Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2016

Z. Li
N. Wu
Q. Zhao
Huawei Technologies
A. Atlas
C. Bowers
Juniper Networks
J. Tantsura
Ericsson

October 19, 2015

Intermediate System to Intermediate System (IS-IS) Extensions for Maximally Redundant Trees (MRT)
draft-ietf-isis-mrt-01

Abstract

This document describes extensions to IS-IS to support the distributed computation of Maximally Redundant Trees (MRT). Example uses of MRT include IP/LDP Fast-Reroute and global protection or live-live for multicast traffic. The extensions indicate what MRT profile(s) each router supports. Different MRT profiles can be defined to support different uses and to allow transition of capabilities. An extension is introduced to flood MRT-Ineligible links, due to administrative policy. This document also defines the Controlled Convergence sub-TLV, which is useful for MRT FRR as well as other applications.

Copyright Notice

Table of Contents

1. Introduction

The IS-IS protocol is specified in [ISO10589], with extensions for supporting IPv4 and IPv6 specified in [RFC1195] and [RFC5308]. Each Intermediate System (IS) (router) advertises one or more IS-IS Link State Protocol Data Units (LSPs) with routing information. Each LSP is composed of a fixed header and a number of tuples, each consisting of a Type, a Length, and a Value. Such tuples are commonly known as TLVs, and are a good way of encoding information in a flexible and extensible format.

[I-D.ietf-rtgwg-mrt-frr-architecture] gives a complete solution for IP/LDP fast-reroute using Maximally Redundant Trees (MRT) to provide alternates. This document describes signalling extensions for supporting MRT-FRR in an IS-IS routing domain.

2. Requirements Language

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].

3. Terminology

Redundant Trees (RT): A pair of trees where the path from any node X to the root R along the first tree is node-disjoint with the path from the same node X to the root R along the second tree. These can be computed in 2-connected graphs.

Maximally Redundant Trees (MRT): A pair of trees where the path from any node X to the root R along the first tree and the path from the same node X to the root R along the second tree share the minimum number of nodes and the minimum number of links. Each such shared node is a cut-vertex. Any shared links are cut-links. Any RT is an MRT but many MRTs are not RTs.

MRT Island: From the computing router, the set of routers that support a particular MRT profile and are connected via MRT- eligible links.

GADAG: Generalized Almost Directed Acyclic Graph - a graph which is the combination of the ADAGs of all blocks. Transforming a network graph into a GADAG is part of the MRT algorithm.

MRT-Red: MRT-Red is used to describe one of the two MRTs; it is used to describe the associated forwarding topology and MT-ID. Specifically, MRT-Red is the decreasing MRT where links in the GADAG

are taken in the direction from a higher topologically ordered node to a lower one.

MRT-Blue: MRT-Blue is used to describe one of the two MRTs; it is used to describe the associated forwarding topology and MT-ID. Specifically, MRT-Blue is the increasing MRT where links in the GADAG are taken in the direction from a lower topologically ordered node to a higher one.

4. Overview of IS-IS Signalling Extensions for MRT

As stated in [I-D.ietf-rtgwg-mrt-frr-algorithm], it is necessary for each MRT-Capable router to compute MRT next hops in a consistent fashion. This is achieved by using the same MRT profile and selecting a common and unique GADAG root in the MRT Island which is connected by MRT-Eligible links. Each of these issues will be discussed in the following sections separately.

4.1. Supporting MRT Profiles

The contents and requirements of an MRT profile has been defined in [I-D.ietf-rtgwg-mrt-frr-architecture]. The parameters and behavioral rules contained in an MRT profile define one router's MRT capabilities. Based on common capabilities, one unified MRT Island is built.

The MRT-Capable router MUST advertise its corresponding MRT profiles by IS-IS protocol extension within IS-IS routing domain. The capabilities of the advertiser MUST completely conform to the profile it claimed, including the MT-IDs, the algorithm and the corresponding forwarding mechanism. This advertisement MUST have level scope. A given router MAY support multiple MRT profiles. If so, it MUST advertise each of these profiles in the corresponding IS-IS level.

The default MRT Profile is defined in [I-D.ietf-rtgwg-mrt-frr-architecture]. Its behavior is intended to support IP/LDP unicast and multicast Fast-Reroute. MRT-Capable routers SHOULD support the default MRT profile.

