

IS-IS Working Group
Internet-Draft
Intended status: Standards Track
Expires: June 2, 2011

N. Shen
T. Li
Cisco Systems, Inc.
S. Amante
Level 3 Communications
M. Abrahamsson
Tele2
November 29, 2010

IS-IS Reverse Metric TLV for Network Maintenance Events
draft-amante-isis-reverse-metric-01

Abstract

This document describes an improved IS-IS neighbor management scheme which can be used to enhance network performance by allowing operators to quickly and accurately shift traffic away from a point-to-point or multi-access LAN interface by allowing one IS-IS router to signal to a second, adjacent IS-IS neighbor to adjust its IS-IS metric that should be used to temporarily reach the first IS-IS router during network maintenance events.

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 June 2, 2011.

Copyright Notice

Copyright (c) 2010 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

Internet-Draft

IS-IS Reverse Metric

November 2010

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	3
1.1.	Link Isolation Challenges	3
1.2.	IS-IS Reverse Metric	4
1.3.	Specification of Requirements	4
2.	IS-IS Reverse Metric TLV	4
3.	Elements of Procedure	6
3.1.	Processing Changes to Default Metric	6
3.2.	Processing Changes to Default Metric for Multi-Topology IS-IS	7
3.3.	Multi-Access LAN Procedures	8
3.4.	Order of Operations	9
3.5.	Operational Guidelines	9
4.	Reverse Metric TLV Example Use Case	10
5.	Operational Considerations	11
6.	Security Considerations	11
7.	IANA Considerations	11
8.	Acknowledgements	11
9.	References	12
9.1.	Normative References	12
9.2.	Informative References	12
	Authors' Addresses	12

[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 one issue and a new mechanism for improving it.

[1.1.](#) Link Isolation Challenges

During network maintenance events, operators substantially increase the IS-IS metric simultaneously on both devices attached to the same link. In doing so, the devices generate new Link State Protocol Data Units (LSP's) that are flooded throughout the network and cause all routers to gradually shift traffic onto alternate paths with very little, to no, disruption to in-flight communications by applications or end-users. When performed successfully, this allows the operator to confidently perform disruptive fault diagnosis and restoration on a link without disturbing ongoing communications in the network.

The challenge with the above solution are as follows. First, it is quite common to have routers with several hundred interfaces onboard and individual interfaces that are transferring several hundred Gigabits/second to Terabits/second of traffic. Thus, it is imperative that operators accurately identify the same point-to-point link on two, separate devices in order to increase (and, afterward, decrease) the IS-IS metric appropriately. Second, the aforementioned solution is very time consuming and even more error-prone to perform when its necessary to temporarily remove a multi-access LAN from the network topology. Specifically, the operator needs to configure ALL devices's that have interfaces attached to the multi-access LAN with an appropriately high IS-IS metric, (and then decrease the IS-IS metric to its original value afterward). Finally, with respect to multi-access LAN's, there is currently no method to bidirectionally isolate only a single node's interface on the LAN when performed more fine-grained diagnosis and repairs to the multi-access LAN.

In theory, use of a Network Management System (NMS) could improve the accuracy of identifying the appropriate subset of routers attached to either a point-to-point link or a multi-access LAN as well as signaling from the NMS to those devices, using a network management protocol, to adjust the IS-IS metrics on the pertinent set of interfaces. The reality is that NMS are, to a very large extent, not used within Service Provider's networks for a variety of reasons. In particular, NMS do not interoperate very well across different vendors or even separate platform families within the same vendor.

The risks of misidentifying one side of a point-to-point link or one

or more interfaces attached to a multi-access LAN and subsequently increasing its IS-IS metric are potentially increased latency, jitter or packet loss. This is unacceptable given the necessary performance requirements for a variety of applications, the customer perception for near lossless operations and the associated, demanding Service Level Agreement's (SLA's) for all network services.

[1.2.](#) IS-IS Reverse Metric

This document proposes that the routing protocol itself be the transport mechanism to allow one IS-IS router to advertise to an adjacent node on a point-to-point or multi-access LAN link a "reverse metric" in a IS-IS Hello (IIH) PDU. This would allow an operator to only configure a single router, set a "reverse metric" on a link and have traffic bidirectionally shift away from that link gracefully to alternate, viable paths.

