

Proposed Update to BGP Link-State SPF NLRI Selection Rules

draft-dong-lsvr-bgp-spf-selection-00

Jie Dong @Huawei

Jinqiang Chen @Huawei

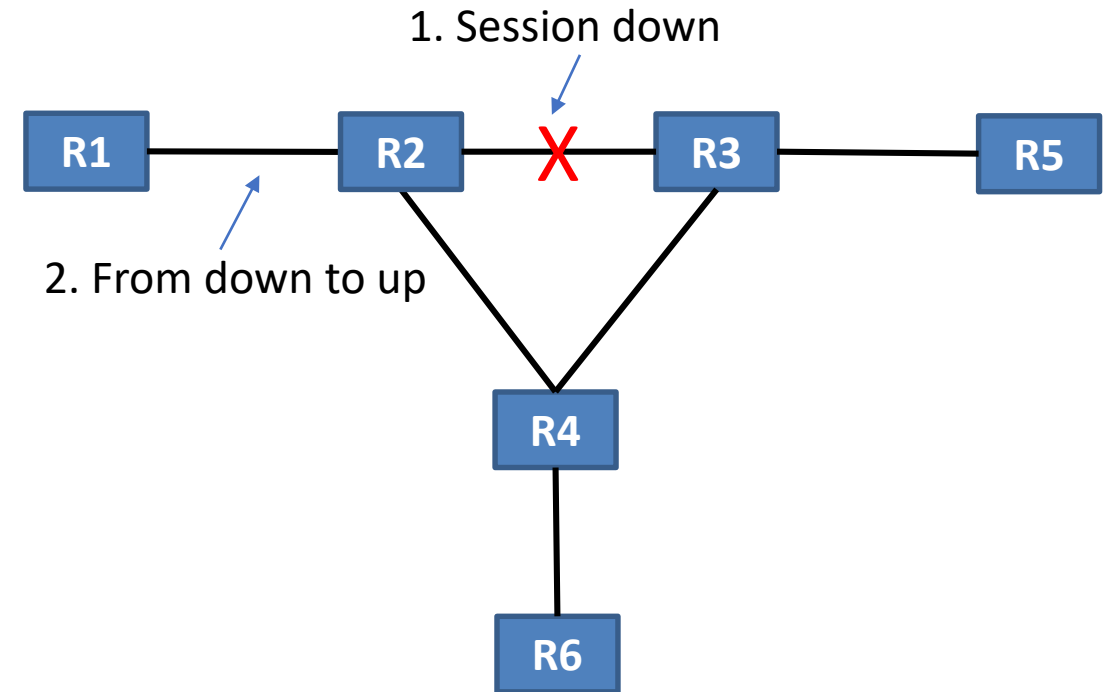
Sheng Fang @Huawei

Motivation

- BGP-LS SPF was designed for link-state information distribution and SPF path calculation in MSDC scenarios
 - Leverages the mechanisms of the BGP base protocol and BGP-LS extensions
- The NLRI selection rules for all BGP-LS SPF NLRI are defined as below:
 1. NLRI originated by directly connected BGP SPF peers are preferred
 2. The NLRI with the most recent Sequence Number TLV, i.e., highest sequence number is selected
 3. The NLRI received from the BGP SPF speaker with the numerically larger BGP Identifier is preferred
- In some cases, these rules may not be enough to provide deterministic selection result
- In some failure cases, these rules may cause the distribution of the latest link-state information be delayed
 - Which would result in delayed route convergence in the network
- This document describes the problem scenarios, and proposes some update to the selection rules of BGP-LS SPF NLRI

