IGP for Network High Availability

draft-chen-lsr-ctr-availability-00

Huaimo Chen (Futurewei)
Mehmet Toy (Verizon)
Aijun Wang (China Telecom)
Lei Liu (Fujitsu)
Xufeng Liu (Volta Networks)
Controller Cluster

Controller A (Primary) → Controller B (Secondary)

Tunnel from PE1 to PE4

Controller Cluster

Controller A (Primary) → Controller C (Third) → Controller N (n-th)

Tunnel from PE1 to PE4

Tunnel from PE1 to PE4

Issues by failures in cluster:
- Split into separated groups
- Out of Synchronization
- Multiple primary groups selected to control network

Extend IGP to resolve these
Overview of Mechanism

- Every controller has IGP to NE(s) as information channel

After failures in cluster:
- Live controller has information channel to NE(s)
- Information on controller is advertised via the channel
- Primary group is selected correctly to control network
Information on Controller

Normally, A (Primary) advertises the information about the controllers connected to it:

\[ C = 1, \ A's \ current \ Position = 1, \ A's \ OldPosition = 1, \ A's \ Priority, \ NoControllers = n, \ A's \ ID, \ B's \ ID, \ldots \]

After failures in cluster, for each separated group
- Intent primary, secondary controller, and so on are elected
- Intent primary controller advertises information about its group
- Every group has information about others. Primary group is selected.
- In case of tie, group with the highest old position controller (e.g., the old primary controller) wins in one policy
Extensions to OSPF

A new TLV, called OSPF Controllers TLV, is defined in RI LSA

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type (TBD1)</td>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flags</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Position</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OldPosition</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Priority</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reserved</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NoControllers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Controller 1 ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Controller n ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+-----------------+-----------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OSPF Controllers TLV</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Flag (8 bits):** One flag bit, C-bit, is defined. When set, it indicates that the position is the position of the current active primary controller.
- **Position (8 bits):** It indicates the current/intent position of the controller in the controller cluster or group. 1: primary (first) controller, 2: secondary controller, ...
- **OldPosition (8 bits):** It indicates the old position of the controller in the controller cluster before it is split.
- **Priority (8 bits):** It indicates the priority of the controller to be elected as a primary controller.
- **NoControllers (8 bits):** It indicates the number of controllers
- **Controller i ID (32 bits):** It represents the identifier (ID) of controller i at position i (i = 1, ..., n) in the cluster or group.
Extensions to IS-IS

Similar to OSPF, a new IS-IS Controllers TLV, is defined in LSP

```
<table>
<thead>
<tr>
<th>Type (TBD2)</th>
<th>Length</th>
<th>Flags</th>
<th>C</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OldPosition</td>
<td>Priority</td>
<td>NoControllers</td>
<td>Reserved</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controller 1 ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>:</td>
</tr>
<tr>
<td>Controller n ID</td>
</tr>
</tbody>
</table>
```

IS-IS Controllers TLV

The meaning of each of the fields such as Flags, Position is the same as the one of the corresponding field in the OSPF Controllers TLV.
Recovery Procedure

- Cluster of n controllers: A, B, …, N with position 1, 2, …, n respectively
- Failures split cluster into:
  Group 1: A, C;
  Group 2: B, N

- Normally, A originates LS containing a Controllers TLV:
  C=1, Position=1, OldPosition = 1, A’s Priority, NoControllers=n, A’s ID, B’s ID, …, and N’s ID.

- After failures, intent primary in each separated group originates LS containing a Controllers TLV
  A in group 1 originates LS containing:
  C=0, Position=1, OldPosition = 1, A’s Priority, NoControllers=2, A’s ID, C’s ID.

  B in group 2 originates LS containing:
  C=0, Position=1, OldPosition = 2, B’s Priority, NoControllers=2, B’s ID, N’s ID.

Group 1 and 2 have the same number of controllers, which is 2. But OldPosition in group 1 is higher than that in group 2. Group 1 is elected as the primary group
Primary controller A in the primary group (i.e., group 1) originates the LS containing
C=1, Position=1, OldPosition = 1, A’s Priority, NoControllers=2, A’s ID, C’s ID.
Next Step

Comments