MPLS Working Group   A. Atlas
Internet-Draft   K. Tiruveedhula
Intended status: Standards Track   C. Bowers
Expires: January 5, 2015   Juniper Networks
                                                          J. Tantsura
                                                             Ericsson
                                                         IJ. Wijnands
                                                  Cisco Systems, Inc.
                                                         July 4, 2014

        LDP Extensions to Support Maximally Redundant Trees
                   draft-atlas-mpls-ldp-mrt-01

Abstract

   This document specifies extensions to LDP to support the creation of
   label-switched paths for Maximally Redundant Trees (MRT). A prime
   use of MRTs is for unicast and multicast IP/LDP Fast-Reroute (MRT-
   FRR).

   The sole protocol extension to LDP is simply the ability to advertise
   an MRT Capability. This document describes that extension and the
   associated behavior expected for LSRs and LERs advertising the MRT
   Capability.

   MRT-FRR uses LDP multi-topology extensions and requires three
   different multi-topology IDs to be allocated from the LDP MT-ID
   space.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   This document describes the LDP signaling extension and associated
   behavior necessary to support the architecture that defines how IP/
   LDP Fast-Reroute can use MRTs [I-D.ietf-rtgwg-mrt-frr-architecture].
   It is necessary to read the architecture in
   [I-D.ietf-rtgwg-mrt-frr-architecture] to understand how and why the
   LDP extensions for behavior are needed.

   At least one common standardized algorithm, such as the lowpoint
   algorithm explained and fully documented in
   [I-D.ietf-rtgwg-mrt-frr-algorithm], is required so that the routers

supporting MRT computation consistently compute the same MRTs. LDP
depends on the IGP to compute the MRTs and alternates. Extensions to
OSPF are defined in [I-D.atlas-ospf-mrt]. Extension to IS-IS are
defined in [I-D.li-isis-mrt]

MRT can also be used to protect multicast traffic via either global
protection or local protection.[I-D.atlas-rtgwg-mrt-mc-arch] An MRT
path can be used to provide node-protection for mLDP traffic via the
mechanisms described in [I-D.wijnands-mpls-mldp-node-protection]; an
MRT path can also be use to provide link protection for mLDP traffic.

For each destination, IP/LDP Fast-Reroute with MRT (MRT-FRR) creates
two alternate destination-based trees separate from the primary next-
hop forwarding used during stable operation. LDP uses the multi-
topology extensions [I-D.ietf-mpls-ldp-multi-topology] to signal FECs
for these two new forwarding topologies, known as MRT-Blue and MRT-
Red.

In order to create MRT paths and support IP/LDP Fast-Reroute, a new
capability extension is needed for LDP. An LDP implementation
supporting MRT must also follow the described rules for originating
and managing FECs related to MRT, as indicated by their multi-
topology ID. Network reconvergence is described in
[I-D.ietf-rtgwg-mrt-frr-architecture] and the worst-cast network
convergence time can be flooded via the extension in Section 7 of
[I-D.atlas-ospf-mrt].

IP/LDP Fast-Reroute using MRTs can provide 100% coverage for link and
node failures in an arbitrary network topology where the failure
doesn't split the network. It can also be deployed incrementally; an
MRT Island is formed of connected supporting routers and the MRTs are
computed inside that island.

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

For ease of reading, some of the terminology defined in
[I-D.ietf-rtgwg-mrt-frr-architecture] is repeated here.

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 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 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. The two MRTs are
   referred to as MRT-Blue and MRT-Red.

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

MRT-Red: MRT-Red is used to describe one of the two MRTs; it is
   used to described 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 described 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.

Rainbow MRT: It is useful to have an MT-ID that refers to the
   multiple MRT topologies and to the default topology. This is
   referred to as the Rainbow MRT MT-ID and is used by LDP to reduce
   signaling and permit the same label to always be advertised to all
   peers for the same (MT-ID, Prefix).

4. Overview of LDP Signaling Extensions for MRT

Routers need to know which of their neighbors support MRT.
Supporting MRT indicates several different aspects of behavior, as
listed below.

1. Support for Multi-Topology (MT) - this MAY also be indicated via
    the Multi-Capability MT Capability
    [I-D.ietf-mpls-ldp-multi-topology].

