Network Working Group
INTERNET-DRAFT
Intended Status: Standard Track
Expires: January 9, 2017

Anil Kumar S N
Gaurav Agrawal
Vinod Kumar S
Huawei Technologies India

July 8, 2016

Maximally Redundant Trees in Segment Routing

draft-agv-rtgwg-spring-segment-routing-mrt-02

Abstract

This document presents a Fast Reroute (FRR) approach aimed at providing link and node protection of node and adjacency segments within the Segment Routing (SR) framework based on Maximally Redundant Trees (MRT) FRR algorithm [RFC 7811].

Fast-Reroute with Maximally Redundant Trees (MRT-FRR) for Segment routing network is to provide link-protection and node- protection with 100% coverage in Segment routing network topology that is still connected after the failure. MRT is computational efficient.

Status of this Memo

Copyright Notice

Table of Contents

1 Introduction

Segment routing MRT FRR is one among the local repair mechanisms for Segment routing network. Another well known local repair mechanism for SR is Topology Independent Fast Reroute which is also capable of restoring end-to-end connectivity in case of a failure of a link or a node, with guaranteed coverage properties. MRT guarantees 100% recovery for single failures when the network is 2-connected. This guaranteed coverage does not depend on the link metrics, which an operator may be using to traffic-engineer the IP network. The link metrics and general network topology are largely decoupled from the guaranteed coverage.

The advantage of MRT over TI-LFA would be the computation complexities involved in MRT is much lesser then TI-LFA with additional cost of memory usage. MRT is best suited for access/aggregate ring network or low end devices which has low computing capacity but could afford to have enough memory to hold more FIB entries.

1.1 Standard Terminology

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 RFC 2119 [RFC2119].

2. Draft Specific Terminology

For ease of reading, some of the terminology defined in [RFC 7811] 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-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 MT-ID: 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).

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

Island Border Router (IBR): A router in the MRT Island that is connected to a router not in the MRT Island and both routers are in a common area or level.

Island Neighbor (IN): A router that is not in the MRT Island but is adjacent to an IBR and in the same area/level as the IBR..

3. MRT segment routing requirements

To extend MRT support to Segment routing following requirement need to be achieved :

1. SR MRT Capabilities must be advertised using IGP extension for SR MRT. Also SR MRT capabilities must be in sync with IGP specific MRT capabilities advertisement. If the peer has not advertised the SR MRT capability, then it indicates that LSR does not support MRT procedures.

2. As specified in MRT Architecture [RFC 7811], both Option 1A and Option 1B can be used for the implementation of SR MRT.

For Option 1A, two additional Prefix SID's/Label for RED and BLUE MT must be advertised in addition to default prefix SID/Label. The IGP extension carrying prefix SID for RED and BLUE MT must have corresponding MT-ID allocated by IANA for default MRT profile.

For Option 1B, Global Unique Context SID/Label for Red & Blue as topology identifier must be used.

4. MRT segment routing overview

Segment routing devices has to undergo no changes with respect to forwarding plane. Segment Routing (SR) allows a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called "segments". These segments are advertised by the link-state routing protocols (IS-IS and OSPF). Prefix segments represent an ECMP-aware shortest-path to a prefix (or a node), as per the state of the IGP topology. Adjacency segments represent a hop over a specific adjacency between two nodes in the IGP.

MRT FRR in segment routing network does not require any additional signaling (other than IGP extensions).

Basically MRT Fast Reroute requires that packets to be forwarded not only on the shortest-path tree, but also on two Maximally Redundant Trees (MRTs), referred to as the MRT-Blue and the MRT-Red. A router that experiences a local failure must also have predetermined which alternate to use. The MRT algorithm is based on those presented in [MRTLinear] and expanded in [EnyediThesis]. Default MRT Profile path calculation uses Lowpoint algorithm to calculate Maximally Redundant Trees.