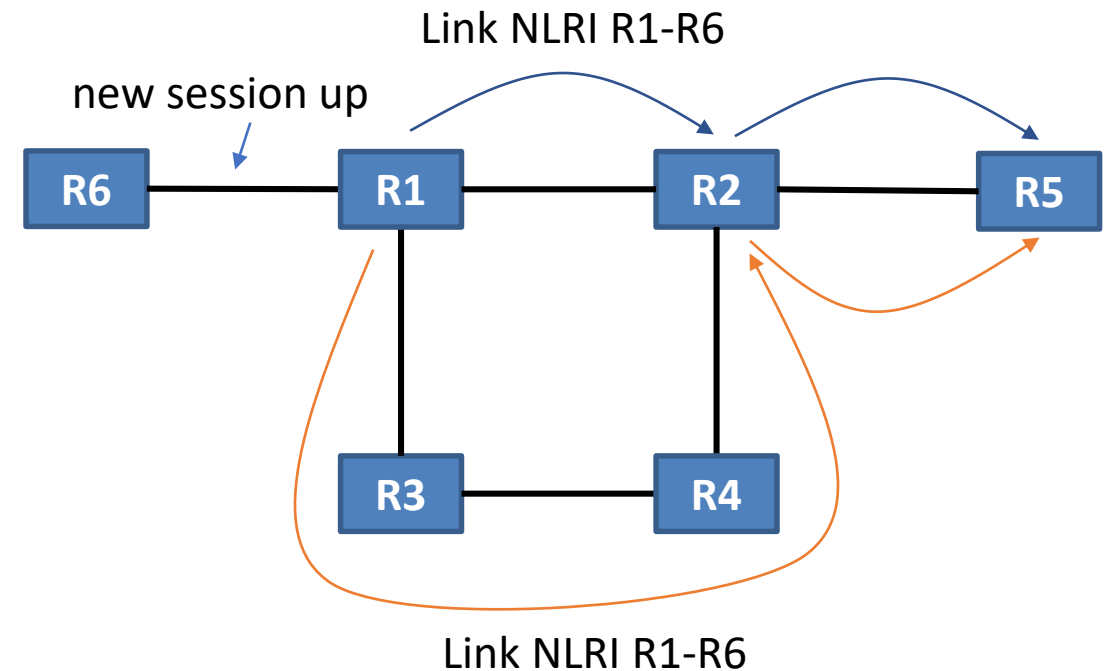
Problem Scenario 1: Delayed Convergence

- A failure of BGP session R2-R3 is detected by R3 (e.g. using BFD or other detection mechanisms)
- To avoid route flaps, R3 will hold all the NLRI received from R1 for NLRIImplicitWithdrawalDelay
- During this period, if the state of link R1-R2 changes from down to up, R2 would generate update for link NLRI R1-R2 with a greater sequence number, and advertise it to its BGP peers
- R3 receives R2's latest link NLRI R1-R2 from R4
- However, **R3 would prefer the link NLRI received from R2 directly** (according to the NLRI selection rule #1)
- Consequently, R3 would not use the latest link NLRI R1-R2 for SPF computation, nor it will advertise it further to its neighbors (R5 in this case)
- This would cause delayed convergence of the network



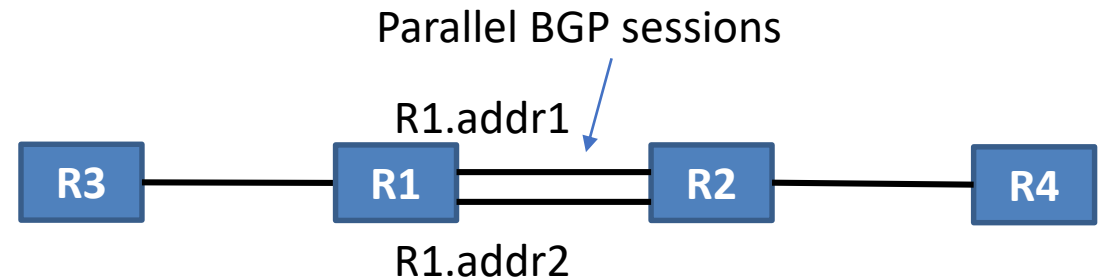
Problem Scenario 2: Redundant Advertisement

- A new BGP session is established between R1 and R6
- R1 advertises the link NLRI R1-R6 to its neighbors
- R2 firstly receives the link NLRI R1-R6 from R1 directly, and advertise it further to its neighbors (R4 and R5)
- R4 receives the link NLRI R1-R6 with the same sequence number from both R3 and R2, and **prefer the one from the peer with larger BGP ID (R3)**
- R4 advertises link NLRI R1-R6 to R2
- R2 prefers the link NLRI received from the peer with larger BGP-ID (R4)
- R2 advertise the link NLRI received from R4 to R5, **which is a redundant advertisement of the same link NLRI**



Problem Scenario 3: Indeterministic Selection Result

- There are two parallel links between R1 and R2, on each link a separate BGP session is established
- For the same NLRI with the same sequence number received from R1 via different sessions, the current NLRI selection rule cannot determine which one is the preferred route
- Thus R2 may select either one from the peer R1.addr1 or R1.addr2 as the best route, and advertises it further to R4
- To facilitate network operation and troubleshooting, it is preferable to have a deterministic result of NLRI selection once the network enters stable state



Proposed Updates to NLRI Selection Rules

- This document proposes to update the selection rules for all BGP-LS SPF NLRI as follows
 1. NLRI originated by directly connected BGP SPF peers SHOULD be preferred.
 2. The NLRI with the most recent Sequence Number TLV, i.e., highest sequence number SHOULD be selected.
 3. For NRIs received from EBGP peers, the NLRI with smaller number of AS numbers in the AS_PATH attribute SHOULD be preferred.
 4. For NRIs received from IBGP peers, the NLRI with smaller number of Cluster IDs in the CLUSTER_LIST attributes SHOULD be preferred.
 5. The NLRI received from the BGP SPF speaker with the numerically larger BGP Identifier SHOULD be preferred.
 6. NLRI received from the BGP SPF peer with the smaller peer address SHOULD be preferred.

Note: for problem scenario 1, further updates to the selection rule needs to be discussed

Next Steps

- Solicit opinion on the problem scenarios and the proposed NLRI selection rules
- Open discussion on possible solutions to solve or mitigate the problem scenario 1 are welcome
- Revise the draft accordingly
 - Or merge this into the BGP-LS SPF base document?

Thank You