2. Understand the Rainbow MRT MT-ID and apply the associated labels
    to all relevant MT-IDs.

3. Advertise the Rainbow MRT MT-ID to the appropriate neighbors for
    the associated prefix.

   4. If acting as LDP egress for a prefix in the default topology,
       also advertise and act as egress for the same prefix in MRT-Red
       and MRT-Blue.

   5. For a FEC learned from a neighbor that does not support MRT,
       originate FECS for MRT-Red and MRT-Blue with the same prefix.
       This MRT Island egress behavior is to support an MRT Island that
       does not include all routers in the area/level.

4.1. MRT Capability Advertisement

   It is not possible to support MRT without supporting the LDP multi-
   topology extensions, but it is possible that the only use of the
   multi-topology extensions is for MRT. In that case, a router MAY not
   negotiate the multi-topology capability and only negotiate the MRT
   Capability with its LDP peer. Negotiation of the MT capability is
   not required with negotiation of the MRT capability.

   [EDITOR NOTE: How do we deal with different abilities for IPv4 and
   IPv6? The MT capability has the Wildcard FEC to indicate this. Do
   we just assume??]

   A new MRT Capability Parameter TLV is defined, which is defined in
   accordance with LDP Capability definition guidelines[RFC5561].

   The LDP MRT capability can be advertised during the LDP session
   initialization or after the LDP session is established.
   Advertisement of the MRT capability indicates support of the
   procedures for establishing the MRT-Blue and MRT-Red LSP paths
   detailed in this document. If the peer has not advertised the MRT
   capability, then it indicates that LSR does not support MRT
   procedures.

   If a router advertises the LDP MRT capability to its peer, but the
   peer has not advertised the MRT capability, then the router MUST NOT
   advertise MRT-related FEC-label bindings to that peer, until that
   peer starts to advertise the MRT capability.

   The following is the format of the MRT Capability Parameter.

```
 0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |U|F| MRT Capability (IANA)     |      Length (= 1)             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |S| Reserved    |
 +-+-+-+-+-+-+-+-+
```
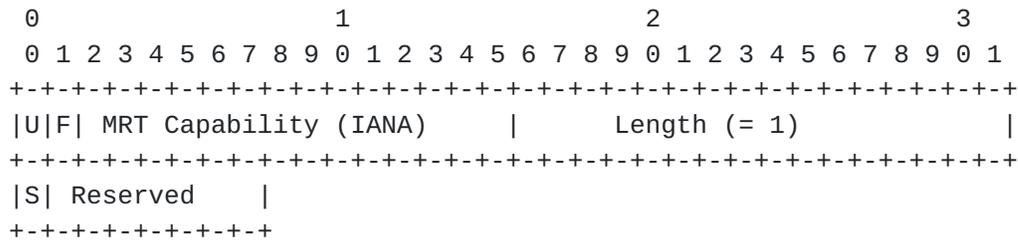
                      MRT Capability TLV Format

   Where:

   U- and F-bits: MUST be 1 and 0, respectively, as per Section 3.
      (Signaling Extensions) of LDP Capabilities [RFC5561].

   MRT Capability: TBA-MRT-LDP-1 (To Be Allocated by IANA)

   S-bit: MUST be 1 if used in LDP "Initialization" message. MAY be
      set to 0 or 1 in dynamic "Capability" message to advertise or
      withdraw the capability respectively.

   Length: The length (in octets) of TLV. Its value is 1.

4.2. Behavior Related to the Rainbow MRT MT-ID

   In Section 10.1 of [I-D.ietf-rtgwg-mrt-frr-architecture], the need to
   advertise different MPLS labels to different neighbors for the same
   FEC is described. This can be shortly summarized as either
   advertising MRT MT-ID differentiated labels to a neighbor or just
   advertising the same MPLS label for the default topology, for MRT-Red
   and MRT-Blue. MRT-supporting neighbors in the same domain as the
   default SPT next-hop get the differentiated MPLS labels; all other
   neighbors do not.

   A second use for the Rainbow MRT MT-ID is for an egress LER to send
   the Rainbow MRT MT-ID with an IMPLICIT_NULL label to indicate
   penultimate-hop-popping for all three types of FECs (IP Prefix FEC,
   MRT-Blue MT-IP Prefix FEC, and MRT-Red MT-IP Prefix FEC).

