", RFC 5309, 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.
- [RFC5883] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for Multihop Paths", RFC 5883, June 2010.
- [RFC6213] Hopps, C. and L. Ginsberg, "IS-IS BFD-Enabled TLV", RFC 6213, April 2011.
- [RFC6232] Wei, F., Qin, Y., Li, Z., Li, T., and J. Dong, "Purge Originator Identification TLV for IS-IS", RFC 6232, May 2011.

Authors' Addresses

Les Ginsberg
Cisco Systems
510 McCarthy Blvd.
Milpitas, CA 95035
USA

Email: ginsberg@cisco.com

Sina Mirtorabi
Cisco Systems
3800 Zankar Road
San Jose, CA 95134
USA

Email: smirtora@cisco.com

Stefano Previdi
Cisco Systems
Via Del Serafico 200
Rome 0144
Italy

Email: sprevidi@cisco.com

Abhay Roy
Cisco Systems
560 McCarthy Blvd.
Milpitas, CA 95135
USA

Email: akr@cisco.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2014

Z. Li
N. Wu
Q. Zhao
Huawei Technologies
A. Atlas
C. Bowers
Juniper Networks
J. Tantsura
Ericsson
October 18, 2013

Intermediate System to Intermediate System (IS-IS) Extensions for
Maximally Redundant Trees (MRT)
draft-li-isis-mrt-00

Abstract

This document describes necessary extensions to IS-IS to support the distributed computation of Maximally Redundant Trees (MRT). Some example uses of the MRTs include IP/LDP Fast-Reroute and global protection or live-live for multicast traffic. The extensions indicate what MRT profile(s) each router supports. Different MRT profiles can be defined to support different uses and to allow transition of capabilities. An extension is introduced to flood MRT-Ineligible links, due to administrative policy.

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 21, 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. Requirements Language	3
3. Terminology	3
4. Overview of IS-IS Signaling Extensions for MRT	3
4.1. Supporting MRT Profiles	4
4.2. Electing GADAG Root	4
4.3. Advertising MRT-Ineligible Links for MRT	5
4.4. Triggering an MRT Computation	5
5. MRT Capability Advertisement	5
5.1. Advertising MRT Capability in IS-IS LSP	6
5.2. MRT Profile sub-TLV in IS-IS Router CAPABILITY TLV	6
5.3. MRT-Ineligible Links sub-TLV in IS-IS Router CAPABILITY TLV	7
6. Handling MRT Capability Sending and Receiving	8
6.1. Advertising MRT extension	8
6.2. Parsing MRT extension	9
7. Backwards Compatibility	9
8. Security Considerations	10
9. IANA Considerations	10
10. References	10
10.1. Normative References	10
10.2. Informative References	10
Authors' Addresses	11

1. Introduction

The IS-IS protocol is specified in [ISO10589], with extensions for supporting IPv4 and IPv6 specified in [RFC1195] and [RFC5308]. Each Intermediate System (IS) (router) advertises one or more IS-IS Link State Protocol Data Units (LSPs) with routing information. Each LSP is composed of a fixed header and a number of tuples, each consisting of a Type, a Length, and a Value. Such tuples are commonly known as TLVs, and are a good way of encoding information in a flexible and extensible format.

[I-D.ietf-rtgwg-mrt-frr-architecture] gives a complete solution for IP/LDP fast-reroute using Maximally Redundant Trees (MRT) to provide alternates. This document describes the necessary signaling extensions for supporting MRT-FRR used in IS-IS routing domain.

2. 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 [RFC2119].

3. Terminology

Redundant Trees (RT): A pair of trees where the path from any node X to the root R along the first tree is node-disjoint with the path from the same node X to the root R along the second tree. These can be computed in 2-connected graphs.

Maximally Redundant Trees (MRT): A pair of trees where the path from any node X to the root R along the first tree and the path from the same node X to the root R along the second tree share the minimum number of nodes and the minimum number of links. Each such shared node is a cut-vertex. Any shared links are cut-links. Any RT is an MRT but many MRTs are not RTs.

MRT Island: From the computing router, the set of routers that support a particular MRT profile and are connected via MRT-eligible links.

GADAG: Generalized Almost Directed Acyclic Graph - a graph which is the combination of the ADAGs of all blocks. Transforming a network graph into a GADAG is part of the MRT algorithm.

MRT-Red: MRT-Red is used to describe one of the two MRTs; it is used to describe the associated forwarding topology and MT-ID. Specifically, MRT-Red is the decreasing MRT where links in the GADAG are taken in the direction from a higher topologically ordered node to a lower one.

MRT-Blue: MRT-Blue is used to describe one of the two MRTs; it is used to describe the associated forwarding topology and MT-ID. Specifically, MRT-Blue is the increasing MRT where links in the GADAG are taken in the direction from a lower topologically ordered node to a higher one.

4. Overview of IS-IS Signaling Extensions for MRT

As stated in [I-D.enyedi-rtgwg-mrt-frr-algorithm], it is necessary for each MRT-Capable router to compute MRT next hops in a consistent fashion. This is achieved by using same MRT profile and selecting the unique root in an MRT Island which is connected by MRT-Eligible links. Each of these issues will be discussed in following sections separately.

4.1. Supporting MRT Profiles

The contents and requirements of an MRT profile has been defined in [I-D.ietf-rtgwg-mrt-frr-architecture]. The parameters and behavioral rules contained in an MRT profile define one router's MRT capabilities. Based on common capabilities, one unified MRT Island is built.

The MRT-Capable router MUST advertise its corresponding MRT profiles by IS-IS protocol extension within IS-IS routing domain. The capabilities of advertiser MUST conform to the profile it claimed completely, especially the MT-IDs, the algorithm and the corresponding forwarding mechanism. This advertisement MUST have level scope. One router MAY support multiple MRT profiles and it MUST advertise these profiles in corresponding IS-IS level. The MT-IDs used in one supported MRT Profile MUST NOT overlap with those MT-IDs used in a different supported MRT Profile.

The default MRT Profile is defined in [I-D.ietf-rtgwg-mrt-frr-architecture]. Its behavior is intended to support IP/LDP unicast and multicast Fast-Reroute. MRT-Capable routers SHOULD support the default MRT profile.

4.2. Electing GADAG Root

As per [I-D.enyedi-rtgwg-mrt-frr-algorithm], a GADAG root MUST be selected for one MRT Island. An unique GADAG root in common-sense among MRT Island routers is a necessity to do MRT computation. Since the selection of the GADAG root can affect the alternates and the traffic through it, the selection rules give network operator a knob to control the alternates and the traffic inside the MRT Island. Relevant discussion for the relationship between GADAG root role and MRT Island alternates is out of the scope of this document.

Each MRT-Capable router MUST advertise its priority for GADAG root selection. One router can only have one priority in the same MRT Island. It can have multiple priorities for different MRT Islands it supports. Routers that are marked as overloaded([RFC3787]) are not qualified as candidate for root selection. A GADAG root is selected by comparing their priorities. The router with the highest priority among those available candidates is the GADAG root with higher system ID as the tie-breaker if priorities are same.

When the current root is out of service or new router with higher priority joined into the MRT Island, the GADAG root MUST be re-selected. A new MRT computation will be triggered because of such a topology change.

4.3. Advertising MRT-Ineligible Links for MRT

For certain administrative or management reason, some links may not be involved into MRT computation. In this scenario, MRT-Capable router MUST claim those MRT-Ineligible links are out of MRT Island scope. If such claim splits current MRT Island then MRT computation has to be done inside the modified MRT Island which the computing router belongs to.

4.4. Triggering an MRT Computation

An MRT Computation can be triggered through topology changes or MRT capability changes of any router in the MRT Island. It is always triggered for a given MRT Profile in the corresponding level. First, the associated MRT Island is determined. Then, the GADAG Root is selected. Finally, the actual MRT algorithm is run to compute the transit MRT-Red and MRT-Blue topologies. Additionally, the router MAY choose to compute MRT-FRR alternates or make other use of the MRT computation results.

Prefixes can be attached and detached and have their associated MRT-Red and MRT-Blue next-hops computed without requiring a new MRT computation.

5. MRT Capability Advertisement

MRT-Capable router MUST identify its MRT capabilities through IS-IS Link State Packet(LSP) in level scope.

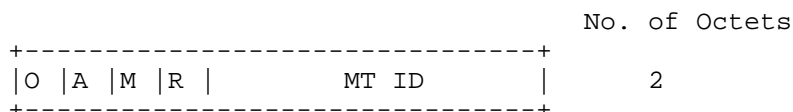
5.1. Advertising MRT Capability in IS-IS LSP

One new M-bit is introduced into TLV 229 to identify router is MRT-Capable. Structure of TLV 229 is stated in [RFC5120] as pictured below:

TYPE: 229

LENGTH: total length of the value field, it SHOULD be 2 times the number of MT components.

VALUE: one or more 2-byte MT components, structured as follows:



Bit M identifies the originator is of MRT-Capable. The MRT-Blue and the MRT-Red alternates will be calculated for the MT identified by MT-ID.

This M-bit MUST be set and checked in LSP fragment 0. An MRT-Capable router MUST advertise this TLV with M-bit set for corresponding MT. For instance, if M-bit is set for MT-ID #0, MRT alternates will be calculated for standard topology.

If only M-bit is advertised for MRT-Capabilities without any other MRT information then the router is regarded as supporting default MRT profile with default GADAG root priority.

