

Network Working Group
Request for Comments: 3786
Category: Informational

A. Hermelin
Montilio Inc.
S. Previdi
M. Shand
Cisco Systems
May 2004

Extending the Number of Intermediate System to Intermediate System (IS-IS) Link State PDU (LSP) Fragments Beyond the 256 Limit

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.

Abstract

This document describes a mechanism that allows a system to originate more than 256 Link State PDU (LSP) fragments, a limit set by the original Intermediate System to Intermediate System (IS-IS) Routing protocol, as described in ISO/IEC 10589. This mechanism can be used in IP-only, OSI-only, and dual routers.

Table of Contents

1.	Introduction	2
1.1.	Keywords	2
1.2.	Definitions of Commonly Used Terms	2
1.3.	Operation Modes	3
1.4.	Overview	4
2.	IS Alias ID TLV (IS-A)	5
3.	Generating LSPs	6
3.1.	Both Operation Modes	6
3.2.	Operation Mode 1 Additives	8
4.	Purging Extended LSP Fragments	10
5.	Modifications to LSP handling in SPF	10
6.	Forming Adjacencies	11
7.	Interoperating between extension-capable and non-capable ISs ..	11
8.	Security Considerations	12
9.	Acknowledgements	12

10. References	12
11. Authors' Addresses	13
12. Full Copyright Statement	14

1. Introduction

In the Intermediate System to Intermediate System (IS-IS) protocol, a system floods its link-state information in Link State PDU (LSP) Data Units, or LSPs for short. These logical LSPs can become quite large, therefore the protocol specifies a means of fragmenting this information into multiple LSP fragments. The number of fragments a system can generate is limited by ISO/IEC 10589 [ISIS-IS0] to 256 fragments, where each fragment's size is also limited. Hence, there is a limit on the amount of link-state information a system can generate.

A number of factors can contribute to exceeding this limit:

- Introduction of new TLVs and sub-TLVs to be included in LSPs.
- The use of LSPs to propagate various types of information (such as traffic-engineering information).
- The increasing number of destinations and AS topologies.
- Finer granularity routing, and the ability to inject external routes into areas [DOMAIN-WIDE].
- Other emerging technologies, such as optical, IPv6, etc.

This document describes mechanisms to relax the limit on the number of LSP fragments.

1.1. Keywords

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 BCP 14, RFC 2119 [BCP14].

1.2. Definitions of Commonly Used Terms

This section provides definitions for terms that are used throughout the text.

Originating System

A router physically running the IS-IS protocol. As this document describes methods allowing a single IS-IS process to advertise its LSPs as multiple "virtual" routers, the Originating System represents the single "physical" IS-IS process.

Normal system-id

The system-id of an Originating System.

Additional system-id

An Additional system-id that is assigned by the network administrator. Each Additional system-id allows generation of 256 additional, or extended, LSP fragments. The Additional system-id, like the Normal system-id, must be unique throughout the routing domain.

Virtual System

The system, identified by an Additional system-id, advertised as originating the extended LSP fragments. These fragments specify the Additional system-id in their LSP IDs.

Original LSP

An LSP using the Normal system-id in its LSP ID.

Extended LSP

An LSP using an Additional system-id in its LSP ID.

LSP set

Logical LSP. This term is used only to resolve the ambiguity between a logical LSP and an LSP fragment, both of which are sometimes termed "LSP".

Extended LSP set

A group of LSP fragments using an Additional system-id, and originated by the Originating System.

Extension-capable IS

An IS implementing the mechanisms described in this document.

1.3. Operation Modes

Two administrative operation modes are provided:

- Operation Mode 1 provides behavior that allows implementations that don't support this extension, to correctly process the extended fragment information, without any modifications. This mode has some restrictions on what may be advertised in the extended LSP fragments. Namely, only leaf information may be advertised in the extended LSPs.

- Operation Mode 2 extends the previous mode and relaxes its advertisement restrictions. Any link-state information may be advertised in the extended LSPs. However, it mandates a change to the way LSPs are considered during the SPF algorithm, in a way that is not compatible with previous implementations.