   The use of the Rainbow-FEC by the ABR for non-best-area
   advertisements is RECOMMENDED. An ABR MAY advertise the label for
   the default topology in separate MRT-Blue and MRT-Red advertisements.
   An LSR advertising the MRT capability MUST recognize the Rainbow MRT
   MT-ID and associate the advertised label with the specific prefix
   with the MRT-Red and MRT-Blue MT-IDs associated with all MRT Profiles
   that advertise LDP as the forwarding mechanism.

The value of the Rainbow MRT MT-ID (TBA-MRT-LDP-2) will be assigned
by IANA from the LDP MT-ID space. Prototype experiments have used
the value 3999.

4.3. MRT-Blue and MRT-Red FECs

To provide MRT support in LDP, the MT Prefix FEC is used.
[I-D.ietf-rtgwg-mrt-frr-architecture] contains the IANA request for
the MRT-Red and MRT-Blue MT-IDs associated with the Default MRT
Profile.

The MT Prefix FEC encoding is defined in
[I-D.ietf-mpls-ldp-multi-topology] and is used without alteration for
signaling MRT-Blue, MRT-Red and Rainbow MRT FECs.

5. LDP MRT FEC Advertisements

This sections describes how and when labels for MRT-Red and MRT-Blue
FECs are advertised. The associated LSPs must be created before a
failure occurs, in order to provide protection paths which are
immediately usable by a PLR.

5.1. Downstream Unsolicited Mode

If the upstream session is negotiated with the MRT capability, the
Egress LER advertises via a Rainbow MRT FEC an allocated MPLS label;
this may be Explicit Null, Implicit Null, or another value.

Based on the MRT algorithm [I-D.ietf-rtgwg-mrt-frr-algorithm], the
IGP computes the MRT-Red and MRT-Blue disjoint paths at Ingress and
Transit LSRs. Once the IGP computes the MRT-Red and MRT-Blue next-
hops, LDP will advertise the Label Mapping for the MRT-Blue and MRT-
Red FECs. If a label is received from a downstream LSR for an MRT-
Red or MRT-Blue FEC where the downstream LSR is capable of MRT, the
MRT-Red FEC or MRT-Blue FEC label is swapped according to the
received downstream label. An LSR may also choose to use the MRT-Red
or MRT-Blue path as an alternate for doing fast-reroute for the local
traffic.

When a downstream router is not capable of MRT, the LSR is an MRT
Island Border Router (IBR) and SHOULD advertise Label Bindings for
the MRT-Red FEC and MRT-Blue FEC as well as the associated normal
topology. The normal topology's primary next-hops will be used to
forward traffic received for the MRT-Red FEC or the MRT-Blue FEC
where the FEC's destination is outside the MRT Island. This
functionality is critical for partial deployment scenarios.

5.2. Downstream On Demand Mode

   After the IGP computes the MRT-Red and MRT-Blue paths, the IGP MAY
   also decide to use either the MRT-Red or MRT-Blue path as a fast-
   reroute alternate for the particular FEC. If so, then when in
   Downstream On Demand Mode, the LSR sends a Label Request for either
   the MRT-Red or MRT-Blue FEC to the downstream LSR. The downstream
   LSR responds by either sending a Label Mapping if available or by
   sending a Label Request to its downstream LSR. Once a Label Mapping
   is received, the associated label may be used as a fast-reroute
   alternate to forward IP and LDP traffic.

   A Label Mapping may be available in the following circumstances:

   o The LSR is acting as Egress

   o A Label Mapping was already received from its downstream router

   o A Label Mapping for the default topology FEC was received and the
      downstream router is not capable of MRT or is in a different MRT
      Island.