5.2. MRT Profile sub-TLV in IS-IS Router CAPABILITY TLV

One new MRT Profile sub-TLV is introduced into IS-IS Router CAPABILITY TLV[RFC4971] to advertise MRT capabilities. Since MRT is per level scope, the S-bit and D-bit of IS-IS Router CAPABILITY TLV MUST be set zero. Structure of MRT Profile sub-TLV is pictured as below:

TYPE: TBD

LENGTH: 8 octets

VALUE:

MT ID (2 octet with 4 bits reserved)

Profile ID (1 octet)

GADAG priority (1 octet)

R R R R	MT ID		2
	Profile ID		1
	GADAG Priority		1
	MRT-Blue MT ID		2
	MRT-Red MT ID		2

12-bit MT ID represents the base MT topology which MRT computation is based on. Profile ID represents the MRT profile this router supports and GADAG Priority is the priority for root selection. The range of this priority is [0, 255] with 128 as the default value. Higher numerical value means higher priority.

Those routers which do not want to be involved into GADAG root selection can have priority 0. If all routers in MRT Island carry the same priority then the one with the highest system ID has to be chosen as GADAG root.

If the MRT-Blue MT-ID is 0, then the value specified in the associated MRT Profile is assumed. If the MRT-Red MT-ID is 0, then the value specified in the associated MRT profile is assumed. The MRT-Blue MT-ID and MRT-Red MT-ID MUST NOT be the reserved values for MT-ID([RFC5120]). The value for MRT-Blue MT-ID and MRT-Red MT-ID MUST be different except for 0. As stated above, the MRT-Blue MT-ID and MRT-Red MT-ID MUST NOT overlap among profiles if multiple MRT-Profile sub-TLVs are advertised.

This sub-TLV can occur multiple times if this router support multiple MRT profiles. This can happen during transition or to support multiple uses of MRT which prefer different profiles.

5.3. MRT-Ineligible Links sub-TLV in IS-IS Router CAPABILITY TLV

As a matter of policy, some links may not be available for the MRT computation, which can prevent alternates or traffic using these links. For instance, policy can be made to prevent fast-rerouted traffic from taking those links.

For a link to be excluded from the MRT computation, it MUST be advertised as sub-TLV in IS-IS Router CAPABILITY TLV which is in level scope with B-bit and D-bit unset. The MRT-Ineligible Link sub-TLV is structured as below:

TYPE: TBD

LENGTH: from 9 to 255 octets

VALUE:

MT ID (2 octet with 4 bits reserved)

System ID and pseudo-node number (7 octet for each MRT-Ineligible Link)

	No. of Octets
+-----+ R R R R MT ID +-----+	2
+-----+ System ID and pseudonode number +-----+	7
+-----+ Default metric +-----+	3
. . .	.
+-----+ System ID and pseudonode number +-----+	7
+-----+ Default metric +-----+	3

Each MRT-Ineligible Link is identified by neighbor's System ID and pseudo-node number and Default metric, same as IS Reachability TLV. This sub-TLV MAY occur multiple times if multiple links are ineligible.

6. Handling MRT Capability Sending and Receiving

The M-bit which identifies router's MRT capability MUST be advertised in LSP fragment 0. Those MRT related sub-TLVs SHOULD be ignored when MRT Capability bit is unset. When changes in MRT capabilities are received, an MRT computation SHOULD be triggered but MAY be delayed for a while to allow reception of all MRT-related information.

6.1. Advertising MRT extension

MRT sub-TLVs are encapsulated in the Router Capability TLV and advertised through LSP PDU for the level-wide. MRT sub-TLVs are optional. If one router does not support MRT, it MUST NOT advertise those sub-TLVs.

Since the advertisement scope of the MRT sub-TLV is level-wide, the D-Bit and S-Bit of the Router Capability TLV MUST be set as 0 when it is advertised. If other sub-TLVs in the Router Capability TLV need different values for those two bits, there MUST be an independent Router Capability TLV for MRT sub-TLVs.

When MRT related information is changed for the router or existing IS-IS LSP mechanisms are triggered for refreshing or updating, MRT sub-TLVs MUST be advertised if the router is MRT-Capable.

For administrative policies or reasons, certain links can not be involved into MRT Computation. MRT-Ineligible sub-TLV is used to advertise these links among MRT Island.

6.2. Parsing MRT extension

MRT extension MUST NOT affect the peer setup and the routing calculation of the standard topology.

MRT sub-TLVs SHOULD be validated like other sub-TLVs when received. MRT sub-TLVs SHOULD also be taken for the checksum calculation and authentication.

If MT-ID conflict is found for MRT-Red or MRT-blue from multiple sub-TLVs then those associated sub-TLVs MUST be ignored.

Links advertised in MRT-Ineligible sub-TLV MUST be precluded from MRT Computation. The removal of those links may change the computing router's MRT Island significantly.

7. Backwards Compatibility

The M-bit for MRT capability, the MRT Profile sub-TLV and the MRT-Ineligible Link sub-TLV defined in this document SHOULD NOT introduce any interoperability issues. Routers that do not support these MRT extensions SHOULD silently ignore them. Alternates or traffic MUST NOT be affected in current IS-IS routing domain.

8. Security Considerations

This IS-IS extension is not believed to introduce new security concerns.

9. IANA Considerations

Please allocate a value from the IS-IS Router CAPABILITY TLV[RFC4971] for the MRT Profile sub-TLV, and for the MRT-Ineligible Link sub-TLV.

10. References

10.1. Normative References

[ISO10589] ISO. Intermediate System to Intermediate System Routing Exchange Protocol for Use in Conjunction with the Protocol for Providing the Connectionless-Mode Network Service. ISO 10589, 1992.

[I-D.enyedi-rtgwg-mrt-frr-algorithm]
Envedi, G., Csaszar, A., Atlas, A., cbowers@juniper.net, c., and A. Gopalan, "Algorithms for computing Maximally Redundant Trees for IP/LDP Fast- Reroute", draft-enyedi-rtgwg-mrt-frr-algorithm-03 (work in progress), July 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.

[RFC3137] Retana, A., Nguyen, L., White, R., Zinin, A., and D. McPherson, "OSPF Stub Router Advertisement", RFC 3137, June 2001.

[RFC3787] Parker, J., "Recommendations for Interoperable IP Networks using Intermediate System to Intermediate System (IS-IS)", RFC 3787, May 2004.

10.2. Infomative References

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.

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

[RFC3787] Parker, J., "Recommendations for Interoperable IP Networks using Intermediate System to Intermediate System (IS-IS)", RFC 3787, May 2004.

- [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.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.

Authors' Addresses

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

Email: lizhenbin@huawei.com

Nan Wu
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: eric.wu@huawei.com

Quintin Zhao
Huawei Technologies
125 Nagog Technology Park
Acton, MA 01719
USA

Alia Atlas
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Email: akatlas@juniper.net

Chris Bowers
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: cbowers@juniper.net

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
USA

Email: jeff.tantsura@ericsson.com

Network Working Group
Internet Draft
Intended status: Proposed Standard
Expires: April 24, 2014

B. Liu
Huawei Technologies Co., Ltd
October 21, 2013

ISIS Auto-Configuration
draft-liu-isis-auto-conf-00.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 April 24, 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 potentially benefit from the self-configuring mechanisms.

Table of Contents

1. Introduction	3
2. IS-IS Default Configuration	3
3. IS-IS NET Generation.....	4
4. IS-IS NET Duplication Detection and Resolution	4
4.1. Router-Hardware-Fingerprint TLV	4
4.2. NET Duplication Detection and Resolution	5
5. Security Considerations	5
6. IANA Considerations	5
7. Acknowledgments	5
8. References	6
8.1. Normative References	6

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.

The auto-configuring mechanisms are designed based on IPv6-only environment. Some IPv4 environments might also be applicable, but they are not specifically considered.

The following aspects of IS-IS auto-configuration are described:

1. IS-IS Default Configuration
2. IS-IS NET self-generation
3. IS-IS Adjacency Formation

However, this draft does not provide a completely configuration-free alternative to the IS-IS protocol, since some plan work by human so far is very difficult to be achieved through algorithm. 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.

2. IS-IS Default Configuration

- o IS-IS SHOULD be enabled on all interfaces in a router 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.

3. 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 the following three parts:

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".

4. 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 would need to be changed. IS-IS auto-configuring does not consider this situation.

For the non-neighbor NET duplication detection within an area, this document utilizes a TLV as following to do it.

4.1. Router-Hardware-Fingerprint TLV

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

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

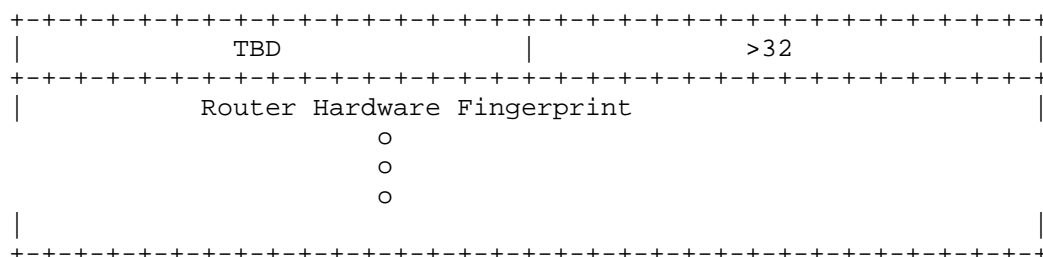


Figure 1 Router-Hardware-Fingerprint TLV Format

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. Note that, since the TLV is to detect MAC address based NET duplication, the TLV content MUST NOT use MAC address only again. Implementations SHOULD use other information exclude MAC address.

4.2. NET Duplication Detection and Resolution

The Router-Hardware-Fingerprint TLV MUST be included in the first originated level-1 LSP by every auto-configuring routers. An IS-IS auto-configuring router MUST compare a received self-originated LSP's Router-Hardware-Fingerprint TLV against its own one. If the they are not equal, there is a NET duplication and the Router with the numerically smaller router hardware fingerprint MUST generate a new NET.

After selecting a new NET, the LSP with the prior duplicate NET MUST be purged. And any IS-IS neighbor adjacencies MUST be reestablished.

5. Security Considerations

TBD.

6. IANA Considerations

The Router Hardware Fingerprint TLV type code needs an assignment.