These modes are configured on a per-level and area basis. That is, all LSPs considered in the same SPF instance **MUST** use the same Mode. There is no restriction that an L1/L2 IS operates in the same mode, for both its L1 and L2 instances. It can use Mode 1 for its L1 LSPs, and Mode 2 for its L2 LSPs, or vice versa.

Mode 1 has the only advantage of being backwards compatible with older implementations. It does have restrictions which are considered drawbacks. Therefore, routers should operate in Mode 1 only if backwards compatibility is desired. Otherwise, it is recommended to run in Mode 2.

Routers **MAY** implement Operational Mode 2 without supporting running in Operational Mode 1. They will still interoperate correctly with routers that support both modes.

1.4. Overview

Using Additional system-ids assigned by the administrator, the Originating System can advertise the excess link-state information in extended LSPs under these Additional system-ids. It would do so as if other routers, or "Virtual Systems", were advertising this information. These extended LSPs will also have the specified limit on their LSP fragments; however, the Originating System may generate extended LSPs under numerous Virtual Systems.

For Operation Mode 1, 0-cost adjacencies are advertised from the Originating System to its Virtual System(s). No adjacencies (other than back to the Originating System) are advertised in the extended LSPs. As a consequence, the Virtual Systems are 'stub', meaning they can only be reached through their Originating System. Therefore, older implementations do not need modifications in order to correctly process these extended LSPs.

For both modes, each LSP (set) created by a node will contain in its fragment-0 a new TLV (IS Alias ID TLV) that contains the Normal system-id and PN Number of the Original LSP created by the router. Extension-capable ISs can then use this information and store the original and extended LSPs as one logical LSP.

The only sections that deal only with Mode 1 additions are 3.2, 3.2.1, and 3.2.2. All other sections relate to both modes.

2. IS Alias ID TLV (IS-A)

The proposed IS-A TLV allows extension-capable ISs to recognize all LSPs of an Originating System, and combine the original and extended LSPs for the purpose of SPF computation. It identifies the Normal system-id of the Originating System.

The proposed IS Alias ID TLV is type 24, and its format is as follows:

x CODE - 24.

x LENGTH - total length of the value field.

x VALUE -

	No. of Octets
+-----+ Normal system-id +-----+	6
+-----+ Pseudonode number +-----+	1
+-----+ Sub-TLVs length +-----+	1
+-----+ : Sub-TLVs : : : : : +-----+	0-247

Normal system-id

The Normal system-id of the LSP set, as described in section 1.2 of this document.

Pseudonode number

The Pseudonode number of the LSP set. LSPs with the same Normal system-id and Pseudonode number are considered in SPF as one logical LSP, as described in section 5 of this document.

Sub-TLVs length

Total length of all sub-TLVs.

Sub-TLVs

A series of tuples with the following format:

	No. of Octets
+-----+ Sub-type	1
+-----+ Length	1
+-----+ Value	0-245
+-----+ :	
+-----+ :	
+-----+	

Sub-type

Type of the sub-TLV

Length

Total length of the value field

Value

Type-specific TLV payload.

For an explanation on sub-TLV handling, see [ISIS-TE].

Without sub-TLVs, this structure consumes 8 octets per LSP set. This TLV **MUST** be included in fragment 0 of every LSP set belonging to an Originating System running in either Mode 1 or Mode 2. Currently, there are no sub-TLVs defined.

For a complete list of used IS-IS TLV numbers, see [ISIS-CODES].

3. Generating LSPs

3.1. Both Operation Modes

Under both modes, the Originating System **MUST** include information binding the Original LSP and the Extended ones. It can do this since it is trivially an extension-capable IS. This is to ensure other extension-capable routers correctly process the extra information in their SPF calculation. This binding is advertised via a new IS Alias ID TLV, which is advertised in all fragment 0 of Original and Extended LSPs.