[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.](#) IS-IS Reverse Metric TLV

The Reverse Metric TLV is composed of 1 octet for the Type, 1 octet that specifies the number of bytes in the Value field and a variable-length Value field. The Value field starts with a 1 octet field of

Flags followed by a 3 octet field containing an IS-IS Metric and, lastly, a 1 octet Traffic Engineering (TE) sub-TLV length field representing the length of a variable number of Extended Intermediate System (IS) Reachability sub-TLV's. If the 'S' bit in the Flags field is set to 1, then the Value field MUST also contain data of 1 or more Extended IS Reachability sub-TLV's.

The Reverse Metric TLV is optional. The Reverse Metric TLV may be present in any IS-IS Hello PDU. A sender MUST only transmit a single Reverse Metric TLV in a IS-IS Hello PDU.

TYPE: TBD
 LENGTH: variable (5 - 255 octets)
 VALUE:
 Flags (1 octet)
 Metric (3 octets)
 TE sub-TLV length (1 octet)

TE sub-TLV data (0 - 250 octets)

Flags

```

  0 1 2 3 4 5 6 7
+--+--+--+--+--+--+
| Reserved  |S|W|
+--+--+--+--+--+--+

```

Figure 1: Flags

The Reverse Metric TLV Type is TBD. Please refer to IANA Considerations, in [Section 7](#), for more details.

The Metric field contains a 24-bit unsigned integer of an IS-IS metric a neighbor SHOULD add to the existing, configured "default metric" contained within its IS Neighbors TLV or Extended IS Reachability TLV's for point-to-point links, or Pseudonode LSP by the Designated Intermediate System (DIS) for multi-access LAN's, back toward the router that originated this Reverse Metric TLV. Refer to "Elements of Procedure", below in [Section 3](#), for details of how an IS-IS router should process the Metric field in a Reverse Metric TLV.

There is currently only two Flag bits defined.

W bit (0x01): The "Whole LAN" bit is only used in the context of multi-access LAN's. When a Reverse Metric TLV is transmitted from a (non-DIS) node to the DIS, if the "Whole LAN" bit is set (1), then a DIS SHOULD add the received Metric value in the Reverse Metric TLV to each node's existing "default metric" in the Pseudonode LSP. If the "Whole LAN" bit is not set (0), then a DIS SHOULD add the received Metric value in the Reverse Metric TLV to the existing "default metric" in the Pseudonode LSP for the single node from whom the Reverse Metric TLV was received. Please refer to "Multi-Access LAN Procedures", in [Section 3.3](#), for additional details. The W bit MUST be unset (0) when a Reverse Metric TLV is transmitted in a IIH PDU onto a point-to-point link to an IS-IS neighbor.

S bit (0x02): The "TE sub-TLV" bit MUST be set (1) when an IS-IS router wishes to signal that its neighbor alter parameters contained in the neighbor's Traffic Engineering "Extended IS Reachability TLV", as defined in [\[RFC5305\]](#). This document defines that only the "Traffic Engineering Default Metric" sub-TLV, sub-TLV Type 18, may be sent toward neighbors in the Reverse Metric TLV, because that is used in Constrained Shortest Path First (CSPF) computations. Upon receipt of this TE sub-TLV in a Reverse Metric TLV, a node SHOULD add the received TE default metric to its existing, configured TE default metric within its Extended IS Reachability TLV. Use of other sub-

TLV's is outside the scope of this document.

The S bit MUST NOT be set (0) when an IS-IS router does not have TE sub-TLV's that it wishes to send to its IS-IS neighbor.

[3.](#) Elements of Procedure

[3.1.](#) Processing Changes to Default Metric

The Metric field, in the Reverse Metric TLV, is a "default metric" that will either be in the range of 0 - 63 when a "narrow" IS-IS metric is used (IS Neighbors TLV, Pseudonode LSP) [\[RFC1195\]](#) or in the range of 0 - ($2^{24} - 2$) when a "wide" Traffic Engineering metric value is used, (Extended IS Reachability TLV) [\[RFC5305\]](#). It is

RECOMMENDED that implementations, by default, place the appropriate maximum default metric value, 63 or $(2^{24} - 2)$, in the Metric field and TE Default Metric sub-TLV of the Reverse Metric TLV, since the most common use is to remove the link from the topology, except for use as a last-resort path.