Just as packets routed on a hop-by-hop basis require that each router compute a shortest-path tree that is consistent, it is necessary for each router to compute the MRT-Blue next hops and MRT-Red next hops in a consistent fashion.

A router’s Labeled Forwarding Information Base (L-FIB) will continue to contain primary next hops segment entries for the current shortest-path tree for forwarding traffic. In addition, a router’s L-FIB will contain primary next hops segments for the MRT-Blue for forwarding received traffic on the MRT-Blue and primary next hops segments for the MRT-Red for forwarding received traffic on the MRT-Red.

Within a link-state IGP domain, an SR-capable IGP node advertises segments for its attached prefixes and adjacencies. These segments are called IGP segments or IGP SIDs. They play a key role in Segment Routing and use-cases as they enable the expression of any topological path throughout the IGP domain. Such a topological path is either expressed as a single IGP segment or a list of multiple IGP segments. After running MRT lowpoint algorithm IGP will advertise two more additional labels as MRT-BLUE and MRT-RED for each such IGP segments.

Since segment Routing is directly applied to the MPLS architecture with no change on the forwarding plane and The ingress node of an SR domain encodes an ordered list of segments as a stack of labels. By default The ingress node encode only default path labels. The protecting router after detecting the node or link failure switches the top label with MRT label[MRT-RED or MRT-BLUE is selected based on algorithm] for the same destination. Till packet reaches the destination MRT colored label path is followed by the packet.

5. Requirements for SR MRT implementation

REQ1 : IGP Extension to carry the Segment Routing Node MRT Capability in addition to exiting IGP extension carrying IGP MRT Capability

REQ2 : IGP Extension to carry Red & Blue MRT SR Segments in addition to existing Default SR Segment

3 Security Considerations

None of the security consideration are identified

4 IANA Considerations

None of the IANA consideration are identified

5 References

5.1 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-05 (work in progress), July 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-05 (work in progress), January 2015.

[RFC5036]
Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <http://www.rfc-editor.org/info/rfc5036>.

[RFC5561]
Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL. Le Roux, "LDP Capabilities", RFC 5561, DOI 10.17487/RFC5561, July 2009, <http://www.rfc-editor.org/info/rfc5561>.

[RFC6420]
Cai, Y. and H. Ou, "PIM Multi-Topology ID (MT-ID) Join Attribute", RFC 6420, DOI 10.17487/RFC6420, November 2011, <http://www.rfc-editor.org/info/rfc6420>.

[RFC7307]
Zhao, Q., Raza, K., Zhou, C., Fang, L., Li, L., and D. King, "LDP Extensions for Multi-Topology", RFC 7307, DOI 10.17487/RFC7307, July 2014, <http://www.rfc-editor.org/info/rfc7307>.

[EnyediThesis]
Enyedi, G., "Novel Algorithms for IP Fast Reroute", Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics Ph.D. Thesis, February 2011, <https://repozitorium.omikk.bme.hu/bitstream/handle/10890/1040/ertekezes.pdf>.

[MRTLinear]
Enyedi, G., Retvari, G., and A. Csaszar, "On Finding Maximally Redundant Trees in Strictly Linear Time", IEEE Symposium on Computers and Communications (ISCC), 2009, <http://opti.tmit.bme.hu/~enyedi/ipfrr/distMaxRedTree.pdf>.

5.2. Informative References

[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-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-ietf-isis-mrt-00 (work in progress), February 2015.

[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-00 (work in progress), January 2015.

[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, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

Authors' Addresses

Anil Kumar S N
Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area,
Kundalahalli Village,
Whitefield,
Bangalore - 560066

EMail: anil.ietf@gmail.com

Gaurav Agrawal

Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area,
Kundalahalli Village,
Whitefield,
Bangalore - 560066

EMail: gaurav.agrawal@huawei.com

Vinod Kumar S
Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area,
Kundalahalli Village,
Whitefield,
Bangalore - 560066

EMail: vinods.kumar@huawei.com