                         Anycast-SID FRR in SR
                   draft-chen-spring-anycast-sid-frr-02

Abstract

   This document specifies the fast redundancy protection mechanism,
   aimed at providing protection of the links and domain boundary nodes
   for network that use segment routing.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   This document extends the use of Anycast-SID FRR to provide links and
   domain boundary nodes that use segment routing.

2. Conventions used in this document

   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. Problem statement

```
   +----------------------------+  +--------------+  +------------------+
   |     SID:20   SID:30        |  |              |  | SID:60           |
   |        A2-----A3--Link311-GW11------C1------GW21------A6           |
   |       / \    /             |  |\     / \     /|  |         \       |
   |      /   \  /        LinkL  |  \    /   \   / |  |          \      |
   |SID:10/     \/     Adj-SID 3|  |  \ /     \ /  |  |           \ SID:80 |
   |    A1      /\              SID:100   \     /  SID:200       A8     |
   |     \    /  \              |  | / \    / \ |  |            /       |
   |      \  /    \             |  |/   \  /   \|  |           /        |
   |       \/      \            |  |/     \ /    \|  |         /         |
   |        A4--------A5---------GW12------C2------GW22-------A7         |
   |       SID:40  SID:50       |  |              |  |  SID:70          |
   +----------------------------+  +--------------+  +------------------+
```
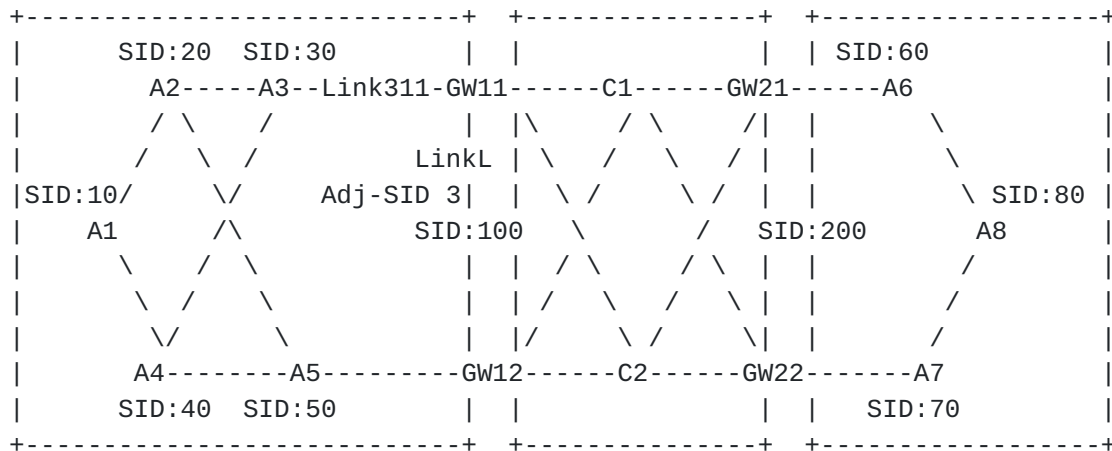
                            Figure 1

   The figure above describes a network example with two groups of the
   domain boundary nodes. The GW11 and GW12 are in the same anycast
   group. They are all configured with the same anycast prefix and the
   same prefix-sid 100, in addition, GW11 has node-sid 110 and GW12 has

node-sid 120. Suppose that the metric of link between two anycast
node is large while other links' metirc are small. From A1/A2/A3
perspective, GW11 is an active anycast node and GW12 is a standby,
from A4/A5 perspective, GW12 is active an anycast node , andGW11 is a
standby.

On the A3 node, it can select the primary or backup TI-LFA
FRR[I-D.ietf-rtgwg-segment-routing-ti-lfa] forwarding path (to
destination node GW12) which not through the active anycast node GW11
as the backup path to anycast-sid 100. In this example it will
select the direct next-hop A2 which is the primary path to
destination node GW12.

Supposed that an SID list is {100, 200, 80} which represents the SR-
TE path from A1 to A8.