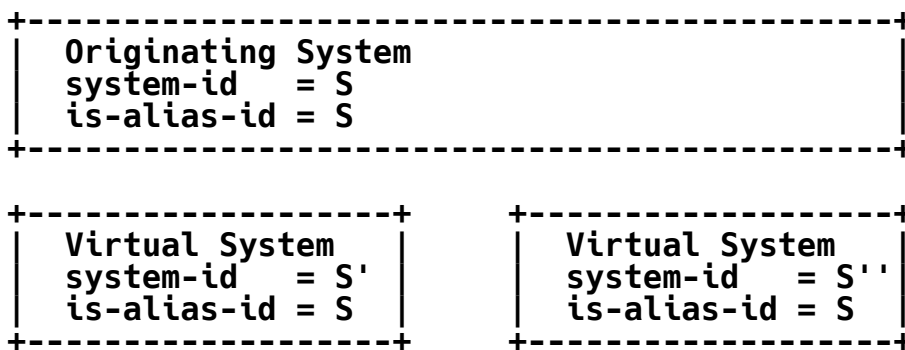


Figure 1. Advertising binding between all of a system's LSPs (both modes). S' and S'' are configured as Additional system-ids.

When new extended LSP fragments are generated, these fragments should be generated as specified in ISO/IEC 10589 [ISIS-ISO]. Furthermore, a system **SHOULD** treat its extended LSPs the same as it treats its original LSPs, with the exceptions noted in the following sections. Specifically, creating, flooding, renewing, purging and all other operations are similar for both Original and Extended LSPs, unless stated otherwise. The Extended LSPs will use one of the Additional system-ids configured for the router, in their LSP ID.

Extended LSPs fragment zero should be regarded in the same special manner as specified in ISO/IEC 10589 for LSPs with number zero, and should include the same type of extra information as specified in ISO/IEC 10589 and RFC 1195 [ISIS-IP]. So, for example, when a system reissues its LSP fragment zero due to an area address change, it should reissue all extended LSPs fragment zero as well.

An extended LSP fragment zero **MUST** be generated for every extended LSP set, to allow a router's SPF calculation to consider those fragments in that set. See section 5 for details.

3.1.1. The Attached Bits

The Attached (ATT) bits **SHOULD** be set to zero for all four metric types, on all Extended LSPs. This is due to the following: if a Virtual System is reachable, so is its Originating System. It is preferable, then, that an L1 IS chooses the Originating System and not the Virtual System as its nearest L2 exit point, as connectivity to the Virtual System has a higher probability of being lost (as a result of the extended LSP no longer being advertised). This could cause unnecessary computations on some implementations.

3.1.2. The Partition Repair Bit

The Partition Repair (P) bit SHOULD be set to zero on all extended LSPs. This is for the same reasons as for the Attached bits.

3.1.3. ES Neighbors TLV

ISO/IEC 10589 [ISIS-ISO] section 7.3.7 specifies inserting an ES Neighbor TLV in L1 LSPs, with the system ID of the router. RFC 1195 [ISIS-IP] relieves IP-only routers of this requirement. However, for routers that do insert this ESN TLV in L1 LSPs (whether IP-only or OSI-capable), then in an extended LSP, the ESN TLV should include the relevant Additional system-id. Furthermore, OSI-capable routers should accept packets destined for this Additional system-id.

3.1.4. Overload Bit

The overload bit should be set consistently across all LSPs, original and extended, belonging to an Originating System, and should reflect the Originating System's overload state.

3.1.5. Other Fields and TLVs

Other fields and TLVs not mentioned above remain the same, both for original and extended LSPs.

3.2. Operation Mode 1 Additions

The following additions apply only to routers generating LSPs in Mode 1. Routers, which are configured to operate in Operation Mode 2, SHOULD NOT apply these additions to their advertisements.

Under Operation Mode 1, adjacencies from the Originating System to its Virtual Systems are advertised using the standard neighbor TLVs. The metric for these connections MUST be zero, since the cost of reaching a Virtual System is the same as the cost of reaching its Originating System.

To older implementations, Virtual Systems would appear reachable only through their Originating System, hence loss of connectivity to the Originating System means loss of connectivity to all of its information, including that advertised in its extended LSPs. Furthermore, the cost of reaching information advertised in non-extended LSPs is the same as the cost of reaching information advertised in the new extended LSPs, with an additional hop.