4.2. Selecting the GADAG Root

As per [I-D.ietf-rtgwg-mrt-frr-algorithm], a GADAG root MUST be selected for one MRT Island. An unique GADAG root in common among MRT Island routers is a necessity to do MRT computation. Since the selection of the GADAG root can affect the quality of paths for traffic sent over MRT Red and Blue trees, the GADAG Root Selection Priority and the GADAG Root Selection Policy gives a network operator the ability to influence the selection of the GADAG root.

Each MRT-Capable router MUST advertise its priority for GADAG root selection. A router can only have one priority in a given MRT Island. It can have multiple priorities for different MRT Islands it supports. Routers that are marked as overloaded([RFC3787]) are not qualified as candidate for root selection.

The GADAG Root Selection Policy (defined as part of an MRT profile) may make use of the GADAG Root Selection Priority value advertised in the MRT Profile in the IS-IS Router CAPABILITY TLV. For example, the GADAG Root Selection Policy for the default MRT profile is the following: Among the routers in the MRT Island and with the highest priority advertised, an implementation MUST pick the router with the highest Router ID to be the GADAG root.

When the current root is out of service or new router with higher priority joined into the MRT Island, the GADAG root MUST be re-selected. A new MRT computation will be triggered because of such a topology change.

4.3. Advertising MRT-Ineligible Links for MRT

For administrative or management reasons, it may be desirable to exclude some links from the MRT computation. In this scenario, MRT-Capable router MUST claim those MRT-Ineligible links are out of MRT Island scope. If such claim splits current MRT Island then MRT computation has to be done inside the modified MRT Island which the computing router belongs to.

4.4. Triggering an MRT Computation

A MRT Computation can be triggered through topology changes or MRT capability changes of any router in the MRT Island. It is always triggered for a given MRT Profile in the corresponding level. First, the associated MRT Island is determined. Then, the GADAG Root is selected. Finally, the actual MRT algorithm is run to compute the transit MRT-Red and MRT-Blue topologies. Additionally, the router MAY choose to compute MRT-FRR alternates or make other use of the MRT computation results.

Prefixes can be attached and detached and have their associated MRT-Red and MRT-Blue next-hops computed without requiring a new MRT computation.

5. MRT Capability Advertisement

An MRT-Capable router MUST identify its MRT capabilities through IS-IS Link State Packets(LSPs) in level scope.

5.1. MRT Profile sub-TLV in the IS-IS Router CAPABILITY TLV

A new MRT Profile sub-TLV is introduced into the IS-IS Router CAPABILITY TLV (TLV #242 defined in [RFC4971]) to advertise MRT capabilities. Since MRT has per level scope, the S-bit and D-bit of IS-IS Router CAPABILITY TLV MUST be set to zero. The structure of the MRT Profile sub-TLV is shown below:

TYPE: TBA-MRT-ISIS-1 (To Be Allocated by IANA)

LENGTH: 4

VALUE:

MT-ID (2 octets with 4 bits reserved)

MRT Profile ID (1 octet)

GADAG Root Selection Priority (1 octet)

```
                                           Number of octets
+-------------------------------+
|R |R |R |R |        MT-ID      |                2
+-------------------------------+
| MRT Profile ID                |                1
+-------------------------------+
| GADAG Root Selection Priority |                1
+-------------------------------+
```

MT-ID is a 12-bit field containing the multi-topology ID. As discussed in Section 5.3, this document specifies that the MT-ID field MUST be set to zero.

The MRT Profile ID represents the MRT profile this router supports.

The GADAG Root Selection Priority is the priority for selection as GADAG root. A router advertising the MRT Profile sub-TLV MUST specify a GADAG Root Selection Priority. The range of this priority is [0, 255]. The RECOMMENDED default value of the GADAG Root Selection Priority is 128. The GADAG Root Selection Policy defined as part of a given MRT profile determines how the GADAG Root Selection Priority value is used.

This sub-TLV can occur multiple times if a router supports multiple MRT profiles. This can happen during a network transition. Or it can be used to support different uses of MRT within the same network which may perform better with different profiles.

