Network Working Group
Request for Comments: 3787
Category: Informational

J. Parker, Ed.
Axiowave Networks

May 2004

Recommendations for Interoperable IP Networks using Intermediate System to Intermediate System (IS-IS)

Status of this Memo

Copyright Notice

Abstract

This document discusses a number of differences between the Intermediate System to Intermediate System (IS-IS) protocol used to route IP traffic as described in RFC 1195 and the protocol as it is deployed today. These differences are discussed as a service to those implementing, testing, and deploying the IS-IS Protocol to route IP traffic. A companion document describes the differences between the protocol described in ISO 10589 and current practice.

Table of Contents

1. Introduction

Interior Gateway Protocols such as IS-IS are designed to provide timely information about the best routes in a routing domain. The original design of IS-IS, as described in ISO 10589 [1] has proved to be quite durable. However, a number of original design choices have been modified. This document describes some of the differences between the protocol as described in RFC 1195 [2] and the protocol that can be observed on the wire today. A companion document describes the differences between the protocol described in ISO 10589 and current practice [8].

The key words "MUST", "MUST NOT", "SHOULD", "SHOULD NOT" and "MAY" in this document are to be interpreted as described in RFC 2119 [3].

2. Acknowledgments

This document is the work of many people, and is the distillation of over a thousand mail messages. Thanks to Vishwas Manral, who pushed to create such a document. Thanks to Danny McPherson, the original editor, for kicking things off. Thanks to Mike Shand, for his work in creating the protocol, and his uncanny ability to remember what everything is for. Thanks to Micah Bartell and Philip Christian, who showed us how to document difference without displaying discord. Thanks to Les Ginsberg, Neal Castagnoli, Jeff Learman, and Dave Katz, who spent many hours educating the editor. Thanks to Radia Perlman, who is always ready to explain anything. Thanks to Satish Dattatri, who was tenacious in seeing things written up correctly, and to Bryan Boulton for his work on the IP adjacency issue. Thanks to Russ White, whose writing improved the treatment of every topic he touched. Thanks to Shankar Vemulapalli, who read several drafts with close attention. Thanks to Don Goodspeed, for his close reading of the text. Thanks to Michael Coyle for identifying the quotation from Jan L.A. van de Snepscheut. Thanks for Alex Zinin’s ministrations behind the scenes. Thanks to Tony Li and Tony Przygienda, who kept us on track as the discussions veered into the weeds. And thanks to all those who have contributed, but whose names I have carelessly left from this list.

3. Unused Features

Some features defined in RFC 1195 are not in current use.

3.1. Inter-Domain Routing Protocol Information TLV, Code 131

RFC 1195 defines an Inter-Domain Routing Protocol Information TLV, with code 131, designed to convey information transparently between boundary routers. TLV 131 is not used, and MUST be ignored if received.

3.2. Authentication TLV, Code 133

RFC 1195 defines an authentication TLV, code 133, which contains information used to authenticate the PDU. This TLV has been replaced by TLV 10, described in "IS-IS Cryptographic Authentication" [4]. TLV 133 is not used, and MUST be ignored.

4. Overload Bit

To deal with transient problems that prevent an IS from storing all the LSPs it receives, ISO 10589 defines an LSP Database Overload condition in section 7.3.19. When an IS is in Database Overload condition, it sets a flag called the Overload Bit in the non-pseudonode LSP number Zero that it generates. Section 7.2.8.1 of ISO 10589 instructs other systems not to use the overloaded IS as a transit router. Since the overloaded IS does not have complete information, it may not be able to compute the right routes, and routing loops could develop. However, an overloaded router may be used to reach End Systems directly attached to the router, as it may provide the only path to an End System.

The ability to signal reduced knowledge is so useful that the meaning of this flag has been overloaded. In a Service Provider’s network, when a router running BGP and IS-IS reboots, BGP might take more time to converge than IS-IS. Thus the router may drop traffic for destinations not yet learned via BGP. It is convenient to set the Overload Bit until BGP has converged, as described in "Intermediate System to Intermediate System (IS-IS) Transient Blackhole Avoidance" [6].

