Link State Vector Routing Working Group
Internet-Draft
Intended status: Informational
Expires: 4 September 2025

J. Dong
J. Chen
S. Fang
Huawei Technologies

3 March 2025

Proposed Update to BGP Link-State SPF NLRI Selection Rules
draft-dong-lsvr-bgp-spf-selection-02

Abstract

For network scenarios such as Massively Scaled Data Centers (MSDCs), BGP is extended for Link-State (LS) distribution and the Shortest Path First (SPF) algorithm based calculation. BGP-LS-SPF leverages the mechanisms of both BGP protocol and BGP-LS protocol extensions, with new selection rules defined for BGP-LS-SPF NLRI. This document proposes some updates to the BGP-LS-SPF NLRI selection rules, so as to improve the route updates and convergence, while consistent SPF computation result can still be achieved. This document updates the NLRI selection rules in I-D.ietf-lsvr-bgp-spf.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 4 September 2025.

Copyright Notice

Table of Contents

1. Introduction

For network scenarios such as Massively Scaled Data Centers (MSDCs), BGP is extended for Link-State (LS) distribution and the Shortest Path First (SPF) algorithm based calculation. BGP-LS-SPF leverages the mechanisms of both BGP protocol and BGP-LS protocol extensions, with new selection rules for BGP-LS-SPF NLRI defined in [I-D.ietf-lsvr-bgp-spf]. For all BGP-LS-SPF NLRIs, the NLRI selection rules 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.

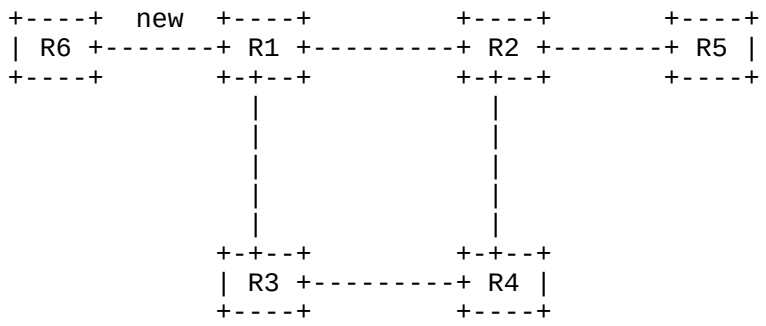
In some network scenarios, these rules may not be enough to provide optimized route convergence. This document proposes some updates to the BGP-LS-SPF NLRI selection rules, so as to improve the route convergence while consistent SPF computation result can still be achieved.

This document firstly describes the network scenarios in which the existing NLRI selection rules are considered not enough. Then it provides suggested updates to the BGP-LS-SPF NLRI selection rules.

2. Network Scenarios Which Triggered This Update

2.1. Unnecessary Redundant Advertisement

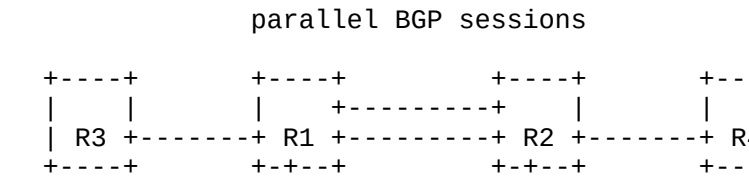
According to the rules in [I-D.ietf-lsvr-bgp-spf], for the BGP-LS-SPF NLRIs with the same sequence number, the NLRI received from the numerically larger BGP ID is preferred. While in some cases, this may cause unnecessary redundant advertisement of the same NLRI.



As shown in the example in Figure 2, a new BGP session is established between R1 and R6, and R1 advertise the link NLRI of R1-R6 to its neighboring nodes (R2 and R3). 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 of R1-R6 with the same sequence number from both R3 and R2, and according to the NLRI selection rules, R4 would prefer the NLRI received from R3 according to the rule of numerically larger BGP ID, then R4 advertises this link NLRI of R1-R6 to R2. R2 would also prefer the NLRI received from R4 according to the rule of numerically larger BGP ID, and further advertises this link NLRI to R5, which is a redundant advertisement of its previous advertisement of the same link NLRI.

2.2. Parallal BGP-LS-SPF Peers

In some scenarios, BGP single-hop peering model is used between directly connected BGP nodes. When two or more parallel links exists between the BGP nodes, multiple BGP sessions are established between the peering nodes, and each session will be used for the distribution of BGP-LS-SPF NLRIs.



As shown in the example of Figure 3, there are two parallel links between R1 and R2, and a separate BGP session is established on each link. Based on the existing BGP-LS-SPF NLRI selection rules, from R2's perspective, for the same NLRI with the same sequence number, either the route received from peer R1.1, or the route received from peer R1.2 may be selected as the best. No matter what tie-breaking rule is used, depending on the order of the routes received from R1, R2 may need to advertise duplicated NLRIs to R4.

3. Update to BGP-LS-SPF Selection Rules

As the BGP-LS-SPF NLRIs are used to distribute the link-state information, which are then used for the SPF computation, BGP attributes which are used in BGP best path selection (e.g. AS_PATH) for BGP address families other than BGP-LS are not considered in the computation of BGP-LS-SPF. The consistency of BGP-LS-SPF computation result only relies on the sequence number associated with the BGP-LS-SPF NLRIs. For network scenarios where optimized route convergence is more desirable, route updates due to the changes in BGP attributes which are not considered in the SPF computation, although may help to achieve deterministic NLRI selection, is considered not quite necessary.

Thus 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. NLRI received from a BGP-LS-SPF peer with the same sequence number as the one of the current selected NLRI SHOULD not be selected.

The new rule 3 can help to solve the duplicated advertisement problem as described in section 2.

4. IANA Considerations

This document makes no request of IANA.

5. Security Considerations

The mechanism described in this document provide updates to the NLRI selection rules for BGP-LS-SPF. It does not introduce any additional security considerations than those described in [RFC4271] and [RFC4272].

6. Acknowledgements

The authors would like to thank Haibo Wang, Jun Ge and Li Zhang for the valuable discussion and suggestions.

7. References

7.1. Normative References

[I-D.ietf-lsvr-bgp-spf] Patel, K., Lindem, A., Zandi, S., and W. Henderickx, "BGP Link-State Shortest Path First (SPF) Routing", Work in Progress, Internet-Draft, draft-ietf-lsvr-bgp-spf-51, 23 January 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-lsvr-bgp-spf-51>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

7.2. Informative References

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <https://www.rfc-editor.org/info/rfc4272>.

Authors' Addresses

Jie Dong
Huawei Technologies
China
Email: jie.dong@huawei.com

Jinqiang Chen
Huawei Technologies
China
Email: chenjinqiang@huawei.com

Sheng Fang
Huawei Technologies
China
Email: fangsheng@huawei.com