A given router SHOULD NOT advertise multiple MRT Profile sub-TLVs with the same MRT profile ID. If a router receives multiple MRT Profile sub-TLVs with the same MRT profile ID from the same originator, it MUST use one with the lowest value of GADAG Root Selection Priority.

5.1.1. A note on the M-bit in OSPF

[I-D.ietf-ospf-mrt] defines MRT signalling extensions for OSPF. In addition to the OSPF version of MRT Profile sub-TLV (which is carried in the OSPF Router Information LSA), it also defines the M-bit (which is carried in the OSPF Router-LSA.) As discussed in [I-D.ietf-ospf-mrt], the M-bit in the Router-LSA ensures that all routers in the area have up-to-date information if a router is downgraded to a software version that does not support MRT and the Router Information LSA.

The problematic scenario that requires the M-bit in the OSPF Router-LSA does not occur in IS-IS. The presence or absence of an IS-IS Router CAPABILITY TLV containing a given MRT Profile sub-TLV in the IS-IS Link State PDU provides enough information for all routers to determine which remote routers support MRT, even after a software version downgrade of remote routers. Therefore, this document does not define a corresponding M-bit for IS-IS.

5.2. MRT-Ineligible Link sub-TLV in the Extended IS Reachability TLV

As a matter of policy, an operator may choose to mark some links as ineligible for carrying MRT traffic. For instance, policy can be made to prevent fast-rerouted traffic from taking those links.

For a link to be marked as ineligible for use in the MRT calculation, it MUST be advertised including the MRT-Ineligible Link sub-TLV in the Extended IS Reachability TLV (TLV #22 defined in [RFC5305]) corresponding to the ineligible link. The MRT-Ineligible Link sub-TLV is a zero-length sub-TLV as shown below:

TYPE: TBA-MRT-ISIS-2 (To Be Allocated by IANA)

LENGTH: 0

VALUE: None

The presence of the zero-length MRT-Ineligible Link sub-TLV in the Extended IS Reachability TLV indicates that the associated link MUST NOT be used in the MRT calculation for the default topology for all profiles.

5.3. A Note on Multi-Topology Routing

[RFC5120] specifies how to support multi-topology routing in IS-IS. The current document specifies how to signal support for MRT in the default routing topology only. The format of the extensions in this document allows the MRT Profile sub-TLV and the MRT-Ineligible Link sub-TLV to be scoped to topologies with non-zero MT-ID at some point in future. However, this document restricts these extensions to apply only to the default topology. The MRT Profile sub-TLV is restricting by explicitly requiring the MT-ID field to be set to zero. The MRT-Ineligible Link sub-TLV is restricted by only allowing it to be included in the Extended IS Reachability TLV (TLV#22) which is scoped to the default topology, and not allowing it to appear TLV#222 (the topology-scoped version of the Extended IS Reachability TLV).

Lifting these restrictions is for further study. Note that the MRT signalling extensions in this document can co-exist with IS-IS multi-topology routing, with the limitation that MRT Red and Blue forwarding trees can only be built for the default topology.

The format of the Controlled Convergence sub-TLV (discussed below) also contains an MT-ID field, allowing a Controlled Convergence sub-TLV to be scoped to a particular topology. This document does not place restrictions on the MT-ID field in the Controlled Convergence sub-TLV. The Controlled Convergence sub-TLV has potential applications that are not limited to MRT, and a topology-scoped FIB compute/install time makes sense on its own.

6. Controlled Convergence sub-TLV in IS-IS Router CAPABILITY TLV

Section 12.2 of [I-D.ietf-rtgwg-mrt-frr-architecture] describes the need to wait for a configured or advertised period after a network failure to insure that all routers are using their new shortest path trees. Similarly, avoiding micro-forwarding loops during convergence [RFC5715] requires determining the maximum among all routers in the area of the worst-case route computation and FIB installation time. More details on the specific reasoning and need for flooding this value are given in [I-D.atlas-bryant-shand-lf-timers].

A new Controlled Convergence sub-TLV is introduced into the IS-IS Router CAPABILITY TLV (TLV #242 defined in [RFC4971]) to advertise the worst-case time for a router to compute and install all IS-IS routes in the level after a change to a stable network. This advertisement has per level scope, so the S-bit and D-bit of IS-IS Router CAPABILITY TLV MUST be set to zero. The advertisement is scoped by MT-ID, allowing a router supporting multi-topology routing to advertise a different worst-case FIB compute/install time for each

topology. This makes sense as the SPF computations and FIB installations for each IGP topology can potentially be done independently of one another, and may have different worst-case compute and install times.