An implementation SHOULD use the Overload Bit to signal that it is not ready to accept transit traffic.

An implementation SHOULD not set the Overload bit in PseudoNode LSPs that it generates, and Overload bits seen in PseudoNode LSPs SHOULD be ignored. This is also discussed in the companion document on ISO interoperability [8].

RFC 1195 makes clear when describing the SPF algorithm for IP routers in section C.1.4 that directly connected IP subnetworks are reachable when an IS is overloaded.

Note that the End Systems neighbors of the system P includes IP reachable address entries included in the LSPs from system P.

When processing LSPs received from a router which has the Overload bit set in LSP number Zero, the receiving router SHOULD treat all IP reachability advertisements as directly connected and use them in its SPF computation.

Since the IP prefixes that an overloaded router announces will be treated as directly attached, an overloaded router SHOULD take care in selecting which routes to advertise in the LSPs it generates.

5. Migration from Narrow Metrics to Wide

The IS-Neighbors TLV (TLV 2) as defined in ISO 10589 and the IP Reachability TLV (TLV 128/TLV 130) as defined in RFC 1195 provide a 6 bit metric for the default link metric to the listed neighbor. This metric has proved too limited. The Extended IS-Neighbors TLV (TLV 22) and the Extended IP Reachability TLV (TLV 135) are defined in "IS-IS extensions for Traffic Engineering" [5]. The Extended IS-Neighbors TLV (TLV 22) defines a 24 bit metric, and the Extended IP Reachability TLV (TLV 135) defines a 32 bit metric for IP Networks and Hosts.

If not all devices in the IS-IS domain support wide metrics, narrow metrics MUST continue to be used. Once all devices in the network are able to support the new TLVs containing wide metrics, the network can be migrated to the new metric style, though care must be taken to avoid routing loops.

We make the following assumptions about the implementation:

(1) Each system can generate and understand both narrow and wide metrics.

(2) The implementation can run the SPF algorithm on an LSP DB with instances of both metric styles.

(3) If there are two metric styles for a link or IP prefix, it will pick one of them as the true cost for the link.

To compare the different variants of the narrow metric with wide metrics, we need an algorithm that translates External and Internal narrow metrics into a common integer range. Since we have different computations for the L1 and L2 routes, we only need to map metrics from a single level.

In RFC 1195 section 3.10.2, item 2c) states that the IP prefixes located in "IP External Reachability" with internal-metric and IP prefixes located in "IP Internal Reachability" with internal-metric have the same preference. As defined in "Domain-wide Prefix Distribution with Two-Level IS-IS", the Most Significant Bit on an L1 metric tells us if the route has been leaked down, but does not change the distance. Thus we will ignore the MSBit.

We interpret the default metric as an 7 bit quantity. Metrics with the external bit set are interpreted as metrics in the range [64..127]. Metrics with the external bit clear are interpreted as metrics in the range [0..63].

5.1. Transition Algorithm

To facilitate a smooth transition between the use of narrow metrics exclusively to the use of wide metrics exclusively, the following steps must be taken, in the order below.

(1) All routers advertise Narrow Metrics as defined in ISO 10589, and consider narrow metrics only in their SPF computation.

(2) Each system is configured in turn to send wide metrics as well as narrow metrics. The two metrics for the same link or IP prefix SHOULD agree.

(3) When all systems are advertising wide metrics, make any changes necessary on each system to consider Wide Metrics during the SPF, and change MaxPathMetric to 0xFE000000.

(4) Each system is configured in turn to stop advertising narrow metrics.

(5) When the network is only using wide metrics, metrics on individual links may be rescaled to take advantage of the larger metric.

5.2. Dealing with Non-Equal Metrics

The algorithm above assumes that the metrics are equal, and thus needs to make no assumption about which metric the SPF algorithm uses. This section describes the changes that should be made to the SPF algorithm when both Narrow and Wide metric styles should be considered. Using a common algorithm allows different implementations to compute the same distances independently, even if the wide and narrow metrics do not agree.