Time1: the active anycast node GW11 DOWN.

Time2: Anycast FRR take effect in the event of node GW11 has failed,
the flow will be encapsulated with node-sid of GW12 and directed to
next-hop A2.

Time3: route convergence executed, depending on the convergence speed
of the nodes, A3 may regard GW12 as the new originator source of
anycast prefix, so that any flow that match the anycast prefix will
be forwarded to direct next-hop A2 to destination node GW12.
However, A2 may not converge so quickly, it will still regard GW11 as
the originator source of the anycast prefix, its forwarding entry of
anycast-prefix is still to destination GW11 and the next hop is A3.
There is a loop here. So A3 has responsibility to generate a micro-
loop avoided path {GW12, anycast-prefix} to anycast prefix, that is,
it must insert the new originator source to the unloop path.

4. Proposal

4.1. Domain boundary nodes protection

The solution consists of three parts.

o Configure the same anycast prefix and associated prefix-sid for
   each domain boundary node that forms redundant protection, then
   the anycast prefix and associated prefix-sid with Anycast-Group
   flag should be advertised to the neighbor node.

o Create the anycast-group forwarding entry (i.e. FRR entry) after
   the direct neighbor node of the domain boundary nodes receive the
   prefix-sid with Anycast-Group flag advertisement. The anycast-
   group forwarding entry includes the forwarding information which

points to each of the domain boundary node, then the forwarding
entry pointing to the main domain boundary (one of the direct
connected boundary nodes from the PLR) is set to the active state,
and others are set to the backup state. Only the direct neighbor
of the domain boundary nodes need to set up the anycast-group
forwarding entry. The anycast-group forwarding entry may also be
created on PLR by default, when it received an anycast-prefix
advertisement from two or more originator source.

o if the neighbor node detects the main domain boundary node
   failure, the neighbor node immediately activates the backup entry.
   Note that the backup entry contains the node-sid of the slave
   boundary node, and the packet will be forwarded based on the node-
   sid, not the anycast prefix-sid again.

4.2. example

In figure 1, considering that the GW11 DOWN,then

Time1:GW11 DOWN.

Time2: Primary or backup TI-LFA
[I-D.ietf-rtgwg-segment-routing-ti-lfa]provides protection in the
event of GW11 has failed on the A3 node.

Time3: When A3 detects GW11 failure, and the anycast-sid 100 is the
top Label in the label stack.The anycast-sid 100 is swapped with the
node-sid 120 (node-sid of the GW12) according to the anycast-group
forwarding entry. Pachets is forwarded to next-hop A2.

5. Security Considerations

TBD.

6. Acknowledgements

TBD.

7. Normative references

[I-D.ietf-rtgwg-segment-routing-ti-lfa]
          Litkowski, S., Bashandy, A., Filsfils, C., Decraene, B.,
          Francois, P., Voyer, D., Clad, F., and P. Camarillo,
          "Topology Independent Fast Reroute using Segment Routing",
          draft-ietf-rtgwg-segment-routing-ti-lfa-02 (work in
          progress), January 2020.

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

   [RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol
              Label Switching Architecture", RFC 3031,
              DOI 10.17487/RFC3031, January 2001,
              <https://www.rfc-editor.org/info/rfc3031>.

   [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L.,
              Decraene, B., Litkowski, S., and R. Shakir, "Segment
              Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,
              July 2018, <https://www.rfc-editor.org/info/rfc8402>.

   [RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C.,
              Bashandy, A., Gredler, H., and B. Decraene, "IS-IS
              Extensions for Segment Routing", RFC 8667,
              DOI 10.17487/RFC8667, December 2019,
              <https://www.rfc-editor.org/info/rfc8667>.

Authors' Addresses

   Ran Chen
   ZTE Corporation

   Email: chen.ran@zte.com.cn

   Shaofu Peng
   ZTE Corporation

   Email: peng.shaofu@zte.com.cn

   Jie Han
   ZTE Corporation

   Email: han.jie@zte.com.cn