Link State Routing
Internet-Draft
Intended status: Standards Track
Expires: October 30, 2022

K. Talaulikar
Arrcus Inc
P. Psenak
Cisco Systems, Inc.
H. Johnston
AT&T Labs

April 28, 2022

OSPF Reverse Metric
draft-ietf-lsr-ospf-reverse-metric-05

Abstract

This document specifies the extensions to OSPF that enable a router to use link-local signaling to signal the metric that receiving neighbor(s) should use for a link to the signaling router. The signaling of this reverse metric, to be used on the link to the signaling router, allows a router to influence the amount of traffic flowing towards itself and in certain use cases enables routers to maintain symmetric metric on both sides of a link between them.

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 October 30, 2022.

Table of Contents

1. Introduction

Routers running the Open Shortest Path First (OSPFv2) [RFC2328] and OSPFv3 [RFC5340] routing protocols originate a Router-LSA (Link State Advertisement) that describes all its links to its neighbors and includes a metric that indicates its "cost" of reaching the neighbor over that link. Consider two routers R1 and R2 that are connected via a link. The metric for this link in direction R1->R2 is configured on R1 and in the direction R2->R1 is configured on R2. Thus the configuration on R1 influences the traffic that it forwards towards R2 but does not influence the traffic that it may receive from R2 on that same link.

This document describes certain use cases where a router is required to signal what we call the "reverse metric" (RM) to its neighbor to adjust the routing metric in the inbound direction. When R1 signals its reverse metric on its link to R2, then R2 advertises this value as its metric to R1 in its Router-LSA instead of its locally configured value. Once this information is part of the topology, then all other routers do their computation using this value which

results in the desired change in the traffic distribution that R1 wanted to achieve towards itself over the link from R2.

This document describes extensions to OSPF Link-Local Signaling (LLS) [RFC5613] to signal OSPF reverse metrics. Section 4 specifies the LLS Reverse Metric TLV and Section 5 specifies the LLS Reverse TE Metric TLV. The related procedures are specified in Section 6.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Use Cases

This section describes certain use cases that OSPF reverse metric helps address. The usage of the OSPF reverse metric need not be limited to these cases and is intended to be a generic mechanism.

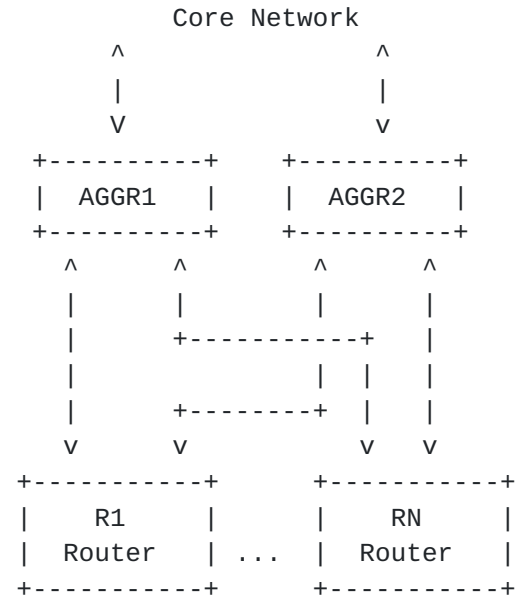


Figure 1: Reference Dual Hub and Spoke Topology

Consider a deployment scenario where, as shown in Figure 1, a bunch of routers R1 through RN, are dual-home connected to AGGR1 and AGGR2 that are aggregating their traffic towards a core network.

2.1. Link Maintenance

Before network maintenance events are performed on individual links, operators substantially increase (to maximum value) the OSPF metric simultaneously on both routers attached to the same link. In doing so, the routers generate new Router LSAs that are flooded throughout the network and cause all routers to gradually shift traffic onto alternate paths with very little or no disruption to in-flight communications by applications or end-users. When performed successfully, this allows the operator to confidently perform disruptive augmentation, fault diagnosis, or repairs on a link without disturbing ongoing communications in the network.

In deployments such as a hub and spoke topology as shown in Figure 1, it is quite common to have routers with several hundred interfaces and individual interfaces that move anywhere from several hundred gigabits/second to terabits/second of traffic. The challenge in such conditions is that the operator must accurately identify the same point-to-point link on two separate devices to increase (and afterward decrease) the OSPF metric appropriately and to do so in a coordinated manner. When considering maintenance for PE-CE links when a large number of CE routers connect to a PE router, an additional challenge related to coordinating access to the CE routers may arise when the CEs are not managed by the provider.