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.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, October 2008.
- [OSPFv3AC] Lindem, A., and J. Arkko, "OSPFv3 Auto-Configuration", Work in Progress, October 2013

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

IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

S. Previdi, Ed.
C. Filsfils
A. Bashandy
Cisco Systems, Inc.
H. Gredler
Juniper Networks, Inc.
S. Litkowski
Orange
October 21, 2013

IS-IS Extensions for Segment Routing
draft-previdi-isis-segment-routing-extensions-04

Abstract

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF).

This draft describes the necessary IS-IS extensions that need to be introduced for Segment Routing.

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	4
2. Segment Routing Identifiers	4
2.1. SID/Label Sub-TLV	4
2.2. Prefix Segment Identifier (Prefix-SID Sub-TLV)	5
2.3. Adjacency Segment Identifier (Adj-SID) Sub-TLV	8
2.3.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV	9
2.3.2. Adjacency Segment Identifiers in LANs	10
2.4. SID/Label Binding TLV	12
2.4.1. Flags	13
2.4.2. Weight	13
2.4.3. Range	14
2.4.4. Prefix Length, Prefix	15
2.4.5. SID/Label Sub-TLV	15
2.4.6. ERO Metric sub-TLV	16
2.4.7. IPv4 ERO subTLV	16
2.4.8. IPv6 ERO subTLV	17
2.4.9. Unnumbered Interface ID ERO subTLV	17
2.4.10. IPv4 Backup ERO subTLV	18
2.4.11. IPv6 Backup ERO subTLV	19
2.4.12. Unnumbered Interface ID Backup ERO subTLV	19
2.4.13. Prefix ERO and Prefix Backup ERO subTLV path semantics	20
3. Router Capabilities	21
3.1. SR-Capabilities Sub-TLV	21
3.2. SR-Algorithm Sub-TLV	22
4. IANA Considerations	23
5. Manageability Considerations	24
6. Security Considerations	24
7. Contributors	24
8. Acknowledgements	24
9. References	24
9.1. Normative References	24

9.2. Informative References	25
Authors' Addresses	26

1. Introduction

Segment Routing (SR) allows for a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF). Two types of segments are defined, Prefix segments and Adjacency segments. Prefix segments represent an ecmp-aware shortest-path to a prefix, as per the state of the IGP topology. Adjacency segments represent a hop over a specific adjacency between two nodes in the IGP. A prefix segment is typically a multi-hop path while an adjacency segment, in most of the cases, is a one-hop path. SR's control-plane can be applied to both IPv6 and MPLS data-planes, and do not require any additional signaling (other than the regular IGP). For example, when used in MPLS networks, SR paths do not require any LDP or RSVP-TE signaling. Still, SR can interoperate in the presence of LSPs established with RSVP or LDP.

This draft describes the necessary IS-IS extensions that need to be introduced for Segment Routing.

Segment Routing architecture is described in [I-D.filsfils-rtgwg-segment-routing].

Segment Routing use cases are described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

2. Segment Routing Identifiers

Segment Routing architecture ([I-D.filsfils-rtgwg-segment-routing]) defines different types of Segment Identifiers (SID). This document defines the IS-IS encodings for the IGP-Prefix-SID, the IGP-Adjacency-SID, the IGP-LAN-Adjacency-SID and the Binding-SID.

2.1. SID/Label Sub-TLV

The SID/Label Sub-TLV is present in multiple Sub-TLVs defined in this document and contains a SID or a MPLS Label. The SID/Label Sub-TLV has the following format:



Type: 1

Length: variable (3 or 4)

SID/Label: if length is set to 3 then the 20 rightmost bits represent a MPLS label. If length is 4 then the value represents a 32 bits SID.

2.2. Prefix Segment Identifier (Prefix-SID Sub-TLV)

A new IS-IS Sub-TLV is defined: the Prefix Segment Identifier Sub-TLV (Prefix-SID Sub-TLV).

The Prefix-SID Sub-TLV carries the Segment Routing IGP-Prefix-SID as defined in [I-D.filsfils-rtgwg-segment-routing]. The 'Prefix SID' must be unique within a given IGP domain. The 'Prefix SID' is an index to determine the actual SID/label value inside the set of all advertised SID/label ranges of a given router. A receiving router uses the index to determine the actual SID/label value in order to construct forwarding state to a particular destination router.

In many use-cases a 'stable transport' IP Address is overloaded as an identifier of a given node. Because the IP Prefixes may be re-advertised into other levels there may be some ambiguity (e.g. Originating router vs. L1L2 router) for which node a particular IP prefix serves as identifier. The Prefix-SID Sub-TLV contains the necessary flags to dissambiguate IP Prefix to node mappings. Furthermore if a given node has several 'stable transport' IP addresses there are flags to differentiate those among other IP Prefixes advertised from a given node.

A Prefix-SID Sub-TLV is associated to a prefix advertised by a node and MAY be present in any of the following TLVs:

TLV-135 (IPv4) defined in [RFC5305].

TLV-235 (MT-IPv4) defined in [RFC5120].

TLV-236 (IPv6) defined in [RFC5308].

TLV-237 (MT-IPv6) defined in [RFC5120].

The Index inside the Prefix-SID Sub-TLV MUST be preserved when an IP Reachability TLV gets propagated across level boundaries.

The Prefix-SID Sub-TLV has the following format:

0									1									2									3								
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1				
Type									Length									Flags									Algorithm								
SID/Index																																			

where:

Type: 3

Length: variable.

Flags: 1 octet field of following flags:

0	1	2	3	4	5	6	7
R	N	P					

where:

R-Flag: Re-advertisement flag. If set, then the prefix to which this Prefix-SID is attached, has been propagated by the router either from another level (i.e.: from level-1 to level-2 or the opposite) or from redistribution (e.g.: from another protocol).

N-Flag: Node-SID flag. Optional and, if set, then the Prefix-SID refers to the router identified by the prefix. Typically, the N-Flag is set on Prefix-SIDs attached to a router loopback address. The N-Flag is set when the Prefix-SID is a Node-SID as described in [I-D.filsfils-rtgwg-segment-routing].

P-Flag: no-PHP flag. If set, then the penultimate hop MUST NOT pop the Prefix-SID before delivering the packet to the node that advertised the Prefix-SID.

Other bits: MUST be zero when originated and ignored when received.

Algorithm: the router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Examples of these algorithms are metric based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The Algorithm field allows a router to advertise algorithms that router is currently using. SR-Algorithm TLV has following structure: one octet identifying the algorithm to which the Prefix-SID is associated. Currently, the following value has been defined:

0: Shortest Path First (SPF) algorithm based on link metric.

Definitions and use of algorithms in Segment Routing are described in [I-D.filsfils-rtgwg-segment-routing]

SID/Index: 32 bit index defining the offset in the SID/Label space advertised by this router using the encodings defined in Section 3.1.

Multiple Prefix-SIDs Sub-TLVs MAY appear on the same prefix in which case each SID is encoded as a separate Sub-TLV. When multiple Prefix-SID Sub-TLVs are present, the receiving router MUST use the first encoded SID and MAY use the subsequent ones.

The No-PHP flag MUST be set on the Prefix-SIDs associated with reachability advertisements which were originated by other routers and leaked (either from Level-1 to Level-2 or vice versa).

The R-Flag MUST be set for prefixes that are not local to the router and either:

advertised because of propagation (Level-1 into Level-2);

advertised because of leaking (Level-2 into Level-1);

advertised because redistribution (e.g.: from another protocol).

In the case where a Level-1-2 router has local interface addresses configured in one level, it may also propagate these addresses into the other level. In such case, the Level-1-2 router MUST NOT set the R bit. The R-bit MUST be set only for prefixes that are not local to

the router and advertised by the router because of propagation and/or leaking.

The N-Flag is used in order to define a Node-SID. A router MAY set the N-Flag only if all of the following conditions are met:

The prefix to which the Prefix-SID is attached is local to the router. I.e.: the prefix is configured on one of the local interfaces. (e.g.: 'stable transport' loopback).

The prefix to which the Prefix-SID is attached MUST have a Prefix length of either /32 (IPv4) or /128 (IPv6).

The router MUST ignore the N-Flag on a received Prefix-SID if the prefix has a Prefix length different than /32 (IPv4) or /128 (IPv6).

The router behavior determined by the P, R and N flags are described in [I-D.filsfils-rtgwg-segment-routing].

2.3. Adjacency Segment Identifier (Adj-SID) Sub-TLV

A new IS-IS Sub-TLV is defined: the Adjacency Segment Identifier Sub-TLV (Adj-SID Sub-TLV).

The Adj-SID Sub-TLV is an optional Sub-TLV carrying the Segment Routing IGP-Adjacency-SID as defined in [I-D.filsfils-rtgwg-segment-routing] with flags and fields that may be used, in future extensions of Segment Routing, for carrying other types of SIDs.

IS-IS adjacencies are advertised using one of the IS-Neighbor TLVs below:

TLV-22 [RFC5305]

TLV-222 [RFC5120]

TLV-23 [RFC5311]

TLV-223 [RFC5311]

TLV-141 [RFC5316]

Multiple Adj-SID Sub-TLVs MAY be associated with a single IS-neighbor. Examples where more than one Adj-SID may be used per IS-neighbor are described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

2.3.1. Adjacency Segment Identifier (Adj-SID) Sub-TLV

The following format is defined for the Adj-SID Sub-TLV:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Type      |      Length      |      Flags      |      Weight      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     SID/Label Sub-TLV (variable)                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

where:

Type: 31

Length: variable.

Flags: 1 octet field of following flags:

```

      0 1 2 3 4 5 6 7
+-----+-----+
| F | B |           |
+-----+-----+

```

where:

F-Flag: Address-Family flag. If unset, then the Adj-SID refers to an adjacency with outgoing IPv4 encapsulation. If set then the Adj-SID refers to an adjacency with outgoing IPv6 encapsulation.

