

# Carry congestion status in BGP extended community

draft-li-idr-congestion-status-community-05.txt

Zhenqiang Li  
China Mobile

Jie Dong  
Huawei Technologies

# Scenario to be addressed

- To aid traffic steering, network operators want to know the congestion status of the exit links.

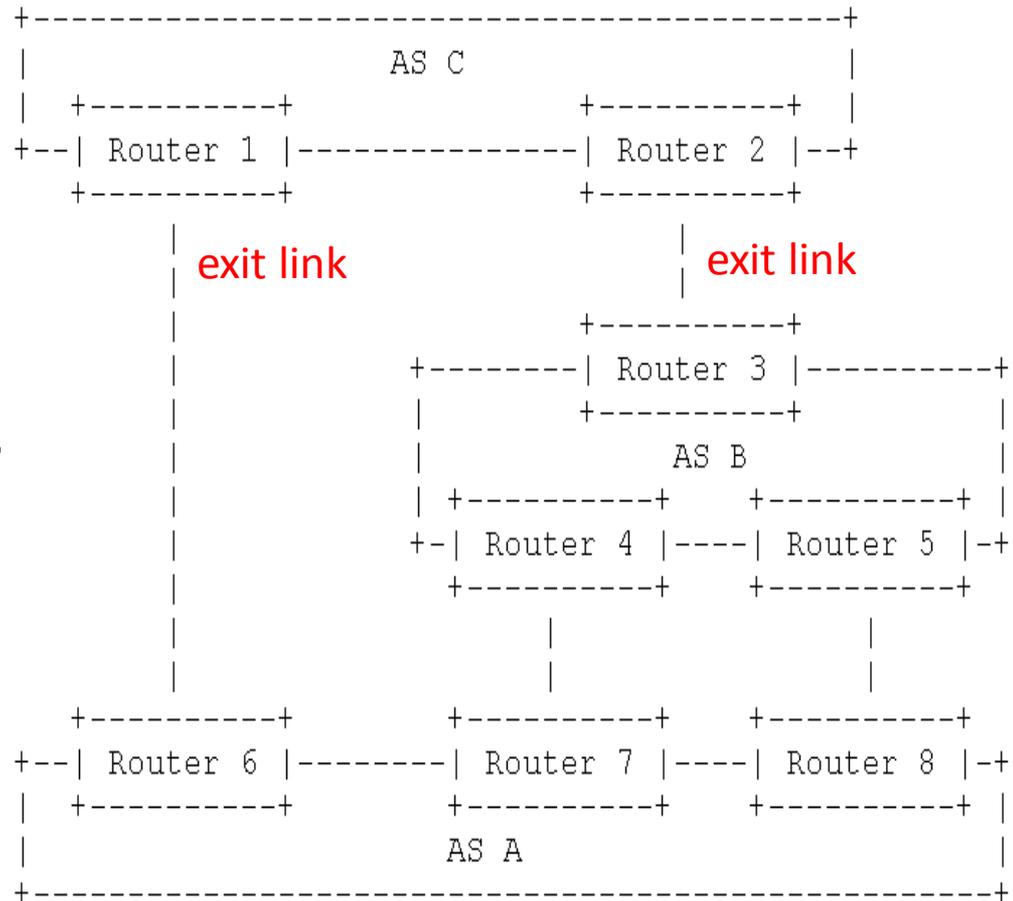
- Scenario 1: Within one AS.

- If the exit routers, Router 7 and 8, tell their iBGP peers in AS A the congestion status of the exit links, the peers in turn can steer some outgoing traffic toward the less loaded exit link.

- Scenario 2: Across multiple ASes.

- Due to cost or network performance, AS A prefers AS B to access AS C.
- If AS A knows the link between router 2 and 3 is congested, it can steer some traffic towards AS C from AS B to the directly connected link between router 1 and 6.

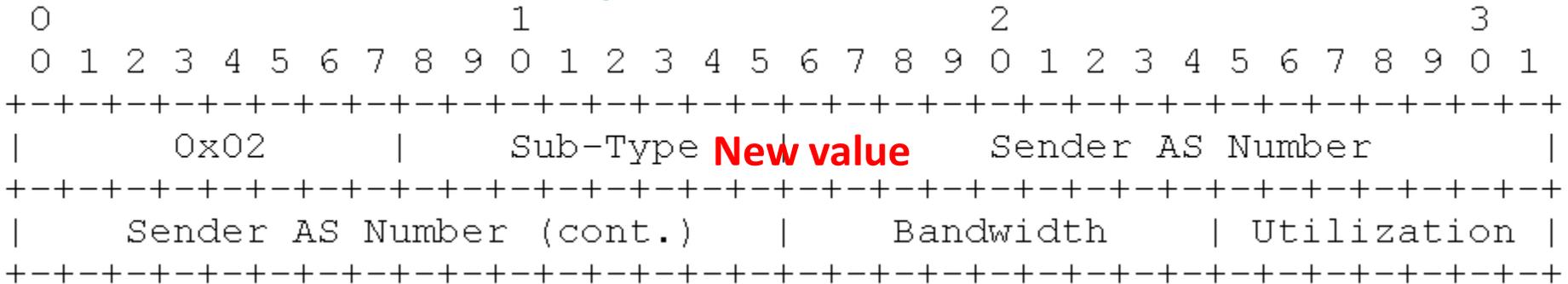
- Authors from France Telecom also specified this kind of requirement in **Constrained Multiple BGP Paths**.



# Suggested Solution

- A new extended community is introduced in this document to deliver the link congestion status to other BGP speakers.
- To satisfy the two application scenarios, we define the congestion status extended community as a optional transitive attribute.
- For SDN controller, besides BGP LS, it can also use the congestion status extended community to get the utilization information of the exit links.
- 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



- 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.

# Next steps

- Accepted by the working group?

**Thanks**