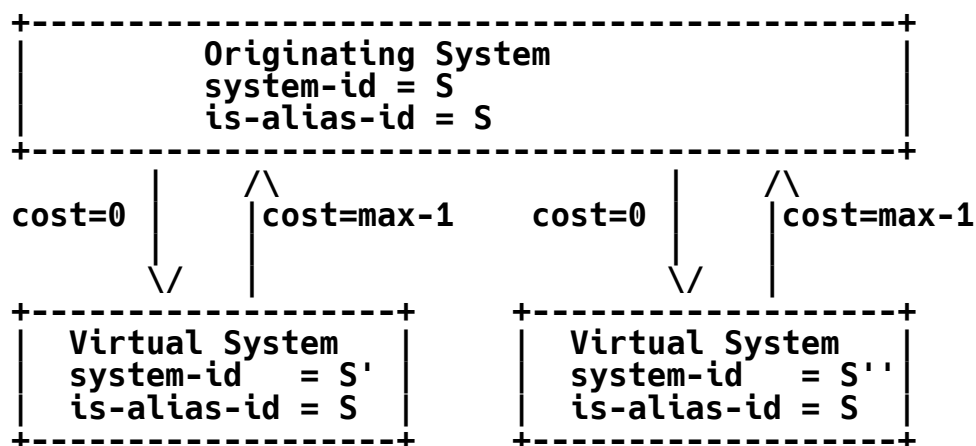


Figure 2. Advertising connections to Virtual Systems under Operation Mode 1. S' and S'' are configured as Additional system-ids.

Under Operation Mode 1, only "leaf" information, i.e., information that serves only as leaves in a shortest path tree, can be advertised in extended LSPs.

When an Extended LSP belonging to Additional system-id S' is first created, the Original LSP MUST specify S' as a neighbor, with metric set to zero. This is in order to consider the cost of reaching the Virtual System S' the same as the cost of reaching its Originating System. Furthermore, the Extended LSP MUST specify the Normal system-id as a neighbor. The metric SHOULD be set to MaxLinkMetric - 1 (this is only for uniformity purpose, any metric greater than zero is acceptable). This is in order to satisfy the two-way connectivity check on other routers. Where relevant, this adjacency SHOULD be considered as point-to-point.

Note, that the restriction specified in ISO/IEC 10589 section 7.2.5 holds: if an LSP Number zero of the Originating System is not present, none of that system's neighbor entries would be processed during SPF, hence none of its extended LSPs would be processed as well.

3.2.1. IS Neighbors TLV (Mode 1 Only)

An Extended LSP must specify only the Originating System as a neighbor, with the metric set to (MaxLinkMetric - 1). Where relevant, this adjacency should be considered as point-to-point. Other neighbors MUST NOT be specified in an Extended LSP, because

those other neighbors would only specify the Originating System and not the Virtual System, and hence would not satisfy the bi-directionality check in the SPF computation.

3.2.2. Originating System in the Overload State in (Mode 1 Only)

If the Originating System is in the overload state, information in the extended LSPs will not be processed by other routers in their SPF computation. This is because in Mode 1, extended LSPs are reachable only through adjacencies from the Original LSP. If this LSP has set its OL bit, adjacencies will not be processed in the SPF computation.

This side effect should be taken into consideration when operating in Mode 1.

4. Purging Extended LSP Fragments

ISO/IEC 10589 [ISIS-ISO] section 7.3.4.4 note 25 suggests that an implementation keeps the number of LSP fragments within a certain limit based on the optimal (minimal) number of fragments needed. Section 7.3.4.6 also recommends that an IS purge its empty LSPs to conserve resources. These recommendations hold for the extended LSP fragments as well. However, an extended LSP fragment zero should not be purged until all of the fragments in its set (i.e., belonging to a particular Additional system-id), are empty as well. This is to ensure implementations consider the fragments in their SPF computations, as specified in section 7.2.5.

In Operational Mode 1, when all the extended LSP fragments of a particular Additional system-id S' have been purged, the Originating System SHOULD remove the neighbor information to S' from its original LSPs.

5. Modifications to LSP handling in SPF

This section describes modifications to the way extension-capable ISs handle LSPs for the SPF computation.