B-Flag: Backup flag. If set, the Adj-SID refers to an adjacency being protected (e.g.: using IPFRR or MPLS-FRR) as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

Other bits: MUST be zero when originated and ignored when received.

Weight: 1 octet. The value represents the weight of the Adj-SID for the purpose of load balancing. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

SID/Label Sub-TLV: contains the SID/Label value as defined in Section 2.1.

An SR capable router MAY allocate an Adj-SID for each of its adjacencies and SHOULD set the B-Flag when the adjacency is protected by a FRR mechanism (IP or MPLS) as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

The F-flag is used in order for the router to advertise the outgoing encapsulation of the adjacency the Adj-SID is attached to. Use cases of the use of the F-flag are described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

2.3.2. Adjacency Segment Identifiers in LANs

In LAN subnetworks, the Designated Intermediate System (DIS) is elected and originates the Pseudonode-LSP (PN-LSP) including all neighbors of the DIS.

When Segment Routing is used, each router in the LAN MAY advertise the Adj-SID of each of its neighbors. Since, on LANs, each router only advertises one adjacency to the DIS (and doesn't advertise any other adjacency), each router advertises the set of Adj-SIDs (for each of its neighbors) inside a newly defined Sub-TLV part of the TLV advertising the adjacency to the DIS (e.g.: TLV-22).

The following new Sub-TLV is defined: LAN-Adj-SID (Type 32) containing the set of Adj-SIDs the router assigned to each of its LAN neighbors.

The format of the LAN-Adj-SID Sub-TLV is as follows:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Type      |  Length    |  Flags      |  Weight      |
+-----+-----+-----+-----+-----+-----+-----+-----+
|
|                                     System-ID (6 octets)
|                                     +-----+-----+-----+-----+
|                                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|
|                                     SID/Label Sub-TLV (variable)
+-----+-----+-----+-----+-----+-----+-----+-----+

```

where:

Type: 32

Length: variable.

Flags: 1 octet field of following flags:

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

where:

F-Flag: Address Family flag. If unset, then the Adj-SID refers to an adjacency with outgoing IPv4 encapsulation. If set then the Adj-SID refers to an adjacency with outgoing IPv6 encapsulation.

B-Flag: Backup flag. If set, the LAN-Adj-SID refers to an adjacency being protected (e.g.: using IPFRR or MPLS-FRR) as described in [I-D.filsfils-rtgwg-segment-routing-use-cases].

Other bits: MUST be zero when originated and ignored when received.

Weight: 1 octet. The value represents the weight of the Adj-SID for the purpose of load balancing. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

System-ID: 6 octets of IS-IS System-ID of length "ID Length" as defined in [ISO10589].

SID/Label Sub-TLV: contains the SID/Label value as defined in Section 2.1.

Multiple LAN-Adj-SID Sub-TLVs MAY be encoded.

In case one TLV-22/23/222/223 (reporting the adjacency to the DIS) can't contain the whole set of LAN-Adj-SID Sub-TLVs, multiple advertisements of the adjacency to the DIS MUST be used, MUST have the same metric and SHOULD be inserted within the same LSP fragment.

Each router within the level, by receiving the DIS PN LSP as well as the non-PN LSP of each router in the LAN, is capable of reconstructing the LAN topology as well as the set of Adj-SID each router uses for each of its neighbors.

2.4. SID/Label Binding TLV

The SID/Label Binding TLV MAY be originated by any router in an IS-IS domain. The router may advertise a SID/Label binding to a FEC along with at least a single 'nexthop style' anchor. The protocol supports more than one 'nexthop style' anchor to be attached to a SID/Label binding, which results into a simple path description language. In analogy to RSVP the terminology for this is called an 'Explicit Route Object' (ERO). Since ERO style path notation allows to anchor SID/label bindings to both link and node IP addresses any label switched path, can be described. Furthermore also SID/Label Bindings from external protocols can get easily re-advertised.

The SID/Label Binding TLV may be used for advertising SID/Label Bindings and their associated Primary and Backup paths. In one single TLV either a primary ERO Path, a backup ERO Path or both are advertised. If a router wants to advertise multiple parallel paths then it can generate several TLVs for the same Prefix/FEC. Each occurrence of a Binding TLV with respect with a given FEC Prefix has accumulating and not canceling semantics. Due the space constraints in the 8-Bit IS-IS TLVs an originating router MAY encode a primary ERO path in one SID/Label Binding TLV and the backup ERO path in a second SID/Label Binding TLV. Note that the FEC Prefix and SID/Label Sub-TLV MUST be identical in both TLVs.

The SID/Label Binding TLV has type TBA and has the following format:

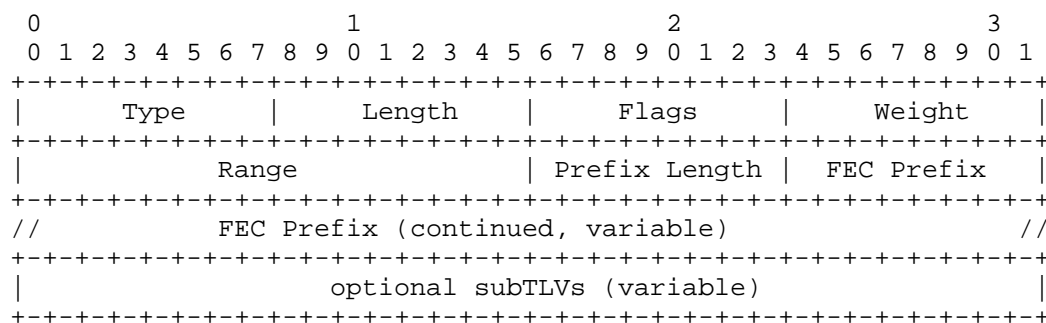


Figure 1: SID/Label Binding TLV format

- o Type: 149
- o Length: variable.
- o 1 octet of flags

- o 1 octet of Prefix length
- o 0-16 octets of FEC Prefix
- o 2 octets of Range
- o sub-TLVs, where each sub-TLV consists of a sequence of:
 - * 1 octet of sub-TLV type
 - * 1 octet of length of the value field of the sub-TLV
 - * 0-255 octets of value

2.4.1. Flags

Flags: 1 octet field of following flags:

```

  0 1 2 3 4 5 6 7
+---+---+---+---+---+---+
|F|M|X|S|   |   |
+---+---+---+---+---+---+

```

where:

F-Flag: Address Family flag. If unset, then the Prefix FEC carries an IPv4 Prefix. If set then the Prefix FEC carries an IPv6 Prefix.

M-Flag: Mirror Context flag. Set if the advertised SID/path corresponds to a mirrored context.

X-Flag: Index flag. Set if the value of the SID/Label Sub-TLV carries an index. Unset if the value of the SID/Label Sub-TLV carries a local SID/Label.

S-Flag: subTLV present 'S' flag: Set if there are subTLVs present.

Other bits: MUST be zero when originated and ignored when received.

2.4.2. Weight

Weight: 1 octet: The value represents the weight of the path for the purpose of load balancing. The use of the weight is defined in [I-D.filsfils-rtgwg-segment-routing].

2.4.3. Range

The 'Range' field provides the ability to specify a range of addresses and their associated Prefix SIDs. It is essentially a compression scheme to distribute a continuous Prefix and their continuous, corresponding SID/Label Block. If a single SID is advertised then the range field MUST be set to one. For range advertisements > 1, the number of addresses that need to be mapped into a Prefix-SID and the starting value of the Prefix-SID range.

Example 1: if the following router addresses (loopback addresses) need to be mapped into the corresponding Prefix SID indexes.

Router-A: 192.0.2.1/32, Prefix-SID: Index 1
 Router-B: 192.0.2.2/32, Prefix-SID: Index 2
 Router-C: 192.0.2.3/32, Prefix-SID: Index 3
 Router-D: 192.0.2.4/32, Prefix-SID: Index 4

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type										Length										0	0	1	1	Weight															
Range = 4																				/32										192									
.0										.2										.1										Sub-TLV Type									
Sub-TLV Length																														1									

Example-2: If the following prefixes need to be mapped into the corresponding Prefix-SID indexes:

10.1.1/24, Prefix-SID: Index 51
 10.1.2/24, Prefix-SID: Index 52
 10.1.3/24, Prefix-SID: Index 53
 10.1.4/24, Prefix-SID: Index 54
 10.1.5/24, Prefix-SID: Index 55
 10.1.6/24, Prefix-SID: Index 56
 10.1.7/24, Prefix-SID: Index 57

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	2	3	4	5	6	7	8	9
Type										Length										0 0 1 1										Weight									
Range = 7																				/24										10									
.1										.1										Sub-TLV Type										Sub-TLV Length									
																														51									

It is not expected that a network operator will be able to keep fully continuous FEC Prefix / SID/Index mappings. In order to support noncontinuous mapping ranges an implementation MAY generate several instances of Binding TLVs.

For example if a router wants to advertise the following ranges:

Range 16: { 192.168.1.1-15, Index 1-15 }

Range 6: { 192.168.1.22-27, Index 22-27 }

Range 41: { 192.168.1.44-84, Index 80-120 }

A router would need to advertise three instances of the Binding TLV.

2.4.4. Prefix Length, Prefix

The 'FEC Prefix' represents the Forwarding equivalence class at the tail-end of the advertised path. The 'FEC Prefix' does not need to correspond to a routable prefix of the originating node.

The 'Prefix Length' field contains the length of the prefix in bits. Only the most significant octets of the Prefix FEC are encoded. I.e. 1 octet for FEC prefix length 1 up to 8, 2 octets for FEC prefix length 9 to 16, 3 octets for FEC prefix length 17 up to 24 and 4 octets for FEC prefix length 25 up to 32, . . . , 16 octets for FEC prefix length 113 up to 128.

2.4.5. SID/Label Sub-TLV

The SID/Label Sub-TLV (Type 1) contains the SID/Label value as defined in Section 2.1. It MUST be present in every SID/Label Binding TLV.

2.4.6. ERO Metric sub-TLV

ERO Metric sub-TLV (Type 2) is a Sub-TLV of the SID/Label Binding TLV.

The ERO Metric sub-TLV carries the cost of an ERO path. It is used to compare the cost of a given source/destination path. A router SHOULD advertise the ERO Metric sub-TLV. The cost of the ERO Metric sub-TLV SHOULD be set to the cumulative IGP or TE path cost of the advertised ERO. Since manipulation of the Metric field may attract or distract traffic from and to the advertised segment it MAY be manually overridden.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
Type										Length										Metric																			
Metric (continued)																																							

ERO Metric sub-TLV format

where:

Type: 2

Length: 4

Metric: 4 bytes

2.4.7. IPv4 ERO subTLV

The IPv4 ERO subTLV (Type 3) describes a path segment using IPv4 address style of encoding. Its semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

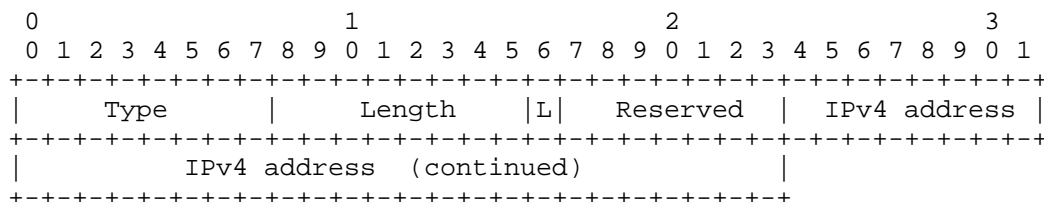


Figure 2: IPv4 ERO subTLV format

2.4.8. IPv6 ERO subTLV

The IPv6 ERO subTLV (Type 4) describes a path segment using IPv6 Address style of encoding. Its semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

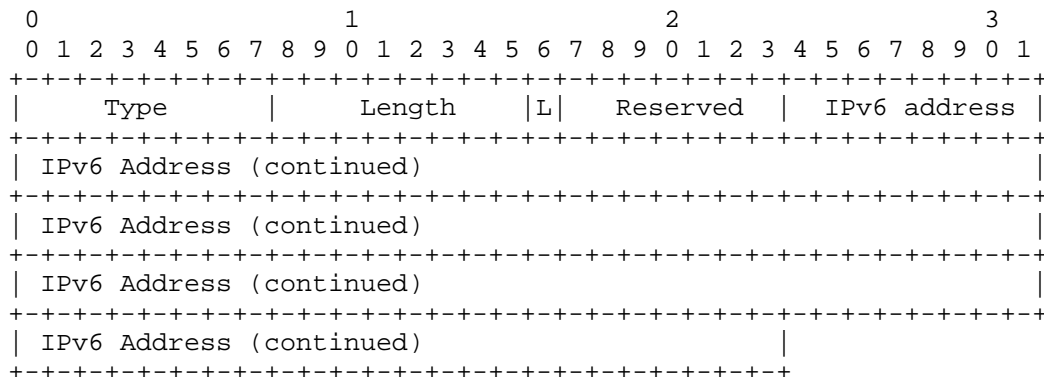


Figure 3: IPv6 ERO subTLV format

2.4.9. Unnumbered Interface ID ERO subTLV

The appearance and semantics of the 'Unnumbered Interface ID' have been borrowed from Section 4 [RFC3477].

The Unnumbered Interface-ID ERO subTLV (Type 5) describes a path segment that spans over an unnumbered interface. Unnumbered interfaces are referenced using the interface index. Interface indices are assigned local to the router and therefore not unique within a domain. All elements in an ERO path need to be unique within a domain and hence need to be disambiguated using a domain unique Router-ID.

The 'Router-ID' field contains the router ID of the router which has assigned the 'Interface ID' field. Its purpose is to disambiguate the 'Interface ID' field from other routers in the domain.

IS-IS supports two Router-ID formats:

- o (TLV 134, 32-Bit format) [RFC5305]
- o (TLV 140, 128-Bit format) [RFC6119]

The actual Router-ID format gets derived from the 'Length' field.

- o For 32-Bit Router-ID width the subTLV length is set to 8 octets.
- o For 128-Bit Router-ID width the subTLV length is set to 20 octets.

The 'Interface ID' is the identifier assigned to the link by the router specified by the router ID.

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

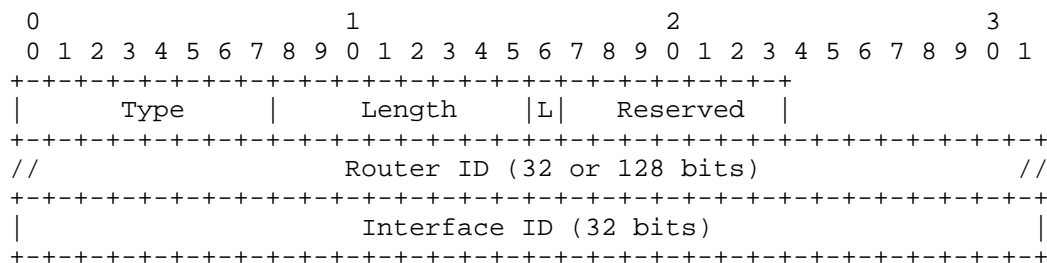


Figure 4: Unnumbered Interface ID ERO subTLV format

2.4.10. IPv4 Backup ERO subTLV

The IPv4 Backup ERO subTLV (Type 6) describes a Backup path segment using IPv4 Address style of encoding. Its appearance and semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

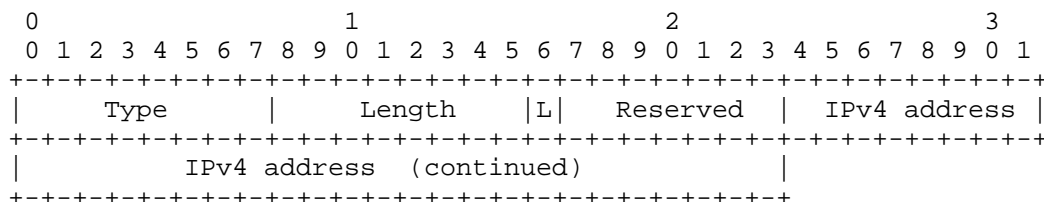


Figure 5: IPv4 Backup ERO subTLV format

2.4.11. IPv6 Backup ERO subTLV

The IPv6 Backup ERO subTLV (Type 7) describes a Backup path segment using IPv6 Address style of encoding. Its appearance and semantics have been borrowed from [RFC3209].

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

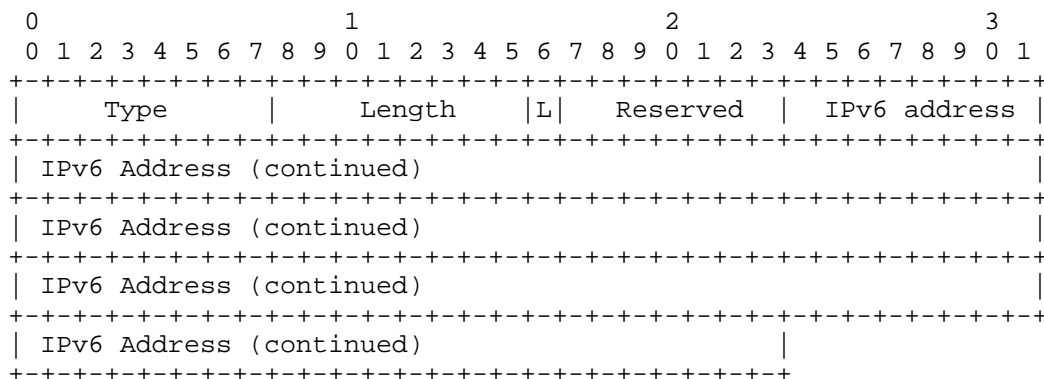


Figure 6: IPv6 Backup ERO subTLV format

2.4.12. Unnumbered Interface ID Backup ERO subTLV

The appearance and semantics of the 'Unnumbered Interface ID' have been borrowed from Section 4 [RFC3477].

The Unnumbered Interface-ID Backup ERO subTLV (Type 8) describes a Backup LSP path segment that spans over an unnumbered interface. Unnumbered interfaces are referenced using the interface index. Interface indices are assigned local to the router and therefore not unique within a domain. All elements in an ERO path need to be unique within a domain and hence need to be disambiguated using a domain unique Router-ID.

The 'Router-ID' field contains the router ID of the router which has assigned the 'Interface ID' field. Its purpose is to disambiguate the 'Interface ID' field from other routers in the domain.

IS-IS supports two Router-ID formats:

- o (TLV 134, 32-Bit format) [RFC5305]
- o (TLV 140, 128-Bit format) [RFC6119]

The actual Router-ID format gets derived from the 'Length' field.

- o For 32-Bit Router-ID width the subTLV length is set to 8 octets.
- o For 128-Bit Router-ID width the subTLV length is set to 20 octets.

The 'Interface ID' is the identifier assigned to the link by the router specified by the router ID.

The 'L' bit in the Flags is a one-bit attribute. If the L bit is set, then the value of the attribute is 'loose.' Otherwise, the value of the attribute is 'strict.'

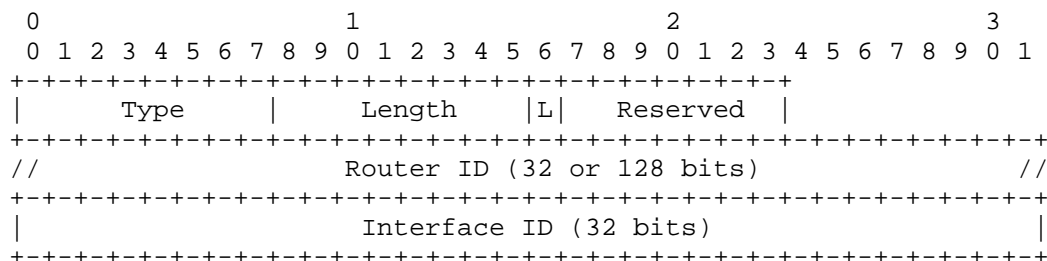


Figure 7: Unnumbered Interface ID Backup ERO subTLV format

2.4.13. Prefix ERO and Prefix Backup ERO subTLV path semantics

All 'ERO' and 'Backup ERO' information represents an ordered set which describes the segments of a path. The last ERO subTLV describes the segment closest to the egress point of the path. Contrary the first ERO subTLV describes the first segment of a path. If a router extends or stitches a label switched path it MUST prepend the new segments path information to the ERO list. The same ordering applies for the Backup ERO labels. An implementation SHOULD first encode all primary path EROs followed by the bypass EROs.

3. Router Capabilities

3.1. SR-Capabilities Sub-TLV

Segment Routing requires each router to advertise its SR data-plane capability and the range of SID/Label values it uses for Segment Routing. Data-plane capabilities and SID/Label ranges are advertised using the newly defined SR-Capabilities Sub-TLV inserted into the IS-IS Router Capability TLV-242 that is defined in [RFC4971].

The Router Capability TLV specifies flags that control its advertisement. The SR Capabilities Sub-TLV MUST be propagated throughout the level and need not to be advertised across level boundaries. Therefore Router Capability TLV distribution flags MUST be set accordingly, i.e.: the S flag MUST be unset.

The SR Capabilities Sub-TLV (Type 2) is optional, MAY appear multiple times inside the Router Capability TLV and has following format:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|  Type   |      Length   |      Flags   |      Range   |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Range (cont.) |  SID/Label Sub-TLV (variable size)  |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

where:

Type: 2

Length: variable.

Flags: 1 octet of flags. The following are defined:

```

      0
      0 1 2 3 4 5 6 7
+-----+-----+
| I | V |
+-----+-----+

```

where:

I-Flag: IPv4 flag. If set, then the router is capable of outgoing IPv4 encapsulation on all interfaces.

V-Flag: IPv6 flag. If set, then the router is capable of outgoing IPv6 encapsulation on all interfaces.

Range: 2 octets value defining the number of values of the range from the starting value defined in the SID/Label Sub-TLV.

SID/Label Sub-TLV: SID/Label value as defined in Section 2.1.

If multiple occurrence of the SR-Capabilities Sub-TLV are advertised by the same router, only the Flags in the first occurrence of the Sub-TLV are to be taken into account.

3.2. SR-Algorithm Sub-TLV

The router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Examples of these algorithms are metric based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The SR-Algorithm Sub-TLV (Type 15) allows the router to advertise the algorithms that the router is currently using. The following value has been defined:

0: Shortest Path First (SPF) algorithm based on link metric.

The SR-Algorithm Sub-TLV is inserted into the IS-IS Router Capability TLV-242 that is defined in [RFC4971].

The Router Capability TLV specifies flags that control its advertisement. The SR-Algorithm MUST be propagated throughout the level and need not to be advertised across level boundaries. Therefore Router Capability TLV distribution flags MUST be set accordingly, i.e.: the S flag MUST be unset.

The SR-Algorithm Sub-TLV is optional, it MAY only appear a single time inside the Router Capability TLV. If the SID-Label Capability Sub-TLV is advertised then the SR-Algorithm Sub-TLV MUST also be advertised.

It has following format:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+
|  Type           |      Length       |
+-----+-----+-----+-----+
| Algorithm 1     | Algorithm 2     | Algorithm ... | Algorithm n |
+-----+-----+-----+-----+
```

where:

Type: 15

Length: variable.

Algorithm: 1 octet of algorithm Section 2.2

4. IANA Considerations

This documents request allocation for the following TLVs and subTLVs.

PDU	TLV	subTLV	Type	subType	#Occurence
LSP	IS Neighbor		22, 23, 222, 223		>=0
		SID/Label		31	>0
		LAN		32	>0
LSP	IP reachability	SID/Label	135, 235, 236, 237		>=0
LSP	SID/MPLS Binding	SID/Label		3	>0
			149		>=0
		SID/Label		1	>0
		ERO Metric		2	1
		IPv4 ERO		3	>=0
		IPv6 ERO		4	>=0
		Unnumbered Interface		5	>=0
		ID ERO			
		IPv4 Backup ERO		6	>=0
		IPv6 Backup ERO		7	>=0
		Unnumbered Interface ID Backup ERO		8	>=0
LSP	Router Capability		242		>=0
		SR Capability		2	>=0

		SR		15		1	
		Algorithm					
+-----+	+-----+	+-----+	+-----+	+-----+	+-----+	+-----+	+-----+

Table 1: IANA allocations

The SID/MPLS Binding TLV requires a new sub-registry. Type value 149 has been assigned, with a starting sub-TLV value of 1, range from 1-255, and managed by Expert Review.

5. Manageability Considerations

TBD

6. Security Considerations

TBD

7. Contributors

The following people gave a substantial contribution to the content of this document: Martin Horneffer, Bruno Decraene, Igor Milojevic, Rob Shakir, Saku Ytti and Wim Henderickx.

8. Acknowledgements

We would like to thank Les Ginsberg, Dave Ward, Dan Frost, Stewart Bryant and Pierre Francois for their contribution to the content of this document.

Many thanks to Yakov Rekhter and Ina Minei for their contribution on earlier incarnations of the "Binding / MPLS Label TLV" in [I-D.gredler-isis-label-advertisement].

9. References

9.1. Normative References

[ISO10589]

International Organization for Standardization,
"Intermediate system to Intermediate system intra-domain
routing 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.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.
- [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.
- [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.
- [RFC5311] McPherson, D., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, February 2009.
- [RFC5316] Chen, M., Zhang, R., and X. Duan, "ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5316, December 2008.
- [RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, February 2011.

9.2. 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-00 (work in progress), June 2013.

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

Filsfils, C., Francois, P., Previdi, S., Decraene, B.,
Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R.,
Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe,
"Segment Routing Use Cases",
draft-filsfils-rtgwg-segment-routing-use-cases-01 (work in
progress), July 2013.

[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.

Authors' Addresses

Stefano Previdi (editor)
Cisco Systems, Inc.
Via Del Serafico, 200
Rome 00142
Italy

Email: sprevidi@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels,
BE

Email: cfilsfil@cisco.com

Ahmed Bashandy
Cisco Systems, Inc.
170, West Tasman Drive
San Jose, CA 95134
US

Email: bashandy@cisco.com

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

Email: hannes@juniper.net

Stephane Litkowski
Orange
FR

Email: stephane.litkowski@orange.com

IS-IS for IP Internets
Internet-Draft
Intended status: Standards Track
Expires: April 27, 2015

P. Sarkar, Ed.
H. Gredler
S. Hegde
Juniper Networks, Inc.
S. Litkowski
B. Decraene
Orange
Z. Li
Huawei Technologies
H. Raghuvver

October 24, 2014

Advertising Per-node Admin Tags in IS-IS
draft-psarkar-isis-node-admin-tag-03

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 administrative 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 April 27, 2015.

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
2. Administrative Tag	2
3. TLV format	4
3.1. Per-node Admin Tag sub-TLV	4
4. Elements of Procedure	5
5. Applications	6
6. Security Considerations	11
7. IANA Considerations	11
8. Acknowledgments	11
9. References	11
9.1. Normative References	11
9.2. Informative References	11
Authors' Addresses	12

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 been 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 router 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] that has either level-wise (i.e. L1 or L2) or domain-wide flooding scope. Choosing the flooding scope to flood the group tags are defined by the needs of the operator's usage and is a matter of local policy or configuration.

Operator may choose to advertise a set of per-node administrative tags across levels and another set of per-node administrative tags within the specific level. But evidently the same set of per-node administrative tags cannot be advertised both across levels and within a specific level. A receiving IS-IS router will not be able to distinguish between the significance of a per-node administrative tag advertised globally from that of a administrative tag advertised locally if they have the same value associated but different significance across different scopes.

Implementations SHOULD allow configuring one or more 'global' as well as 'level-wide' administrative tags. A operator may only need to advertise and flood a specific per-node administrative tag, either across all levels, or only within a specific level. Hence implementations MUST NOT allow configuring the same per-node administrative tag values in both 'global' and 'level-wide' scopes. However the same administrative tag value MAY be allowed to be configured and advertised for multiple levels with 'level-wide' flooding scope.

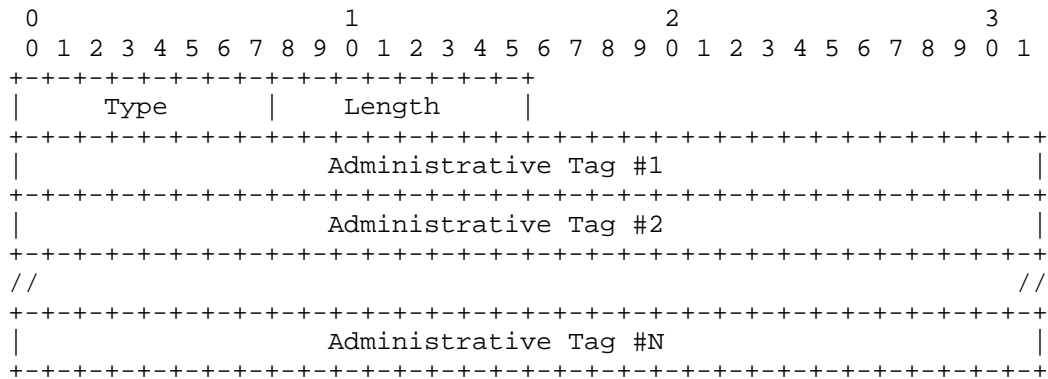
The 'global' per-node administrative tags shall have significance across the entire administrative domain and hence MUST be advertised in a Router-Capability TLV with 'global' scope (i.e. S-bit set to 1), and inserted in the LSP PDUs generated for all levels applicable. The 'level-wide' administrative tags should be copied in to a Router-Capability with 'level-wide' scope only (i.e S-bit reset to 0) and copied into the LSP PDU for the specific level.