The OSPF reverse metric mechanism helps address these challenges. The operator can set the link on one of the routers (generally the hub like AGGR1 or a PE) in a "maintenance mode". This causes the router to advertise the maximum metric for that link and also to signal its neighbor on the same link to advertise maximum metric via the reverse metric signaling mechanism. Once the link maintenance is completed and the "maintenance mode" is turned off, the router returns to using its provisioned metric for the link and also stops the signaling of reverse metric on that link resulting in its neighbor to also revert to its provisioned metric for that link.

2.2. Adaptive Metric Signaling

In Figure 1 above, consider that at some point T, AGGR1 loses some of its capacity towards the core that may result in a congestion issue towards the core and it needs to reduce the traffic towards the core by redirecting some of the load to transit AGGR2 which is not experiencing a similar issue. Altering its link metric towards the R1-RN routers would influence the traffic from the core towards R1-RN but not the other way around as desired.

In such a scenario, the AGGR1 router could signal an incremental OSPF reverse metric to some or all of the R1-RN routers. When the R1-RN

routers add this signaled reverse metric offset to the provisioned metric on their links towards AGGR1, then the path via AGGR2 becomes a better path causing traffic towards the core to be diverted away from AGGR1. Note that the reverse metric mechanism allows such adaptive metric changes to be applied on the AGGR1 as opposed to being provisioned on a possibly large number of R1-RN routers.

The reverse metric mechanism may also be similarly applied between spine and leaf nodes in a CLOS topology deployment.

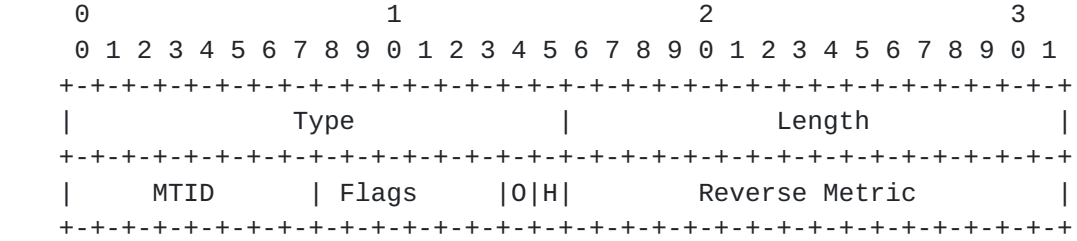
3. Solution

To address the use cases described earlier and to allow an OSPF router to indicate its reverse metric for a specific link to its neighbor(s), this document proposes to extend OSPF link-local signaling to signal the Reverse Metric TLV in OSPF Hello packets. This ensures that the RM signaling is scoped ONLY to each specific link individually. The router continues to include the Reverse Metric TLV in its Hello packets on the link as long as it needs its neighbor to use that metric value towards itself. Further details of the procedures involved are specified in Section 6.

The reverse metric mechanism specified in this document applies only for point-to-point, point-to-multipoint, and hybrid broadcast point-to-multipoint ([RFC6845]) links. It is not applicable for broadcast or non-broadcast-multi-access (NBMA) links since the same objective is achieved there using the OSPF Two-Part Metric mechanism [RFC8042] for OSPFv2. The OSPFv3 solution for broadcast or NBMA links is outside the scope of this document.

4. LLS Reverse Metric TLV

The Reverse Metric TLV is a new LLS TLV. It has following format:



where:

Figure 2: Reverse Metric TLV

Type: 19

Length: 4 octet

MTID : the multi-topology identifier of the link ([RFC4915])

Flags: 1 octet, the following flags are defined currently and the rest MUST be set to 0 on transmission and ignored on reception.

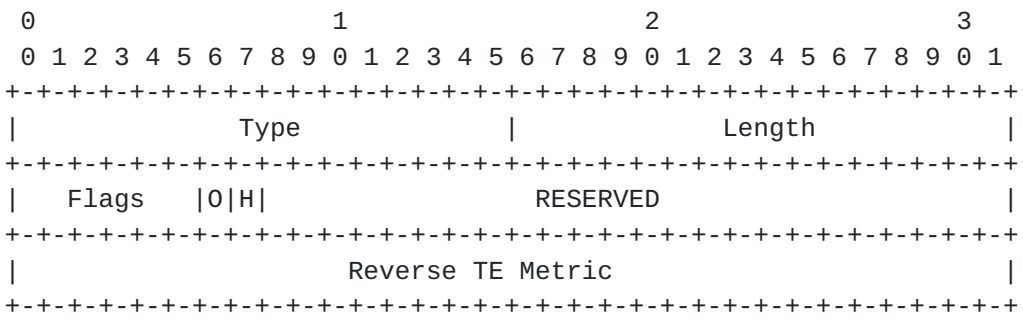
* H (0x1) : Indicates that the neighbor should use the value only if it is higher than its provisioned metric value for the link.