The structure of the Controlled Convergence sub-TLV is shown below:

TYPE: TBA-MRT-ISIS-3 (To Be Allocated by IANA)

LENGTH: 3

VALUE:

MT-ID (2 octets with 4 bits reserved)

FIB compute/install time (1 octet)

```
                                           Number of octets
+--------------------------------+
|R |R |R |R |        MT-ID      |                 2
+--------------------------------+
|  FIB compute/install time      |                 1
+--------------------------------+
```

MT-ID is a 12-bit field containing the multi-topology ID.

The FIB compute/install time is the worst-case time the router may take to compute and install all IS-IS routes in the level after a change to a stable network. The value is an unsigned integer in units of milliseconds.

The FIB compute/install time value sent by a router SHOULD be an estimate taking into account network scale or real-time measurements, or both. Advertisements SHOULD be dampened to avoid frequent communication of small changes in the FIB compute/install time.

A router receiving the Controlled Convergence sub-TLV SHOULD estimate the network convergence time as the maximum of the FIB compute/ install times advertised by the routers in a level for a given MT-ID, including itself. In order to account for routers that do not advertise the Controlled Convergence sub-TLV, a router MAY use a locally configured minimum network convergence time as a lower bound on the computed network convergence time. A router MAY use a locally configured maximum network convergence time as an upper bound on the computed network convergence time.

7. Handling MRT Extensions

7.1. Advertising MRT extensions

MRT sub-TLVs are encapsulated in the Router Capability TLV and advertised using the LS-PDU with level-wide scope. MRT sub-TLVs are optional. If one router does not support MRT, it MUST NOT advertise those sub-TLVs.

Since the advertisement scope of the MRT sub-TLV is level-wide, the D-Bit and S-Bit of the Router Capability TLV MUST be set as 0 when it is advertised. If other sub-TLVs in the Router Capability TLV need different values for those two bits, then the router MUST originate multiple Router Capability TLVs with different bit values.

When MRT-related information is changed for the router or existing IS-IS LSP mechanisms are triggered for refreshing or updating, MRT sub-TLVs MUST be advertised if the router is MRT-Capable.

Based on administrative policy, it may be desirable to exclude certain links from the MRT computation. The MRT-Ineligible sub-TLV is used to advertise which links should be excluded. Note that an interface advertised as MRT-Ineligible by a router is ineligible with respect to all profiles advertised by that router.

7.2. Processing MRT extensions

The processing associated with MRT extensions SHOULD NOT significantly degrade IS-IS adjacency formation and maintenance or the routing calculation for normal shortest path routes.

MRT sub-TLVs SHOULD be validated like other sub-TLVs when received. MRT sub-TLVs SHOULD also be taken into account for checksum calculations and authentication.

Links advertised as MRT-ineligible using the MRT-Ineligible sub-TLV MUST be excluded from the MRT computation. The removal of individual links from the MRT Island can partition an MRT Island or it can significantly modify the construction of the GADAG and the result MRT Red and Blue trees. Therefore, all routers must perform the same computation using the same set of links.

8. Backward Compatibility

The MRT Profile sub-TLV, the MRT-Ineligible Link sub-TLV, and the Controlled Convergence sub-TLV defined in this document are compatible with existing implementations of IS-IS. Routers that do not support these extensions are expected to silently ignore them.

Router that do support these extensions have mechanisms to recognize routers that don't support these extensions (or are not configured to participate in MRT) and behave accordingly.

9. Implementation Status

[RFC Editor: please remove this section prior to publication.]

Please see [I-D.ietf-rtgwg-mrt-frr-architecture] for details on implementation status.

10. Security Considerations

The IS-IS extensions in this document do not introduce new security concerns.

11. Acknowledgements

The authors would like to thank Shraddha Hegde for her useful suggestions.