In deployments using multi-topology routing [RFC5120], since multiple topologies within same IS-IS level share the same flooding scope configuring the same per-node administrative tag across different topologies, SHOULD NOT be allowed. Advertising the same tag value across multiple topologies will lead to same inconsistencies as with the case of advertising same tag value across 'global' and 'level-wide' flooding scope. If there is need to distinguish between the per-node administrative tags used for one topology to another, operators are advised to use disjoint sets of per-node administrative tags across such topologies.

3. TLV format

3.1. Per-node Admin Tag sub-TLV

The new Per-node 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 administrative 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.

4. Elements of Procedure

Meaning of the Per-node administrative tags is generally opaque to IS-IS. Router advertising the per-node administrative tag (or tags) may be configured to do so without knowing (or even explicitly supporting) functionality implied by the tag.

Interpretation of tag values is specific to the administrative domain of a particular network operator. The meaning of a per-node administrative tag is defined by the network local policy and is controlled via the configuration. If a receiving node does not understand the tag value, it ignores the specific tag and floods the Router Capability TLV without any change as defined in [RFC4971].

The semantics of the tag order has no meaning. 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. Tags carried by the administrative tag TLV SHOULD be used to indicate independent characteristics of a node. The TLV SHOULD be considered as an unordered list. Whilst policies may be implemented based on the presence of multiple tags (e.g., if tag A AND tag B are present), they MUST NOT be reliant upon the order of the tags (i.e., all policies should be considered commutative operations, such that tag A preceding or following tag B does not change their outcome).

As mentioned earlier, to avoid incomplete or inconsistent interpretations of the per-node administrative tags the same tag value MUST NOT be advertised by a router in Router Capabilities of different scopes. Implementations MUST NOT allow configuring the same tag value across domain-wide and 'level-wide' scopes. The same tag value MAY be allowed to be configured and advertised under 'level-wide' scope for all levels. A IS-IS Area Border Routers (ABR) participating in both levels 1 and 2 MAY advertise the same tag value in the level-specific Router Capability TLVs with 'level-wide' scope (S-bit reset to 0) generated by it. But the same tag value MUST not be advertised in any of level 1 or level 2 Router-Capability TLV with 'global' scope (S-bit set to 1).

The per-node administrative tags are not meant to be extended by the future IS-IS standards. The new IS-IS extensions MUST NOT require

use of per-node administrative tags or define well-known tag values. Per-node administrative tags are for generic use and do not require IANA registry. The future IS-IS extensions requiring well known values MAY use new Capability sub-TLVs tailored to the needs of the feature, as defined in [RFC4971].

Being part of the Router Capability TLV, the per-node administrative tag sub-TLV MUST be reasonably small and stable. In particular, but not limited to, implementations supporting the per-node administrative tags MUST NOT tie advertised tags to changes in the network topology (both within and outside the IS-IS domain) or reachability of routes.

5. Applications

This section lists several examples of how implementations might use the Per-node administrative tags. These examples are given only to demonstrate generic usefulness of the router tagging mechanism. Implementation supporting this specification is not required to implement any of the use cases. It is also worth noting that in some described use cases routers configured to advertise tags help other routers in their calculations but do not themselves implement the same functionality.

1. Auto-discovery of Services

Router tagging may be used to automatically discover group of routers sharing a particular service.

For example, service provider might desire to establish full mesh of MPLS TE tunnels between all PE routers in the area of MPLS VPN network. Marking all PE routers with a tag and configuring devices with a policy to create MPLS TE tunnels to all other devices advertising this tag will automate maintenance of the full mesh. When new PE router is added to the area, all other PE devices will open TE tunnels to it without the need of reconfiguring them.

2. Policy-based Fast-Reroute

Increased deployment of Loop Free Alternates (LFA) as defined in [RFC5286] poses operation and management challenges. [I-D.ietf-rtgwg-lfa-manageability] proposes policies which, when implemented, will ease LFA operation concerns.

One of the proposed refinements is to be able to group the nodes in IGP domain with administrative tags and engineer the LFA based on configured policies.

(a) Administrative limitation of LFA scope

Service provider access infrastructure is frequently designed in layered approach with each layer of devices serving different purposes and thus having different hardware capabilities and configured software features. When LFA repair paths are being computed, it may be desirable to exclude devices from being considered as LFA candidates based on their layer.

For example, if the access infrastructure is divided into the Access, Distribution and Core layers it may be desirable for a Distribution device to compute LFA only via Distribution or Core devices but not via Access devices. This may be due to features enabled on Access routers; due to capacity limitations or due to the security requirements. Managing such a policy via configuration of the router computing LFA is cumbersome and error prone.

With the Per-node administrative tags it is possible to assign a tag to each layer and implement LFA policy of computing LFA repair paths only via neighbors which advertise the Core or Distribution tag. This requires minimal per-node configuration and network automatically adapts when new links or routers are added.

(b) Optimizing LFA calculations

Calculation of LFA paths may require significant resources of the router. One execution of Dijkstra algorithm is required for each neighbor eligible to become next hop of repair paths. Thus a router with a few hundreds of neighbors may need to execute the algorithm hundreds of times before the best (or even valid) repair path is found. Manually excluding from the calculation neighbors which are known to provide no valid LFA (such as single-connected routers) may significantly reduce number of Dijkstra algorithm runs.

LFA calculation policy may be configured so that routers advertising certain tag value are excluded from LFA calculation even if they are otherwise suitable.

3. Controlling Remote LFA tunnel termination

[I-D.ietf-rtgwg-remote-lfa] proposed method of tunneling traffic after connected link failure to extend the basic LFA coverage and algorithm to find tunnel tail-end routers fitting LFA requirement. In most cases proposed algorithm finds more than

one candidate tail-end router. In real life network it may be desirable to exclude some nodes from the list of candidates based on the local policy. This may be either due to known limitations of the per-node (the router does accept targeted LDP sessions required to implement Remote LFA tunneling) or due to administrative requirements (for example, it may be desirable to choose tail-end router among co-located devices).

The Per-node administrative tag delivers simple and scalable solution. Remote LFA can be configured with a policy to accept during the tail-end router calculation as candidates only routers advertising certain tag. Tagging routers allows to both exclude nodes not capable of serving as Remote LFA tunnel tail-ends and to define a region from which tail-end router must be selected.

4. Mobile backhaul network service deployment

The topology of mobile backhaul network usually adopts ring topology to save fiber resource and it is divided into the aggregate network and the access network. Cell Site Gateways(CSGs) connects the eNodeBs and RNC(Radio Network Controller) Site Gateways(RSGs) connects the RNCs. The mobile traffic is transported from CSGs to RSGs. The network takes a typical aggregate traffic model that more than one access rings will attach to one pair of aggregate site gateways(ASGs) and more than one aggregate rings will attach to one pair of RSGs.

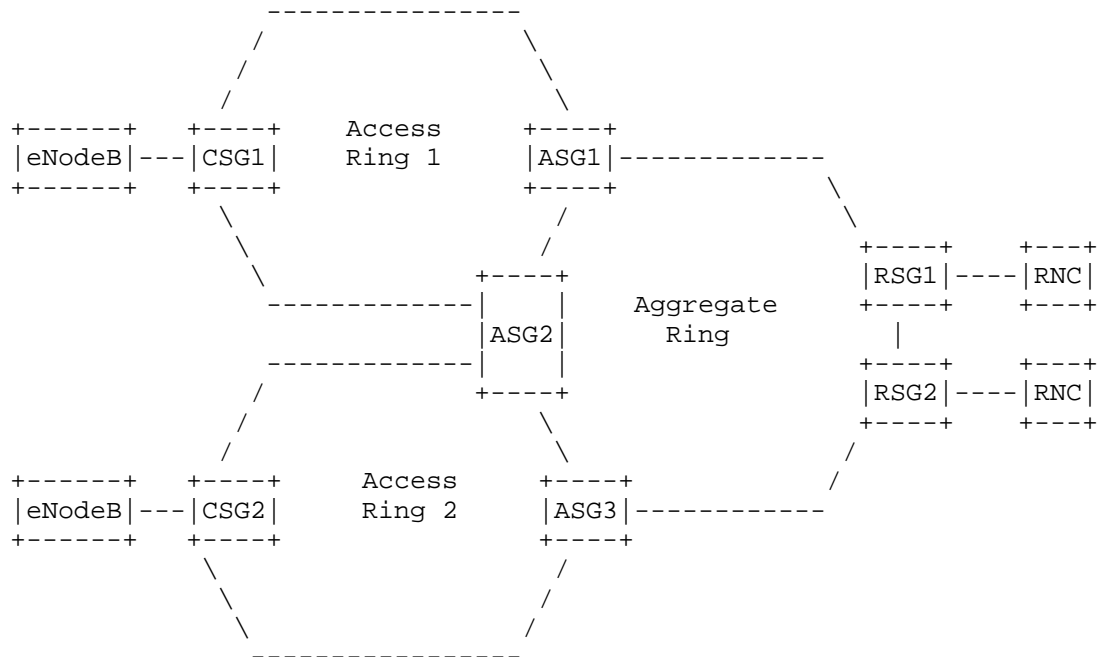


Figure 2: Mobile Backhaul Network

A typical mobile backhaul network with access rings and aggregate links is shown in figure above. The mobile backhaul networks deploy traffic engineering due to the strict Service Level Agreements(SLA). The TE paths may have additional constraints to avoid passing via different access rings or to get completely disjoint backup TE paths. The mobile backhaul networks towards the access side change frequently due to the growing mobile traffic and addition of new eNodeBs. It's complex to satisfy the requirements using cost, link color or explicit path configurations. The per-node administrative tag defined in this document can be effectively used to solve the problem for mobile backhaul networks. The nodes in different rings can be assigned with specific tags. TE path computation can be enhanced to consider additional constraints based on per-node administrative tags.