The standard SPF algorithm proceeds by comparing sums of link costs to obtain a minimal cost path. During transition, there will be more than one description of the same links. We resolve this by selecting the minimum metric for each link. This may give us a path with some links chosen due to a wide metric and some links chosen due to a narrow metric.

The description below is more complex than the implementation needs to be: the implementation may simply select the minimal cost neighbor in TENT, discarding paths to destinations we have already reached, as described in ISO 10589.

The variables MaxPathMetric and MaxLinkMetric SHOULD retain the values defined in Table 2 of section 8 of ISO 10589.

In C.2.5 Step 0 of the description of the SPF algorithm, section b)

d(N) = cost of the parent circuit of the adjacency N

If multiple styles of metric for the link are defined, the cost will be the minimum available cost for the circuit.

In C.2.5 Step 0 of the description of the SPF algorithm, section i)

d(N) = metric of the circuit

If multiple styles of metric for the link are defined, the cost will be the minimum available cost for the circuit.

In C.2.6 Step 1 of the description of the SPF algorithm, section a)

dist(P,N) = d(P) + metric(P,N)

If multiple styles of metric for the neighbor are defined, the cost will be the minimum available cost for the circuit.

6. Intermediate System Hello (ISH) PDU

The original intent of RFC 1195 was to provide a routing protocol capable of handling both CLNS and IPv4 reachability information. To allow CLNS Endstations (ES) to know that they are attached to a router, Intermediate Systems are required to send Intermediate System Hello PDUs (ISH) for End Stations when a point-to-point circuit comes up. Furthermore, an IS is not allowed to send Intermediate System to Intermediate System Hello PDUs (IIH) before receiving an ISH from a peer. This reduces routing protocol traffic on links with a single IS.

For this reason section 5.1 RFC 1195 states:

> "On point-to-point links, the exchange of ISO 9542 ISHs (intermediate system Hellos) is used to initialize the link, and to allow each router to know if there is a router on the other end of the link, before IS-IS Hellos are exchanged. All routers implementing IS-IS (whether IP-only, OSI-only, or dual), if they have any interfaces on point-to-point links, must therefore be able to transmit ISO 9542 ISHs on their point-to-point links."

Section 5.1 RFC 1195 reinforces the need to comply with section 8.2.4 of ISO 10589. However, in an IP Only environment, the original need for the ISH PDU is not present.

A multi-protocol IS that supports the attachment of CLNS ESs over Point to Point circuits must act in accordance with section 8.2.2 ISO 10589 when CLNS functionality is enabled.

An IP only implementation SHOULD issue an ISH PDU as described in section 8.2.3 of ISO 10589. This is to inter-operate with implementations which require an ISH to initiate the formation of an IS-IS adjacency.

An IP Only implementation may issue an IIH PDU when a point to point circuit transitions into an "Up" state to initiate the formation of an IS-IS adjacency, without sending an ISH PDU. However, this may not inter-operate with implementations which require an ISH for adjacency formation.

An IS may issue an IIH PDU in response to the receipt of an IIH PDU in accordance with section 8.2.5.2 ISO 10589, even though it has not received an ISH PDU.

7. The Attached Bit

In section 7.2.9.2 of ISO 10589, an algorithm is described to determining when the attachedFlag should be set on an intermediate system. Some implementations also allow the attachedFlag to be set on Intermediate Systems routing IP traffic when there is a default route in the local routing table, or when some other state is reached that implies a connection to the rest of the network.

8. Default Route

RFC 1195 states in section 1.3:

> Default routes are permitted only at level 2 as external routes (i.e., included in the "IP External Reachability Information" field, as explained in sections 3 and 5). Default routes are not permitted at level 1.

Because of the utility of the default route when dealing with other routing protocols and the ability to influence the exit point from an area, an implementation MAY generate default routes in Level 1.

9. Non-homogeneous Protocol Networks

RFC 1195 assumes that every deployment of IS-IS routers will support a homogeneous set of protocols. It anticipates OSI only, IP only, or dual OSI and IP routers. While it allows mixed areas with, for example, both pure IP and Dual IP and OSI routers, it allows only IP traffic in such domains, and OSI traffic only when pure OSI and Dual IP and OSI routers are present. Thus it provides only lowest common denominator routing.