In order to ensure that an individual TE link is used as a link of last resort during SPF computation, its metric MUST NOT be greater than or equal to $(2^{24} - 1)$ [[RFC5305](#)]. Therefore, a receiver of a Reverse Metric TLV MUST use the numerically smallest value of either the sum of its existing default metric and the Metric value in the Reverse Metric TLV or $(2^{24} - 2)$, as the default metric when updating its Extended IS Reachability TLV and TE default-metric sub-TLV's that it will then flood throughout the IS-IS domain, using normal IS-IS procedures. Likewise, originators of a Pseudonode LSP or IS Neighbors TLV MUST use the numerically smallest value of either the sum of its existing default metric and the Metric value it receives in a Reverse Metric TLV or 63 when updating the corresponding Pseudonode LSP or IS Neighbor TLV before they are flooded. This also applies when an IS-IS router is only configured or capable of sending a "narrow" IS-IS default metric, in the range of 0 - 63, but receives a "wide" Metric value in a Reverse Metric TLV, in the range of 64 - $(2^{24} - 2)$. In this case, the receiving router MUST use the maximum "narrow" IS-IS default metric, 63, as its IS-IS default metric value in its updated IS Neighbor TLV or Pseudonode LSP that it floods.

If an IS-IS router is configured to originate a TE Default Metric sub-TLV for a link, but receives a Reverse Metric TLV from its neighbor that does not contain a TE Default Metric sub-TLV, then the IS-IS router MUST add the value in the Metric field of the Reverse Metric TLV to its own TE Default Metric sub-TLV for that link. The IS-IS router should then flood the updated Extended IS Reachability TLV, including its updated TE Default Metric sub-TLV, using normal

IS-IS procedures.

Routers MUST scan the Metric value and TE sub-TLV's in all subsequently received Reverse Metric TLV's. If changes are observed by a receiver of the Reverse Metric TLV in the Metric value or TE Default Metric sub-TLV value, the receiving router MUST update its advertised IS-IS default metric or Traffic Engineering parameters in the appropriate TLV's, recompute its SPF tree and flood new LSP's to

other IS-IS routers, according to the recommendations outlined in [Section 3.4](#), Order of Operations, below.

If the router does not understand the Reverse Metric TLV or is explicitly configured to ignore received Reverse Metric TLV's, then it MUST NOT update the default metric in its IS Neighbors TLV, Extended IS Reachability TLV, TE Default Metric sub-TLV, Multi-Topology Intermediate Systems TLV or Pseudonode LSP nor execute other procedures that would result from acting on a Reverse Metric TLV, such as recomputing its SPF tree.

[3.2](#). Processing Changes to Default Metric for Multi-Topology IS-IS

The Reverse Metric TLV is applicable to Multi-Topology IS-IS (M-ISIS) [[RFC5120](#)] capable point-to-point links. If an IS-IS router is configured for M-ISIS it MUST send only a single Reverse Metric TLV in IIH PDU's toward its neighbor(s) on the designated link that is about to undergo maintenance. When an M-ISIS router receives a Reverse Metric TLV it MUST add the received Metric value to its default metric in all Extended IS Reachability TLV's for all topologies. If an M-ISIS router receives a Reverse Metric TLV with a TE Default Metric sub-TLV, then the M-ISIS router MUST add the received TE Default Metric value to each of its TE Default Metric sub-TLV's in all of its MT Intermediate Systems TLV's. If an M-ISIS router is configured to advertise TE Default Metric sub-TLV's for one or more topologies, but does not receive a TE Default Metric sub-TLV in a Reverse Metric TLV, then the M-ISIS router MUST add the value in Metric field of the Reverse Metric TLV to each of the TE Default Metric sub-TLV's for all topologies. The M-ISIS should flood its newly updated MT IS TLV's and recompute its SPF/CSPF accordingly.

Multi-Topology IS-IS [[RFC5120](#)] specifies there is no change to construction of the Pseudonode LSP, regardless of the Multi-Topology capabilities of a multi-access LAN. If any MT capable node on the LAN advertises the Reverse Metric TLV to the DIS, the DIS should act according to the "Multi-Access LAN Procedures" in [Section 3.3](#) to update, as appropriate, the default metric contained in the Pseudonode LSP. If the DIS updates the default metric in and floods a new Pseudonode LSP, those default metric values will be applied to all topologies during Multi-Topology SPF calculations.

[3.3](#). Multi-Access LAN Procedures

On a Multi-Access LAN, only the DIS SHOULD act upon information contained in a received Reverse Metric TLV. All non-DIS nodes MUST silently ignore a received Reverse Metric TLV.

In the case of multi-access LAN's, the "W" Flags bit is used to signal from a non-DIS to the DIS whether to change the metric and optionally Traffic Engineering parameters for all nodes in the Pseudonode LSP or a single node on the LAN, (the originator of the Reverse Metric TLV).

