IS-IS Working Group
Internet-Draft
Intended status: Informational
Expires: December 10, 2011

N. Shen
T. Li
Cisco Systems, Inc.
S. Amante
Level 3 Communications
M. Abrahamsson
Tele2

June 8, 2011

IS-IS Operational Enhancements for Network Maintenance Events
draft-ietf-isis-oper-enhance-00.txt

Abstract

This document describes an improved IS-IS neighbor management scheme which can be used to enhance operational experience in terms of convergence speed and finer control of neighbor cost over a LAN.

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 10, 2011.

Table of Contents

1. Introduction 3
1.1. Interface Shutdown Black Hole 3
1.2. LAN of Last Resort 3
1.3. Specification of Requirements 3

2. Sending Hellos with Fast Exit Notification 3

3. Pseudonodes with Non-zero Metrics 4
3.1. Operational Considerations 5

4. Security Considerations 5

5. Acknowledgements 5

6. Normative References 5

Authors' Addresses 6

1. Introduction

The IS-IS [ISO 10589] routing protocol has been widely used in Internet Service Provider IP/MPLS networks. Operational experience with the protocol, combined with ever increasing requirements for lossless operations have demonstrated some operational issues. This document describes those issues and some mechanisms for dealing with those issues. These mechanisms do involve implementation support, but do not require protocol changes.

1.1. Interface Shutdown Black Hole

One of these operationally problematic issues occurs when IS-IS is disabled on only one side of a link. This can result in a significant delay before neighbor(s) on the other end of the same link notice this change. In turn, this can result in several seconds during which traffic is blackholed, until the IS-IS neighbor(s) time out the adjacency and IS-IS reconverges.

1.2. LAN of Last Resort

Another issue stems from a situation when operators want to temporarily make an interface a "last resort" link for transit traffic. This is a straightforward, though cumbersome, operation to perform on a point-to-point link. Each device on the link is reconfigured to use very high metric. This causes traffic to divert to other links in the network. This same operation is more difficult on a multi-access LAN. There, the operator would have to increase the metric on each and every interface attached to the LAN, requiring the reconfiguration of a number of systems.

1.3. Specification of Requirements

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].

2. Sending Hellos with Fast Exit Notification

When an operator shuts down IS-IS on an interface, as described in Section 1.1, there is a significant interval before the change is noticed by all adjacencies and traffic is subsequently re-routed around this link. This delay is unnecessary, as neighbors should not have to wait for the adjacency to timeout, particularly when there exist alternate, viable, paths to downstream neighbors. This delay can be eliminated by carefully removing the adjacency between neighbors prior to actually disabling IS-IS on the interface.

An IS-IS adjacency uses the 3-way handshake protocol as defined in [ISO 10589] for multi-access LANs and [RFC3373] for point-to-point links. In both cases, the IS to IS Hello (IIH) message is used to establish and maintain the adjacency carries the system identifier of the adjacent systems. The receiving system expects to see its own system identifier listed. If not, then it must drop the adjacency.

An implementation that wishes to avoid the issue in Section 1.1 can do so by sending out a final IIH that includes no neighboring system IDs. When this is received, it should cause all neighbors to drop their adjacencies with the router that sent the IIH. This will also cause the systems to update their Link State Protocol Data Units (LSPs), flood them and reconverge to new paths. The technique is known as Fast Exit Notification.

This approach is not guaranteed. If the final IIH is lost on the link, then the neighboring systems will have to wait to time out the adjacency. Since this is unlikely, it is still a useful optimization. Implementations that require an even higher degree of assurance can retransmit the final IIH, possibly multiple times.

3. Pseudonodes with Non-zero Metrics

If an operator wishes to reconfigure a multi-access LAN so that it is only used as a resource of the last resort, then with current mechanisms, the operator must reconfigure each node on the LAN to give the LAN a high metric, as described in Section 1.2. It would be much easier for the operator if they could make a single configuration change that would cause IS-IS to treat the multi-access LAN as a link of last resort.

[ISO 10589] defines the pseudonode LSP as having a metric of zero. This implies that during the Shortest Path First (SPF) calculation, the metric for traversing the LAN is solely based on the metric set by the IS used to access the LAN. Thereby, the pseudonode does not contribute to the cost of traversing the LAN.

However, from the point of view of the SPF calculation, the metric in the pseudonode LSP does not have to be zero. Instead, the metric in a pseudonode LSP could be treated just like a normal LSP and have non-zero metrics to some or all of the systems on the LAN. This can then be used to simplify the operation for turning a LAN into a link of last resort. This could be done by having the Designated Intermediate System (DIS) change all of the metrics within the pseudonode LSP to a high value. This would effectively make the LAN look very 'expensive' and cause SPF calculations to converge to alternate links, if at all possible.

Because this change to the usage of the pseudonode LSP is in direct contradiction to the existing IS-IS specification, extreme caution is necessary. Implementations that would not interpret a non-zero pseudonode metric correctly might cause forwarding loops. As of this writing, we are actively surveying existing known implementations to determine if setting a non-zero metric in a pseudonode LSP will be interpreted properly.

This technique can also be used to divert traffic away from a subset of the nodes on the LAN. If the DIS increases the metric from the pseudonode to a subset of the systems on the LAN, then traffic will avoid exiting the LAN via that subset of systems.

3.1. Operational Considerations

An alternative is to allow any system to temporarily become the DIS, when it is directed to, and set a non-zero metric in the pseudonode LSP(s). This is beneficial because the operator would otherwise first have to determine the current DIS, access that system and reconfigure it. If an implementation wishes to support this, then it can provide an operation that both changes its priority on the LAN so that a node first becomes DIS and then generates a new pseudonode LSP with the non-zero metric.

If there is a concern that the DIS may change, it is prudent to define another node on the same LAN with the second highest priority for becoming DIS. This node can be configured to also set the metric in its pseudonode LSP appropriately if it becomes the new DIS.

4. Security Considerations

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

5. Acknowledgements

The authors would like to thank Mike Shand, Dave Katz, Guan Deng, Ilya Varlashkin, Jay Chen, Peter Ashwood-Smith and Les Ginsberg for their contributions.

6. Normative References

[ISO 10589]
ISO, "Intermediate system to Intermediate system routeing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network

Service (ISO 8473)", ISO/IEC 10589:2002.

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

[RFC3373] Katz, D. and R. Saluja, "Three-Way Handshake for Intermediate System to Intermediate System (IS-IS) Point-to-Point Adjacencies", RFC 3373, September 2002.

Authors' Addresses

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

Email: naiming@cisco.com

Tony Li
Cisco Systems, Inc.
225 West Tasman Drive
San Jose, CA 95134
USA

Email: tli@cisco.com

Shane Amante
Level 3 Communications
1025 Eldorado Blvd
Broomfield, CO 80021
USA

Email: shane@level3.net

Mikael Abrahamsson
Tele2

Email: swmike@swm.pp.se