5. Policy-based Explicit Routing

Partially meshed network provides multiple paths between any two nodes in the network. In a data center environment, the topology is usually highly symmetric with many/all paths having equal

cost. In a long distance network, this is usually less the case for a variety of reasons (e.g. historic, fiber availability constraints, different distances between transit nodes, different roles ...). Hence between a given source and destination, a path is typically preferred over the others, while between the same source and another destination, a different path may be preferred.

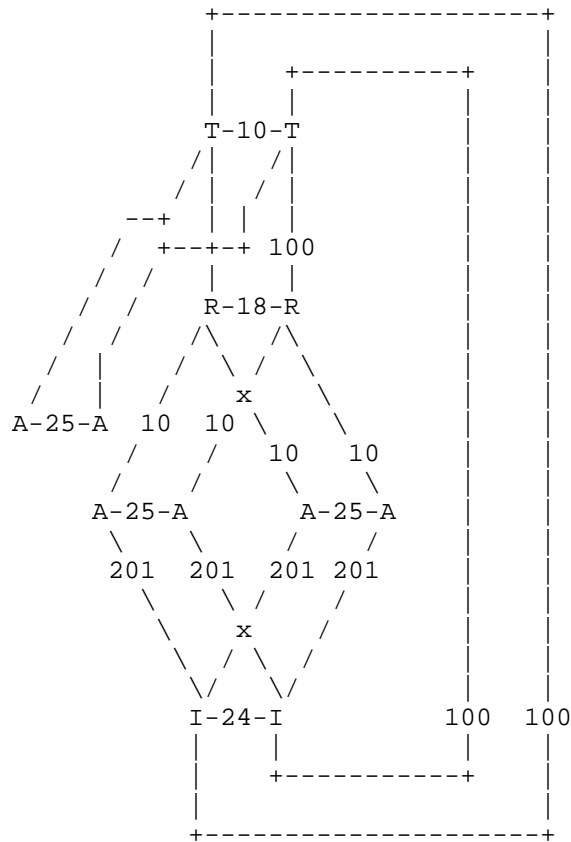


Figure 3: Explicit Routing topology

In the above topology, operator may want to enforce the following high level explicitly routed policies: - Traffic from A nodes to A nodes must not go through I nodes - Traffic from A nodes to I nodes must not go through R and T nodes with per-node

administrative tag, tag A can be configured on all A nodes, (similarly I, R, T), and then configure this single CSPF policy on all A nodes to avoid I nodes for path calculation.

6. Security Considerations

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

7. IANA Considerations

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

i) Per-Node-Admin-Tag Sub-TLV, Type: TBD

8. Acknowledgments

Many thanks to Les Ginsberg, Dhruv Dhody, Uma Chunduri for useful inputs. Thanks to Chris Bowers for providing useful inputs to remove ambiguity related to tag-ordering.

9. References

9.1. Normative References

[ISO10589]
"Intermediate system to Intermediate system intra-domain routing 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.

9.2. Informative References

[I-D.ietf-rtgwg-lfa-manageability]
Litkowski, S., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and p. psarkar@juniper.net, "Operational management of Loop Free Alternates", draft-ietf-rtgwg-lfa-manageability-04 (work in progress), August 2014.

- [I-D.ietf-rtgwg-remote-lfa]
Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote LFA FRR", draft-ietf-rtgwg-remote-lfa-08 (work in progress), September 2014.
- [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.
- [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.
- [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

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com

Bruno Decraene
Orange

Email: bruno.decraene@orange.com

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

Email: lizhenbin@huawei.com

Harish Raghuveer

Email: harish.r.prabhu@gmail.com

Network Working Group
Internet Draft

Category: Standard Track

L. Yong
W. Hao
D. Eastlake
Huawei
A. Qu
MetiaTek
J. Hudson
Brocade
U. Chunduri
Ericsson

Expires: April 2015

October 27, 2014

IS-IS Protocol Extension For Building Distribution Trees
draft-yong-isis-ext-4-distribution-tree-03

Abstract

This document proposes an IS-IS protocol extension to support IGP based multicast transport architecture and solution [IGP-MCAST].

Status of this document

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on April 27, 2015.

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.

Table of Contents

1. Introduction.....	3
1.1. Conventions used in this document.....	3
2. IS-IS Protocol Extension.....	3
2.1. RTADDR sub-TLV.....	3
2.2. RTADDRV6 sub-TLV.....	5
2.3. The Group Address Sub-TLV.....	6
3. Security Considerations.....	7
4. IANA Considerations.....	7
5. Acknowledgements.....	7
6. References.....	7
6.1. Normative References.....	7
6.2. Informative References.....	7

1. Introduction

This document proposes an IS-IS protocol extension to support IGP based multicast transport architecture and solution [IGP-MCAST].

1.1. Conventions used in this document

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. IS-IS Protocol Extension

2.1. RTADDR sub-TLV

This is a sub-TLV that is used in either a Router Capabilities TLV or an MT Capabilities TLV. Each RTADDR sub-TLV contains a root IPv4 address and multicast group addresses that associate to the tree. A router may use multiple RTADDR sub-TLVs to announce multiple root addresses and associated multicast groups with each root. RTADDR sub-TLV format is below.


```

+-----+
|subType=RTADDR | (1 byte)
+-----+
| Length | (1 byte)
+-----+
| Root IPv4 Address |
+-----+
|S|D|RESV | (1 byte)
+-----+
| Tree Priority | (1 byte)
+-----+
|Num of Groups | (1 byte)
+-----+
| Group Address (1) |
+-----+
| Group Mask (1) |
+-----+
~
+-----+
| GROUP Address (N) |
+-----+
| Group Mask (N) |
+-----+

```

Where:

subType: RTADDR (TBD)

Length: variable depending on the number of associated groups

Root IPv4 Address: IPv4 Address for a router that is a tree root

S bit: When set, the rooted tree for single area only. Otherwise, the rooted tree crosses multiple areas.

D bit: When set, the tree root is as of default tree root. Otherwise, the default tree is auto-calculated. [IGP-MCAST] When clear, the tree root is another distribution tree beside the default tree.

RESV: 6 reserved bits. MUST be sent as zero and ignored on receipt.

Tree Priority: An eight bit unsigned integer where larger magnitude means higher priority. Zero means no priority.

Num of Groups: the number of group addresses. When D bit sets, the number of group addresses is 0, which means that indicated tree root is the default tree root (supersede the auto-calculate one).

Group Address: IPv4 Address for the group

Group Mask: Group Mask: multicast groups mask. If the mask bit is a one, the

Group Address bit must match that corresponding bit in the packet destination address to be associated with the tree whose root is given.

One router may be the root for multiple trees. Each tree associates to a set of multicast groups. In this case, a router encodes multiple RTADDR sub-TLVs to announce root addresses, one for each root, in either a Router Capabilities TLV or an MT Capabilities TLV. The group address/mask in different sub-TLVs can overlap. See [IGP-MCAST] for details.

2.2. RTADDRV6 sub-TLV

This sub-TLV is used in an IPv6 network. It has the same format and usage except that the addresses are in IPv6.

ADDR sub-TLV applies to an IPv4 network and GIPV6-ADDR sub-TLV for IPv6 network.

When using a GIP-ADDR or GIPV6-ADDR sub-TLV for IGP multicast, the field VLAN-ID MUST set to zero and be ignored. Other field usage remains the same as [RFC7176]

3. Security Considerations

See Security Considerations in [IGP-MCAST].

4. IANA Considerations

IANA is requested to assign two new sub-TLV numbers for RTADDR and RTADDRV6 as specified in Sections 2.1 and 2.2. These sub-TLVs can be used under both the Router Capability (#242) and MT Capability (#144) TLVs. To avoid confusion, each sub-TLV should be assigned the same sub-Type number under each of these two TLVs.

5. Acknowledgements

Authors like to thank Mike McBride and Linda Dunbar for their valuable inputs.

[Editor note: the previous draft has been split into two drafts: draft-yong-isis-ext-4-distribution-tree-03 and draft-yong-rtgwg-igp-multicast-arch-00 based on AD and chair's suggestion.]

6. References

6.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC2119, March 1997.
- [RFC7176] Eastlake 3rd, D., Senevirathne, T., Ghanwani, A., Dutt, D., and A. Banerjee, "Transparent Interconnection of Lots of Links (TRILL) Use of IS-IS", RFC 7176, May 2014.

6.2. Informative References

- [IGP-MCAST] Yong, L., Hao, W., Eastlake, D., Qu A., Hudson, J., and Chunduri, "IGP Multicast Architecture", draft-yong-rtgwg-igp-multicast-arch-00, work in progress.

Authors' Addresses

Lucy Yong
Huawei USA

Phone: 918-808-1918
Email: lucy.yong@huawei.com

Weiguo Hao
Huawei Technologies
101 Software Avenue,
Nanjing 210012
China

Phone: +86-25-56623144
Email: haoweiguo@huawei.com

Donald Eastlake
Huawei
155 Beaver Street
Milford, MA 01757 USA

Phone: +1-508-333-2270
EMail: d3e3e3@gmail.com

Andrew Qu
MediaTek
San Jose, CA 95134 USA

Email: laodulaodu@gmail.com

Jon Hudson
Brocade
130 Holger Way
San Jose, CA 95134 USA

Phone: +1-408-333-4062
Email: jon.hudson@gmail.com

Uma Chunduri

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

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

