Network Working Group
Request for Comments: 2973
Category: Informational

R. Balay
CoSine Communications
D. Katz
Juniper Networks
J. Parker
Axiowave Networks

October 2000

IS-IS Mesh Groups

Status of this Memo

Copyright Notice

Abstract

This document describes a mechanism to reduce redundant packet transmissions for the Intermediate System to Intermediate System (IS-IS) Routing protocol, as described in ISO 10589. The described mechanism can be used to reduce the flooding of Link State PDUs (Protocol Data Units) (LSPs) in IS-IS topologies. The net effect is to engineer a flooding topology for LSPs which is a subset of the physical topology. This document serves to document the existing behavior in deployed implementations.

The document describes behaviors that are backwards compatible with implementations that do not support this feature.

Table of Contents

1. Overview

In ATM or frame relay networks Intermediate Systems are interconnected using virtual circuits (VCs) which are logical point-to-point links. Some organizations attach multiple Intermediate Systems to form a full "mesh" topology, where every pair of Intermediate Systems are connected by a point-to-point link. In such topologies, IS-IS protocol operation leads to redundant transmission of certain PDUs due to the flooding operation which is illustrated below.

When an Intermediate System gets a new Link State Protocol Data Unit (LSP), it stores it, and prepares to flood it out every circuit except the source circuit. This is done by setting SRM (Send Routing Message) bits held in the local copy of the LSP: there is an SRM for each circuit.

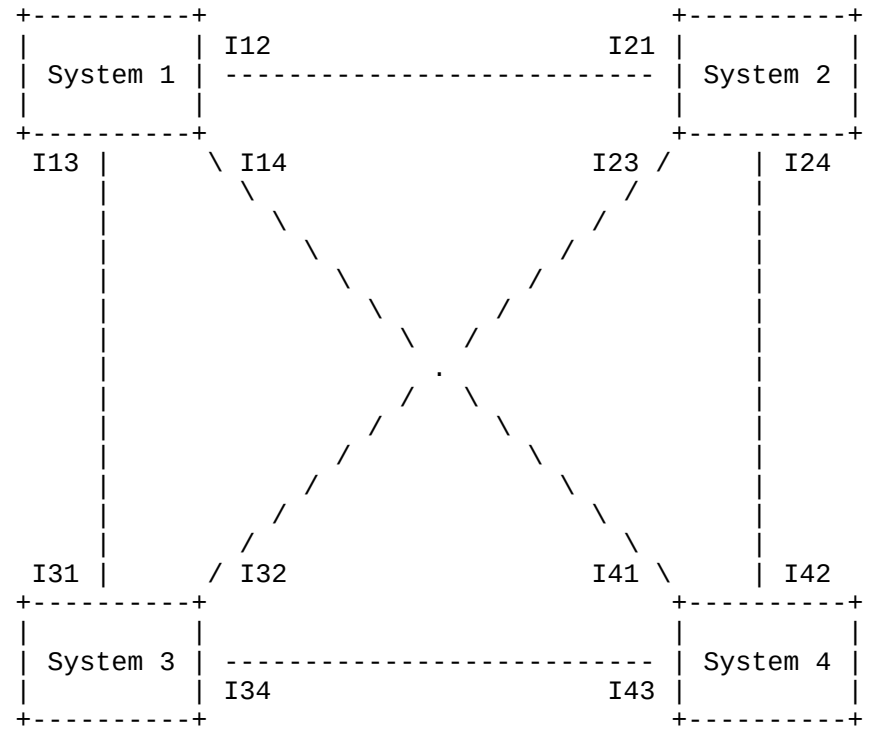

Figure 1. A four node full mesh topology

When System1 regenerates an LSP, it will flood the LSP through the network by marking the SRM bits for circuits I12, I14, and I13. In due course, it will send out the LSP on each circuit.

When System2 receives System1's LSP, it propagates the PDU according to section 7.2.14 of ISO 10589 [1]. It sets the SRM bits on circuits I23 and I24, to flood the LSP to System3 and System4. However, these Intermediate Systems will get the LSP directly from System1. In a full mesh of N Intermediate Systems, the standard protocol mechanism results in N-2 extra transmissions of each LSP, a waste of bandwidth and processing effort, with little gain in reliability.

Mesh groups provide a mechanism to reduce the flooding of LSPs.

2. Definitions of Mesh Groups

A mesh group is defined as a set of point-to-point circuits which provide full connectivity to a set of Intermediate Systems. Each circuit has two new attributes: meshGroupEnabled, which is in state {meshInactive, meshBlocked, or meshSet} and an integer variable meshGroup, which is valid only if the value of meshGroupEnabled attribute is 'meshSet'. Circuits that are in state 'meshSet' and that have the same value of meshGroup are said to be in the same mesh group.

LSPs are not flooded over circuits in 'meshBlocked' state, and an LSP received on a circuit C is not flooded out circuits that belong to C's mesh group.