* O (0x2) : Indicates that the reverse metric value provided is an offset that is to be added to the provisioned metric.

Reverse Metric: 2 octets, the value or offset of reverse metric to replace or be added to the provisioned link metric.

5. LLS Reverse TE Metric TLV

The Reverse TE Metric TLV is a new LLS TLV. It has the following format:



where:

Figure 3: Reverse TE Metric TLV

Type: 20

Length: 4 octet

Flags: 1 octet, the following flags are defined currently and the rest MUST be set to 0 on transmission and ignored on reception.

* H (0x1) : Indicates that the neighbor should use the value only if it is higher than its provisioned TE metric value for the link.

* O (0x2) : Indicates that the reverse TE metric value provided is an offset that is to be added to the provisioned TE metric.

RESERVED: 24-bit field. SHOULD be set to 0 on transmission and MUST be ignored on receipt.

Reverse TE Metric: 4 octets, the value or offset of reverse traffic engineering metric to replace or to be added to the provisioned TE metric of the link.

6. Procedures

When a router needs to signal an RM value that its neighbor(s) should use for a link towards the router, it includes the Reverse Metric TLV in the LLS block of its hello messages sent on that link and continues to include this TLV for as long as it needs its neighbor to use this value. The mechanisms used to determine the value to be used for the RM is specific to the implementation and use case and is outside the scope of this document. For example, the RM value may be derived based on the router's link bandwidth with respect to a reference bandwidth.

A router receiving a hello packet from its neighbor that contains the Reverse Metric TLV on a link SHOULD use the RM value to derive the metric for the link to the advertising router in its Router-LSA. When the O flag is set, the metric value to be advertised is derived by adding the value in the TLV to the provisioned metric for the link. When the O flag is clear, the metric value to be advertised is derived directly from the value in the TLV. When the H flag is set and the O flag is clear, the metric value to be advertised is derived directly from the value in the TLV only when the RM value signaled is higher than the provisioned metric for the link.

A router stops including the Reverse Metric TLV in its hello messages when it needs its neighbors to go back to using their own provisioned metric values. When this happens, a router that had modified its metric in response to receiving a Reverse Metric TLV from its neighbor should revert to using its provisioned metric value.

In certain scenarios, two or more routers may start the RM signaling on the same link. This could create collision scenarios. The following rules MUST be adopted by routers to ensure that there is no instability in the network due to churn in their metric due to signaling of RM:

- The RM value that is signaled by a router to its neighbor MUST NOT be derived from the reverse metric being signaled by any of its neighbors on any of its links.

- The RM value that is signaled by a router MUST NOT be derived from its metric which has been modified on account of an RM signaled from any of its neighbors on any of its links. RM signaling from other routers can affect the router's metric advertised in its Router-LSA. When deriving the RM values that a router signals to its neighbors, it should use its provisioned local metric values not influenced by any RM signaling.

Based on these rules, a router MUST never start, stop, or change its RM metric signaling based on the RM metric signaling initiated by some other router. Based on the local configuration policy, each router would end up accepting the RM value signaled by its neighbor and there would be no churn of metrics on the link or the network on account of RM signaling.

In certain use cases when symmetrical metrics are desired (e.g., when metrics are derived based on link bandwidth), the RM signaling can be enabled on routers on either end of a link. In other use cases (as described in Section 2.1), RM signaling may need to be enabled only on the router at one end of a link.

When using multi-topology routing with OSPF [RFC4915], a router MAY include multiple instances of the Reverse Metric TLV in the LLS block of its hello message - one for each of the topologies for which it desires to signal the reserve metric.

In certain scenarios, the OSPF router may also require the modification of the TE metric being advertised by its neighbor router towards itself in the inbound direction. The Reverse TE Metric TLV, using similar procedures as described above, MAY be used to signal the reverse TE metric for router links. The neighbor SHOULD use the reverse TE metric value to derive the TE metric advertised in the TE Metric sub-TLV of the Link TLV in its TE Opaque LSA [RFC3630].

7. Operational Guidelines

The use of reverse metric signaling does not alter the OSPF metric parameters stored in a router's persistent provisioning database.

If routers that receive a reverse metric advertisement send a syslog message, this will assist in rapidly identifying the node in the network that is advertising an OSPF metric or TE metric different from that which is configured locally on the device.

