Carry congestion status in BGP extended community

draft-li-idr-congestion-status-community-03.txt

Zhenqiang Li
China Mobile
Jie Dong
Huawei Technologies
Scenario to be addressed

• Knowing the congestion status of the exit link is useful to the BGP speakers.

• Router X in AS X, for example, can steer some traffic from link X-L to link X-a when it knows the exit links of backbone network are congested.

• draft-constrained-multiple-path from France Telecom also specified a similar requirement
Suggested Solution

- A new extended community is introduced in this document to deliver the link congestion status to other BGP speakers.

- Congestion status extended community can be used by the BGP speakers within one AS or in other ASes. i.e. it is a optional transitive attribute.

- In a network deployed SDN (Software Defined Network) controller, congestion status extended community can be used by the controller to steer the Internet access traffic among all the exit links from the perspective of the whole network.

- For the network with Route Reflectors (RRs), RRs are RECOMMENDED to enable add-path functionality, since by default RRs only advertise the best route for a specific prefix to their clients.
# Congestion Status Extended Community

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

```
<table>
<thead>
<tr>
<th></th>
<th>0x02</th>
<th>Sub-Type</th>
<th>Sender AS Number</th>
</tr>
</thead>
</table>
```

- It is a sub-type allocated from Transitive Four-Octet AS-Specific Extended Community Sub-Types defined in section 5.2.4 of [RFC7153].
- The "Type" field MUST be 0x02 to indicate this is a Transitive Four-Octet AS-Specific Extended Community.
- The "Sub-Type" field is used to indicate this is a Congestion Status Extended Community. Its value is to be assigned by IANA.
- The "Sender AS Number" field stores the AS number of the BGP speaker who generates this community.
- The "Bandwidth" field is 1 octet. Its value is the bandwidth of the exit link in unit of gbps.
- The "Utilization" field is 1 octet. Its value is the utilization of the exit link in unit of percent.
Application Considerations

• To avoid route oscillation
  • the exit router SHOULD set a threshold. The exit router generates BGP update messages with congestion status extended community only when the link utilization change reaches the threshold.
  • The method similar to BGP Route Flap Damping is RECOMMENDED for the implementations to further reduce the BGP update messages triggered by link utilization change.

• To avoid traffic oscillation
  • Route policy can be set at the exit router. Congestion status extended community is only conveyed for some specific routes or only for some specific BGP peers.
  • If the congestion status extended community is used by a SDN controller, the controller can steer the Internet access traffic among all the exit links from the perspective of the whole network.
IANA Requirements

• One sub-type is solicited to be assigned from Transitive Four-Octet AS-Specific Extended Community Sub-Types registry to indicate the Congestion Status Extended Community defined in this document.
  • 0x06 is suggested.
• WG adoption?
• Thanks