When considering LSPs of an extension-capable IS (identified by the inclusion of the IS Alias ID TLV), the original and extended LSPs are combined to form one large logical LSP. If the LSP belongs to an IS running Operational Mode 1, there might be adjacencies between the original and extended LSPs. These are trivially ignored (since when processing them the large logical LSP is already on PATHS), and does not complicate the SPF. Furthermore, this check should already be implemented (this scenario could occur on error, without this extension).

If LSP fragment 0 of the Original LSP set is missing or its RemainingLifetime is zero, all of the LSPs generated by that Originating System (Extended as well) MUST NOT be considered in the SPF. That is, the large logical LSP is not considered in the SPF. The original LSP fragments are identified when the is-alias-id value is the same as the system-id of those LSPs. If an LSP fragment 0 of an extended LSP set is missing or its RemainingLifetime is zero, only that LSP set MUST NOT be considered in the SPF. These rules are present to ensure consistent SPF results on Mode 1 and Mode 2 LSPs.

Note, that the above behavior is consistent with how previous implementations will interpret Mode 1 LSPs.

6. Forming Adjacencies

It should be noted, that an IS MUST use the system-id of the LSP that will include a neighbor, when forming an adjacency with that neighbor. That is, if a neighbor is to be included in extended LSP S', then S' should be used as the system-id in IS Hellos [3] and IS-IS Hellos when forming an adjacency with that neighbor. This is regardless of the Operational Mode. Of course, in Mode 1 this means that only the Normal system-id will be used when sending hellos.

7. Interoperating between extension-capable and non-extension-capable ISs.

In order to correctly advertise link-state information under Operation Mode 2, all ISs in an area must be extension-capable. However, it is possible to not upgrade every router in the network, if the extended information is not routing information, but rather data that is of use to only a subset of routers (e.g., optical switches using IS-IS could carry optical-specific information in extended LSPs)

If a live network contains routers exceeding the 256 fragment limit, and for some reason the upgrade has to be done incrementally, it is possible to transition the network, using the following steps:

- Upgrade the routers, one-by-one, to run this extension in Operation Mode 1. The other non-extension-capable routers will interoperate correctly.
- When all routers are extension-capable, configure them one-by-one to run in Operation Mode 2. All extension-capable routers interoperate correctly, regardless of what mode they are run in.

Implementations SHOULD support a configuration parameter controlling the LSP origination behavior. The default value of this parameter SHOULD correspond to the behavior described in [ISIS-IS0], i.e., neither of the two modes described in this document should be enabled without explicit configuration when the router software is upgraded with this extension.

8. Security Considerations

This document raises no new security issues for IS-IS.

9. Acknowledgments

The authors would like to thank Tony Li and Radia Perlman for helpful comments and suggestions on the subject.

10. References

10.1. Normative References

- [ISIS-IS0] "Intermediate System to Intermediate System Intra-Domain Routeing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002, Second Edition.
- [ISIS-IP] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.
- [ISIS-TE] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, May 2004.
- [BCP14] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

10.2. Informative References

- [DOMAIN-WIDE] Li, T., Przygienda, T. and H. Smit, "Domain-wide Prefix Distribution with Two-Level IS-IS", RFC 2966, October 2000.
- [ISIS-CODES] Przygienda, T., "Reserved Type, Length and Value (TLV) Codepoints in Intermediate System to Intermediate System", RFC 3359, August 2002.

11. Authors' Addresses

Amir Hermelin
Montilio Inc.
1 Maskit St.
POB 12253
Herzeliya, 46733
ISRAEL

Phone: +972 9 9511944
Fax: +972 9 9542430
EMail: amir@montilio.com

Stefano Previdi
Cisco Systems, Inc.
Via Del Serafico 200
00142 Roma
Italy

Phone: +39 06 5164 4491
EMail: sprevidi@cisco.com

Mike Shand
Cisco Systems
250, Longwater Avenue,
Green Park,
Reading,
RG2 6GB,
UK

Phone: +44 20 8824 8690
EMail: mshand@cisco.com

12. Full Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.