12. IANA Considerations

12.1. MRT Profile and Controlled Convergence sub-TLVs

IANA is requested to allocate values for the following sub-TLVs types in the IS-IS Router CAPABILITY TLV defined in [RFC4971]: the MRT Profile sub-TLV (TBA-MRT-ISIS-1) and Controlled Convergence sub-TLV (TBA-MRT-ISIS-3). The IANA registry for these sub-TLV types is displayed as "sub-TLVs for TLV 242" (http://www.iana.org/assignments/isis-tlv-codepoints/isis-tlv-codepoints.xhtml#isis-tlv-codepoints-242). The new entries in this registry are shown below.

Value	Description	Reference
TBA-MRT-ISIS-1	MRT Profile sub-TLV	[This draft]
TBA-MRT-ISIS-3	Controlled Convergence sub-TLV	[This draft]

12.2. MRT-Ineligible Link sub-TLV

IANA is requested to allocate a value for the following sub-TLVs type in the Extended IS Reachability TLV defined in [RFC5305]: MRT-Ineligible Link sub-TLV (TBA-MRT-ISIS-2). The IANA registry for these sub-TLV types is displayed as "Sub-TLVs for TLVs 22, 23, 141, 222, and 223" (http://www.iana.org/assignments/isis-tlv-codepoints/isis-tlv-codepoints.xhtml#isis-tlv-codepoints-22-23-141-222-223). The new entry in this registry is shown below.

Type	Description	22	23	141	222	223	Ref.
TBA-MRT-ISIS-2	MRT-Ineligible Link sub-TLV	y	y	n	n	n	[This draft]

13. References

13.1. Normative References

[I-D.ietf-rtgwg-mrt-frr-algorithm]
Envedi, G., Csaszar, A., Atlas, A., Bowers, C., and A. Gopalan, "Algorithms for computing Maximally Redundant Trees for IP/LDP Fast- Reroute", draft-ietf-rtgwg-mrt-frr-algorithm-06 (work in progress), October 2015.

[I-D.ietf-rtgwg-mrt-frr-architecture]
Atlas, A., Kebler, R., Bowers, C., Envedi, G., Csaszar, A., Tantsura, J., and R. White, "An Architecture for IP/LDP Fast-Reroute Using Maximally Redundant Trees", draft-ietf-rtgwg-mrt-frr-architecture-07 (work in progress), October 2015.

[RFC3787] Parker, J., Ed., "Recommendations for Interoperable IP Networks using Intermediate System to Intermediate System (IS-IS)", RFC 3787, DOI 10.17487/RFC3787, May 2004, <http://www.rfc-editor.org/info/rfc3787>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <http://www.rfc-editor.org/info/rfc5305>.

13.2. Infomative References

[I-D.atlas-bryant-shand-lf-timers]
K, A. and S. Bryant, "Synchronisation of Loop Free Timer Values", draft-atlas-bryant-shand-lf-timers-04 (work in progress), February 2008.

[I-D.ietf-ospf-mrt]
Atlas, A., Hegde, S., Bowers, C., Tantsura, J., and Z. Li, "OSPF Extensions to Support Maximally Redundant Trees", draft-ietf-ospf-mrt-01 (work in progress), October 2015.

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, DOI 10.17487/RFC1195, December 1990, <http://www.rfc-editor.org/info/rfc1195>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <http://www.rfc-editor.org/info/rfc4915>.

[RFC4971] Vasseur, JP., Ed., Shen, N., Ed., and R. Aggarwal, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, DOI 10.17487/RFC4971, July 2007, <http://www.rfc-editor.org/info/rfc4971>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <http://www.rfc-editor.org/info/rfc5120>.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <http://www.rfc-editor.org/info/rfc5308>.

[RFC5715] Shand, M. and S. Bryant, "A Framework for Loop-Free Convergence", RFC 5715, DOI 10.17487/RFC5715, January 2010, <http://www.rfc-editor.org/info/rfc5715>.

Authors' Addresses

Zhenbin Li
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: lizhenbin@huawei.com

Nan Wu
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: eric.wu@huawei.com

Quintin Zhao
Huawei Technologies
125 Nagog Technology Park
Acton, MA 01719
USA

Alia Atlas
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Email: akatlas@juniper.net

Chris Bowers
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: cbowers@juniper.net

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
USA

Email: jeff.tantsura@ericsson.com