Section 7.3.15.1 clause e.1.ii) of ISO 10589 [1] is modified as follows:

e.1.ii)
if the meshGroupEnabled attribute is 'meshSet' for the circuit C, set the SRMflag for that LSP for all circuits other than C whose meshGroupEnabled attribute is 'meshInactive'. Also set the SRMflag for all circuits in state 'meshSet' whose meshGroup attribute is not the same as C's.

if the meshGroupEnabled attribute is 'meshInactive' for circuit C, set the SRMflag for that LSP for all circuits other than C whose meshGroupEnabled attribute is not 'meshBlocked'.

For robust database synchronization when using mesh groups, the Complete Sequence Number PDUs (CSNPs) are sent periodically on point-to-point links with a mesh group meshEnabled or meshBlocked. Section 7.3.15.3 clause b) of ISO 10589 [1] is modified as follows:

b) If C is a point-to-point circuit (including non-DA DED circuits and virtual links), then

1) If the circuit's attribute is 'meshSet' or 'meshBlocked', then for each valid level, the IS will send a complete set of CSNPs as described for a Designated IS in section 7.3.15.3 clause a).

2) CSNPs are transmitted only at initialization on point-to-point links whose state is 'meshInactive'.

Use of mesh groups at an Intermediate System also modifies the behavior in transmission of generated LSPs. These LSPs are not required to be transmitted over circuits in state 'meshBlocked' at system startup or when the LSP is regenerated. The second sentence of Section 7.3.12 is modified to read:

> "For all the circuits whose meshGroupEnabled attribute is not 'meshBlocked', the IS shall set the SRMflags for that Link State PDU to propagate it on all these circuits. The IS shall clear the SRMflags for circuits whose meshGroupEnabled attribute is 'meshBlocked'."

Some of the transient transmission overhead can be reduced by having an Intermediate System not transmit its copies of the LSPs in database on a circuit start-up/restart if the circuit is 'meshBlocked'. The clause a) in the last part of Section 7.3.17 of ISO 10589, which refers to the point-to-point circuits, is modified as follows:

a) set SRMflag for that circuit on all LSPs if the meshGroupEnabled attribute of the circuit is not 'meshBlocked', and

Numbering of mesh groups provides the ability to divide a large full mesh topology into a smaller group of full mesh sub-topologies (mesh groups). These mesh groups are connected by "transit" circuits which are 'meshInactive', while the remaining circuits between the mesh groups are configured as 'meshBlocked' to reduce flooding redundancy. Use of numbering makes mesh groups more scalable.

3. Drawbacks of Mesh Groups

The mesh group feature described in this document is a simple mechanism to reduce flooding of LSPs in some IS-IS topologies. It relies on a correct user configuration. If a combination of user configuration and link failures result in a partitioned flooding topology, LSPs will not be sent in a timely fashion, which may lead to routing loops or black holes.

The concept of using numbered mesh groups also suffers from the complexity and reliance on static configuration, making the topologies brittle. Loosing a transit link can partition LSP flooding in unpredictable ways, requiring the periodic flooding of CSNPs to synchronize databases. In large networks, CSNPs become large and also consume bandwidth.

The authors are not aware of any networks that have deployed numbered mesh groups: instead, administrators set links to state 'meshBlocked' to prune the flooding topology (also known as "poor man's mesh groups").

Some improvements to mesh groups which have been suggested include:

a) To negotiate or check the mesh group attributes during initialization of an adjacency to verify that the two ends of every circuit hold identical values of the mesh state and mesh number.

b) Dynamic election of active transit links so that a topology could recover from failure of transit circuits.

c) Reduce the flooding of CSNPs by sending them periodically on some meshGroup circuits rather than all circuits.

d) Reduce the size of PDUs required by flooding of CSNPs by sending CSNP summaries: checksums or sequence numbers.

e) A related problem is the unneeded multiple transmissions of LSPs to neighbors that are connected via multiple links. The protocol could use the remote system ID of each adjacency and attempt to send a single copy of each LSP to a neighbor.

Any such improvements are outside the scope of this document, and may be the basis for future work.

4. Interoperation with Mesh Groups

Since mesh groups do not alter the content of packets, an Intermediate System that does not implement mesh groups will not see any different packets or new TLVs. The only impact will be that additional CSNPs will be seen on some point-to-point links. A conformant implementation can be expected to respond correctly to extra CSNPs.

5. Acknowledgments

The original idea for mesh groups is due to Dave Katz. Thanks to Tony Li, Tony Przygienda, Peter Livesey, and Henk Smit for helpful comments.

6. References

[1] ISO/IEC 10589, "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", June 1992.

7. Security Considerations

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

8. Authors' Addresses

Rajesh Balay
CoSine Communications, Inc
1200 Bridge Parkway
Redwood City, CA 94065

EMail: Rajesh.Balay@cosinecom.com

Dave Katz
Juniper Networks
385 Ravendale Drive
Mountain View, CA 94043

EMail: dkatz@juniper.net

Jeff Parker
Axiowave Networks,
100 Nickerson Road,
Marlborough, MA 01752

EMail: jparker@axiowave.com

9. Full Copyright Statement

Acknowledgement

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