A non-DIS node, e.g.: Router B, attached to a multi-access LAN will send a Reverse Metric TLV with the W bit set to 0 to the DIS, when Router B wishes the DIS to add the Metric value to the default metric contained in the Pseudonode LSP specific to just Router B. Other non-DIS nodes, i.e.: Routers C and D, may simultaneously send a Reverse Metric TLV with the W bit set to 0 to request the DIS add their own Metric value to their default metric contained in the Pseudonode LSP. When the DIS receives a properly formatted Reverse Metric TLV with the W bit set to 0, the DIS MUST only add the default metric contained in its Pseudonode LSP for the specific neighbor that sent the Reverse Metric TLV.

It is possible for one node, Router A, to signal to the DIS with the W bit set to 1, in which case the DIS would add the Metric value in the Reverse Metric TLV to all neighbor adjacencies in the Pseudonode LSP and transmit a new Pseudonode LSP to all nodes in the IS-IS domain. Later, a second node on the LAN, Router B, could signal to the DIS with the W bit also set to 1. In this case, the DIS MUST use the highest source MAC address from IIH PDU's containing Reverse Metric TLV's it receives as the tie-breaker to determine the sole Reverse Metric TLV used as the source for the Metric value that will be added to the default metric for all nodes in the Pseudonode LSP. If the source MAC address was highest in IIH PDU's containing a Reverse Metric TLV received from Router B, then the DIS MUST add the Metric value to the default metric of all neighbors in its Pseudonode LSP and flood the LSP to all nodes in the IS-IS domain. On the other hand, if the DIS determines that Router A's IIH PDU's, containing Reverse Metric TLV's, have the highest source MAC address, then the DIS will ignore Router B's Reverse Metric TLV and continue to use the Metric value found in Router A's Reverse Metric TLV to add to the default metric of all neighbors in the Pseudonode LSP. When this occurs, the DIS MAY send a single syslog message or SNMP trap indicating that it has received a Reverse Metric TLV from a neighbor, but is ignoring it due to it being received from a neighbor with a lower MAC address.

Another scenario is that one node, Router A, may signal the DIS with the W bit set to 1. The DIS would add the Metric value to the default metric for all neighbors in the Pseudonode LSP and flood the LSP. Later, a second node on the LAN, Router B, could signal the DIS with the W bit set to 0, which indicates to the DIS that Router B is requesting the DIS only add the Metric value in the Reverse Metric TLV from Router B to the default metric for Router B in the Pseudonode LSP. The DIS MUST honor a neighbor's Reverse Metric TLV to update its individual default metric in the Pseudonode LSP even if the DIS receives prior or later requests to assert a Whole LAN metric from other nodes on the same LAN.

In all cases above, the DIS is MUST use 0 as the base default-metric value for each neighbor contained in the Pseudonode LSP to which the DIS will add the Metric value in the Reverse Metric TLV(s) it receives from neighbors on the LAN.

Local configuration on the DIS to adjust the default metric(s) contained in the Pseudonode LSP, as documented in [\[I-D.shen-isis-oper-enhance\]](#) MUST take precedence over received Reverse Metric TLV's.

[3.4.](#) Order of Operations

When an IS-IS router starts or stops generating a Reverse Metric TLV, it will go through a process of updating its own IS-IS metric and optionally Traffic Engineering parameters in its IS Neighbors TLV, Extended IS Reachability TLV or Pseudonode LSP, flooding updated LSP's (using normal IS-IS mechanisms), recompute its SPF/CSPF tree plus corresponding metrics to IP prefixes, update its FIB and begin advertising the Reverse Metric TLV in IIH PDU's toward its corresponding neighbor(s) on the appropriate link or LAN. Likewise, when IS-IS neighbor(s) start or stop receiving a Reverse Metric TLV, they will go through a similar process. It is critical that devices which implement the Reverse Metric TLV conduct this process in a deterministic order that minimizes the possibilities to generate temporary micro forwarding loops during a metric increase and decrease.

[3.5.](#) Operational Guidelines

A router MUST advertise a Reverse Metric TLV toward a neighbor only for the period during which it wants a neighbor to temporarily update its IS-IS metric or TE parameters.

During the period when a Reverse Metric TLV is used, IS-IS routers

that are generating and receiving a Reverse Metric TLV MUST NOT change their existing IS-IS metric or Traffic Engineering parameters

in their stored (e.g.: hard disk, etc.) configurations, since those parameters are carefully derived from off-line capacity planning tools and are difficult to restore to their original values.