When the link TE metric is raised to the maximum value, either due to the reverse metric mechanism or by explicit user configuration, this SHOULD immediately trigger the CSPF (Constrained Shortest Path First) recalculation to move the TE traffic away from that link.

Implementations SHOULD provide a configuration option to enable the signaling of reverse metric from a router to its neighbors and are RECOMMENDED to provide a configuration option to disable the acceptance of the RM from its neighbors.

If an implementation enables this mechanism by default, it is RECOMMENDED that it be disabled by the operators when not explicitly using it.

For the use case in Section 2.1, a router SHOULD limit the period of advertising reverse metric towards a neighbor only for the duration of a network maintenance window.

8. Backward Compatibility

The signaling specified in this document happens at a link-local level between routers on that link. A router that does not support this specification would ignore the Reverse Metric and Reverse TE Metric LLS TLVs and not update its metric(s) in the other LSAs. As a result, the behavior would be the same as prior to this specification. Therefore, there are no backward compatibility related issues or considerations that need to be taken care of when implementing this specification.

9. IANA Considerations

This specification updates Link Local Signalling TLV Identifiers registry.

IANA is requested to make permanent the following code points that have been assigned via early allocation

o 19 - Reverse Metric TLV

o 20 - Reverse TE Metric TLV

10. Security Considerations

The security considerations for "OSPF Link-Local Signaling" [RFC5613] also apply to the extension described in this document. The usage of the reverse metric TLVs is to alter the metrics used by routers on the link and influence the flow and routing of traffic over the network. Hence, modification of the Reverse Metric and Reverse TE

Metric TLVs may result in misrouting of traffic. If authentication is being used in the OSPF routing domain [RFC5709][RFC7474], then the Cryptographic Authentication TLV [RFC5613] SHOULD also be used to protect the contents of the LLS block.

Receiving a malformed LLS Reverse Metric or Reverse TE Metric TLVs MUST NOT result in a hard router or OSPF process failure. The reception of malformed LLS TLVs or sub-TLVs SHOULD be logged, but such logging MUST be rate-limited to prevent denial-of-service (DoS) attacks.

11. Contributors

Thanks to Jay Karthik for his contributions to the use cases and the review of the solution.

12. Acknowledgements

The authors would like to thank Les Ginsberg, Aijun Wang, Gyan Mishra, and Matthew Bocci for their review and feedback on this document. The authors would also like to thank Acee Lindem for this detailed shepherd's review and comments on this document.

The document leverages the concept of Reverse Metric for IS-IS, its related use cases, and applicability aspects from [RFC8500].

13. References

13.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <https://www.rfc-editor.org/info/rfc2328>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <https://www.rfc-editor.org/info/rfc3630>.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <https://www.rfc-editor.org/info/rfc5340>.

[RFC5613] Zinin, A., Roy, A., Nguyen, L., Friedman, B., and D. Yeung, "OSPF Link-Local Signaling", RFC 5613, DOI 10.17487/RFC5613, August 2009, <https://www.rfc-editor.org/info/rfc5613>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

13.2. Informative References

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5709] Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M., Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic Authentication", RFC 5709, DOI 10.17487/RFC5709, October 2009, <https://www.rfc-editor.org/info/rfc5709>.

[RFC6845] Sheth, N., Wang, L., and J. Zhang, "OSPF Hybrid Broadcast and Point-to-Multipoint Interface Type", RFC 6845, DOI 10.17487/RFC6845, January 2013, <https://www.rfc-editor.org/info/rfc6845>.

[RFC7474] Bhatia, M., Hartman, S., Zhang, D., and A. Lindem, Ed., "Security Extension for OSPFv2 When Using Manual Key Management", RFC 7474, DOI 10.17487/RFC7474, April 2015, <https://www.rfc-editor.org/info/rfc7474>.

[RFC8042] Zhang, Z., Wang, L., and A. Lindem, "OSPF Two-Part Metric", RFC 8042, DOI 10.17487/RFC8042, December 2016, <https://www.rfc-editor.org/info/rfc8042>.

[RFC8500] Shen, N., Amante, S., and M. Abrahamsson, "IS-IS Routing with Reverse Metric", RFC 8500, DOI 10.17487/RFC8500, February 2019, <https://www.rfc-editor.org/info/rfc8500>.

Authors' Addresses

Ketan Talaulikar
Arrcus Inc
India

Email: ketant.ietf@gmail.com

Peter Psenak
Cisco Systems, Inc.
Apollo Business Center
Mlynske nivy 43
Bratislava 821 09
Slovakia

Email: ppsenak@cisco.com

Hugh Johnston
AT&T Labs
USA

Email: hugh_johnston@labs.att.com