RFC 1195 also requires the inclusion of the Protocol Supported TLV with code 129 in IIH and ISH PDUs and in LSP number Zero. IP capable routers MUST generate a Protocol Supported TLV, and MUST include the IP protocol as a supported protocol. A router that does not include the Protocols Supported TLV may be assumed to be a pure OSI router and can be interpreted as implicitly "advertising" support for the OSI protocol.

The requirements of RFC 1195 are ample if networks adhere to this restriction. However, the behavior of mixed networks that do not follow these guidelines is not well defined.

The ITU-T requires that SONET/SDH equipment running the IS-IS protocol must not form an adjacency with a neighbour unless they share at least one network layer protocol in common. Unless this feature is present in every IS in the SONET or SDH DCN network the network may not function correctly. Implementors MAY include this feature if they wish to ensure interoperability with SONET and SDH DCN networks.

Definition of an interoperable strategy for resolving the problems that arise in non-homogeneous protocol networks remains incomplete. Members of the ITU are actively working on a proposal: see "Architecture and Specification of Data Communication Network", [7].

10. Adjacency Creation and IP Interface Addressing

RFC 1195 states that adjacencies are formed without regard to IP interface addressing. However, many current implementations refuse adjacencies based on interface addresses and related issues.

In section 4.2, RFC 1195 requires routers with IP interface addresses to advertise the addresses in an IP Interface Address TLV (132) carried in IIH PDUs. Some implementations will not interoperate with a neighbor router that does not include the IP Interface Address TLV. Further, some implementations will not form an adjacency on broadcast interfaces with a peer who does not share an interface address in some common IP subnetwork.

If a LAN contains a mixture of implementations, some that form adjacencies with all neighbors and some that do not, care must be taken when assigning IP addresses. If not all routers in a LAN are on the same IP subnet, it is possible that DIS election may fail, leading to the election of multiple DISs on a LAN, or no DIS at all. Even if DIS election succeeds, black holes can result because the IS-IS LAN transitivity requirements of section 6.7.3 ISO 10589 are not met.

Unnumbered point to point links do not have IP interface addresses, though they may have other IP addresses assigned to the routers. The IP address assigned to two routers that are neighbors on an unnumbered point to point link do not need to be related. However, some implementations will not form an adjacency on numbered point to point links if the interface addresses of each endpoint are not in the same IP subnetwork. This means that care must be taken in assigning IP interface addresses in all networks.

For an implementation to interoperate in a such mixed environment, it MUST include an IP Interface address (TLV 132) in its IIH PDUs. The network administrator should ensure that there is a common IP subnet assigned to links with numbered interfaces, and that all routers on each link have a IP Interface Addresses belonging to the assigned subnet.

11. Security Considerations

The clarifications in this document do not raise any new security concerns, as there is no change in the underlying protocol described in ISO 10589 [1] and RFC 1195 [2].

The document does make clear that TLV 133 has been deprecated and replaced with TLV 10.

12. References

12.1. Normative References

[1] 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.

[2] Callon, R., "OSI IS-IS for IP and Dual Environment," RFC 1195, December 1990.

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

[4] Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 3567, July 2003.

[5] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, May 2004.

[6] McPherson, D., "Intermediate System to Intermediate System (IS-IS) Transient Blackhole Avoidance", RFC 3277, April 2002.

12.2. Informative References

[7] ITU, "Architecture and Specification of Data Communication Network", ITU-T Recommendation G.7712/Y.1703, November 2001

[8] Parker, J., Ed., "Recommendations for Interoperable Networks using Intermediate System to Intermediate System (IS-IS)", RFC 3719, February 2004.

13. Author's Address

Jeff Parker
Axiowave Networks
200 Nickerson Road
Marlborough, Mass 01752
USA

EMail: jparker@axiowave.com

14. Full Copyright Statement

Intellectual Property

Acknowledgement

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