5.3. Inter-Area

   As discussed in Section 4.2, the Rainbow MRT FEC is defined to
   facilitate signaling the same label for multiple topologies.
   Section 9 of [I-D.ietf-rtgwg-mrt-frr-architecture] recommends that
   traffic leaving an OSPF area or IS-IS level SHOULD use the default
   topology's shortest-path-tree next-hops instead of remaining on the
   MRT-Red or MRT-Blue paths. If an LDP peer is in the same OSPF area
   or IS-IS level as the primary next-hop, then LDP SHOULD advertise
   different label values for a given set of MRT-Red FEC, MRT-Blue FEC,
   and FEC, unless Explicit-Null or Implicit-Null is appropriate. If an
   LDP peer is in a different OSPF area or IS-IS level from the primary
   next-hop, then LDP SHOULD either advertise the same label value for
   the given set of MRT-Red FEC, MRT-Blue FEC, and FEC or advertise a
   single label for the Rainbow MRT FEC, whose behavior is defined in
   Section 4.2.

6. Security Considerations

   This LDP extension is not believed to introduce new security
   concerns. It relies upon the security architecture already provided
   for LDP.

7. IANA Considerations

Please allocate a value for the new LDP Capability TLV from the LDP registry "TLV Type Name Space": MRT Capability TLV (TBA-MRT-LDP-1).

Please allocate a value from the LDP Multi-Topology (MT) ID Name Space [I-D.ietf-mpls-ldp-multi-topology]: Rainbow MRT MT-ID (TBA-MRT-LDP-2).

8. Acknowledgements

The authors would like to thank Ross Callon for his suggestions.

9. References

9.1. Normative References

[I-D.ietf-mpls-ldp-multi-topology]
          Zhao, Q., Raza, K., Zhou, C., Fang, L., Li, L., and D.
          King, "LDP Extensions for Multi Topology", draft-ietf-
          mpls-ldp-multi-topology-12 (work in progress), April 2014.

[I-D.ietf-rtgwg-mrt-frr-architecture]
          Atlas, A., Kebler, R., Bowers, C., Enyedi, G., Csaszar,
          A., Tantsura, J., Konstantynowicz, M., and R. White, "An
          Architecture for IP/LDP Fast-Reroute Using Maximally
          Redundant Trees", draft-rtgwg-mrt-frr-architecture-04
          (work in progress), July 2014.

[RFC5561] Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL.
          Le Roux, "LDP Capabilities", RFC 5561, July 2009.

9.2. Informative References

[I-D.atlas-ospf-mrt]
          Atlas, A., Hegde, S., Bowers, C., and J. Tantsura, "OSPF
          Extensions to Support Maximally Redundant Trees", draft-
          atlas-ospf-mrt-02 (work in progress), July 2014.

[I-D.atlas-rtgwg-mrt-mc-arch]
          Atlas, A., Kebler, R., Wijnands, I., Csaszar, A., and G.
          Envedi, "An Architecture for Multicast Protection Using
          Maximally Redundant Trees", draft-atlas-rtgwg-mrt-mc-
          arch-02 (work in progress), July 2013.

[I-D.ietf-rtgwg-mrt-frr-algorithm]
          Enyedi, G., Csaszar, A., Atlas, A., Bowers, C., and A.
          Gopalan, "Algorithms for computing Maximally Redundant
          Trees for IP/LDP Fast-Reroute", draft-rtgwg-mrt-frr-
          algorithm-01 (work in progress), July 2014.

[I-D.li-isis-mrt]
          Li, Z., Wu, N., Zhao, Q., Atlas, A., Bowers, C., and J.
          Tantsura, "Intermediate System to Intermediate System (IS-
          IS) Extensions for Maximally Redundant Trees(MRT)", draft-
          li-isis-mrt-01 (work in progress), July 2014.

[I-D.wijnands-mpls-mldp-node-protection]
          Wijnands, I., Rosen, E., Raza, K., Tantsura, J., Atlas,
          A., and Q. Zhao, "mLDP Node Protection", draft-wijnands-
          mpls-mldp-node-protection-04 (work in progress), June
          2013.

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

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P.
          Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC
          4915, June 2007.

[RFC5715] Shand, M. and S. Bryant, "A Framework for Loop-Free
          Convergence", RFC 5715, January 2010.

Authors' Addresses

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

   Email: akatlas@juniper.net

   Kishore Tiruveedhula
   Juniper Networks
   10 Technology Park Drive
   Westford, MA 01886
   USA

   Email: kishoret@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

IJsbrand Wijnands
Cisco Systems, Inc.

Email: ice@cisco.com