Routers that receive a Reverse Metric TLV MAY send a syslog message or SNMP trap, in order to assist in rapidly identifying the node in the network that is asserting an IS-IS metric or Traffic Engineering parameters different from that which is configured locally on the device.

It is RECOMMENDED that implementations provide a capability to disable any changes to a node's, or individual interfaces of the node, default metric or Traffic Engineering parameters based upon receipt of properly formatted Reverse Metric TLV's.

4. Reverse Metric TLV Example Use Case

The following is a brief example illustrating one use case of the Reverse Metric TLV. In order to isolate a point-to-point link from the IS-IS network, an operator would configure one router, Router A, attached to a point-to-point link with a "Reverse Metric". This should not affect the configuration of the existing IS-IS default metric previously configured on the router's interface. Assuming Router A is using IS-IS Extensions for Traffic Engineering [[RFC5305](#)], this should trigger Router A to update its Traffic Engineering Default Metric sub-TLV in its own Extended IS Reachability TLV, recompute its SPF tree and corresponding metrics to IP prefixes in the IS-IS domain and begin the process of flooding a new LSP throughout the network. Router A would also begin transmitting a Reverse Metric TLV, with an appropriate Metric value, in an IIH PDU, to its adjacent neighbor, Router B. Upon receipt of the Reverse Metric TLV, Router B would add the received Metric or TE default metric sub-TLV value to its own Traffic Engineering Default Metric sub-TLV, recalculate its SPF tree and associated route topology as well as start flooding a new LSP containing the updated Extended IS Reachability TLV throughout the network. As nodes in the network receive the associated LSP's from Router A and B and recalculate a new SPF tree, and route topology, traffic should gracefully shift

onto alternate paths away from the A-B link; ultimately, after all nodes in the network recompute their SPF tree link A-B should only be used as a link of last-resort. The operator can inspect traffic counters on the A-B interface to determine if the link was successfully isolated from the topology and proceed with necessary fault diagnosis or maintenance of the associated link.

When the maintenance activity is complete, the operator would remove the reverse metric configuration from Router A, which would cease

advertisement of the Reverse Metric TLV in IIH PDU's to Router B. Both routers would revert to their originally configured IS-IS metric, recompute new SPF trees and corresponding metrics to IP prefixes and originate new LSP's. As the new LSP's are received and SPF is recalculated by nodes in the IS-IS domain, traffic should gradually shift back onto link A-B.

[5.](#) Operational Considerations

Since the Reverse Metric TLV may not be recognized by adjacent IS-IS neighbors, operators should inspect input and output traffic throughput counters on the local router to ensure that traffic has bidirectionally shifted away from a link before starting any maintenance activities.

[6.](#) Security Considerations

The enhancement in this document makes it possible for one IS-IS router to manipulate the IS-IS default metric or optionally Traffic Engineering parameters of adjacent IS-IS neighbors. Although IS-IS routers within a single Autonomous System nearly always reside under the control of a single administrative authority, it is highly RECOMMENDED that operators configure authentication of IS-IS PDU's to mitigate use of the Reverse Metric TLV as a potential attack vector, particularly on multi-access LAN's.

[7.](#) IANA Considerations

This document requests that IANA allocate from the IS-IS TLV

Codepoints Registry a new TLV, referred to as the "Reverse Metric" TLV, with the following attributes: IIH = y, LSP = n, SNP = n, Purge = n.

[8.](#) Acknowledgements

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

[9.](#) References

Shen, et al.	Expires June 2, 2011	[Page 11]
--------------	----------------------	-----------

Internet-Draft	IS-IS Reverse Metric	November 2010
----------------	----------------------	---------------

[9.1.](#) Normative References

- [ISO 10589]
ISO, "Intermediate system to Intermediate system routing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589:2002.
- [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.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", [RFC 5305](#), October 2008.

[9.2.](#) Informative References

- [I-D.shen-isis-oper-enhance]
Shen, N., Li, T., Amante, S., and M. Abrahamsson, "IS-IS

Operational Enhancements for Network Maintenance Events",
[draft-shen-isis-oper-enhance-00](#) (work in progress),
October 2010.

Authors' Addresses

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

Email: naiming@cisco.com

Shen, et al.

Expires June 2, 2011

[Page 12]

Internet-Draft

IS-IS Reverse Metric

November 2010

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