LSR Workgroup
Internet-Draft
Intended status: Standards Track
Expires: May 6, 2021

A. Lindem
Cisco Systems
Y. Qu
Futurewei
A. Roy
Arrcus, Inc.
S. Mirtorabi
Cisco Systems

November 2, 2020

OSPF Transport Instance Extensions
draft-acee-lsr-ospf-transport-instance-01

Abstract

OSPFv2 and OSPFv3 include a reliable flooding mechanism to disseminate routing topology and Traffic Engineering (TE) information within a routing domain. Given the effectiveness of these mechanisms, it is convenient to envision using the same mechanism for dissemination of other types of information within the domain. However, burdening OSPF with this additional information will impact intra-domain routing convergence and possibly jeopardize the stability of the OSPF routing domain. This document presents mechanism to relegate this ancillary information to a separate OSPF instance and minimize the impact.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction 3
2. Requirements Language 3
3. Possible Use Cases 3
   3.1. MEC Service Discovery 3
   3.2. Application Data Dissemination 4
   3.3. Intra-Area Topology for BGP-LS Distribution 4
4. OSPF Transport Instance 4
   4.1. OSPFv2 Transport Instance Packet Differentiation 5
   4.2. OSPFv3 Transport Instance Packet Differentiation 5
   4.3. Instance Relationship to Normal OSPF Instances 5
   4.4. Network Prioritization 5
   4.5. OSPF Transport Instance Omission of Routing Calculation . 6
   4.6. Non-routing Instance Separation 6
   4.7. Non-Routing Sparse Topologies 7
      4.7.1. Remote OSPF Neighbor 7
   4.8. Multiple Topologies 8
5. OSPF Transport Instance Information Encoding 8
   5.1. OSPFv2 Transport Instance Information Encoding 9
   5.2. OSPFv3 Transport Instance Information Encoding 9
6. Manageability Considerations 9
7. Security Considerations 9
8. IANA Considerations 9
9. Acknowledgement 9
10. References 9
   10.1. Normative References 9
   10.2. Informative References 10

Authors' Addresses 11

1. Introduction

OSPFv2 [RFC2328] and OSPFv3 [RFC5340] include a reliable flooding mechanism to disseminate routing topology and Traffic Engineering (TE) information within a routing domain. Given the effectiveness of these mechanisms, it is convenient to envision using the same mechanism for dissemination of other types of information within the domain. However, burdening OSPF with this additional information will impact intra-domain routing convergence and possibly jeopardize the stability of the OSPF routing domain. This document presents mechanism to relegate this ancillary information to a separate OSPF instance and minimize the impact.

This OSPF protocol extension provides functionality similar to "Advertising Generic Information in IS-IS" [RFC6823]. Additionally, OSPF is extended to support sparse non-routing overlay topologies Section 4.7.

2. 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.

3. Possible Use Cases

3.1. MEC Service Discovery

Multi-Access Edge Computing (MEC) plays an important role in 5G architecture. MEC optimizes the performance for ultra-low latency and high bandwidth services by providing networking and computing at the edge of the network [ETSI-WP28-MEC]. To achieve this goal, it's important to expose the network capabilities and services of a MEC device to 5G User Equipment UE, i.e. UEs.

The followings are an incomplete list of the kind of information that OSPF transport instance can help to disseminate:

o A network service is realized using one or more physical or virtualized hosts in MEC, and the locations of these service points might change. The auto-discovery of these service locations can be achieved using an OSPF transport instance.

o UEs might be mobile, and MEC should support service continuity and application mobility. This may require service state transferring

and synchronization. OSPF transport instance can be used to synchronize these states.

o Network resources are limited, such as computing power, storage. The availability of such resources is dynamic, and OSPF transport instance can be used to populate such information, so applications can pick the right location of such resources, hence improve user experience and resource utilization.

3.2. Application Data Dissemination

Typically a network consists of routers from different vendors with different capabilities, and some applications may want to know whether a router supports certain functionality or where to find a router supports a functionality, so it will be ideal if such kind of information is known to all routers or a group of routers in the network. For example, an ingress router needs to find an egress router that supports In-situ Flow Information Telemetry (IFIT) [I-D.wang-lsr-igp-extensions-ifit] and obtain IFIT parameters.

OSPF transport instance can be used to populate such router capabilities/functionalities without impacting the performance or convergence of the base OSPF protocol.

3.3. Intra-Area Topology for BGP-LS Distribution

In some cases, it is desirable to limit the number of BGP-LS [RFC5572] sessions with a controller to the a one or two routers in an OSPF domain. However, many times those router(s) do not have full

visibility to the complete topology of all the areas. To solve this problem without extended the BGP-LS domain, the OSPF LSAs for non-local area could be flooded over the OSPF transport instance topology using remote neighbors Section 4.7.1.

4. OSPF Transport Instance

In order to isolate the effects of flooding and processing of non-routing information, it will be relegated to a separate protocol instance. This instance should be given lower priority when contending for router resources including processing, backplane bandwidth, and line card bandwidth. How that is realized is an implementation issue and is outside the scope of this document.

Throughout the document, non-routing refers to routing information that is not used for IP or IPv6 routing calculations. The OSPF transport instance is ideally suited for dissemination of routing information for other protocols and layers.

4.1. OSPFv2 Transport Instance Packet Differentiation

OSPFv2 currently does not offer a mechanism to differentiate Transport instance packets from normal instance packets sent and received on the same interface. However, the [RFC6549] provides the necessary packet encoding to support multiple OSPF protocol instances.

4.2. OSPFv3 Transport Instance Packet Differentiation

Fortunately, OSPFv3 already supports separate instances within the packet encodings. The existing OSPFv3 packet header instance ID field will be used to differentiate packets received on the same link (refer to section 2.4 in [RFC5340]).

4.3. Instance Relationship to Normal OSPF Instances

In OSPF transport instance, we must guarantee that any information we've received is treated as valid if and only if the router sending it is reachable. We'll refer to this as the "condition of reachability" in this document.

The OSPF transport instance is not dependent on any other OSPF instance. It does, however, have much of the same as topology information must be advertised to satisfy the "condition of reachability".

Further optimizations and coupling between an OSPF transport instance and a normal OSPF instance are beyond the scope of this document. This is an area for future study.

4.4. Network Prioritization

While OSPFv2 (section 4.3 in [RFC2328]) are normally sent with IP precedence Internetwork Control, any packets sent by an OSPF transport instance will be sent with IP precedence Flash (B'011'). This is only appropriate given that this is a pretty flashy mechanism.

Similarly, OSPFv3 transport instance packets will be sent with the traffic class mapped to flash (B'011') as specified in ([RFC5340]).

By setting the IP/IPv6 precedence differently for OSPF transport instance packets, normal OSPF routing instances can be given priority during both packet transmission and reception. In fact, some router implementations map the IP precedence directly to their internal packet priority. However, internal router implementation decisions are beyond the scope of this document.

4.5. OSPF Transport Instance Omission of Routing Calculation

Since the whole point of the transport instance is to separate the routing and non-routing processing and fate sharing, a transport instance SHOULD NOT install any IP or IPv6 routes. OSPF routers SHOULD NOT advertise any transport instance LSAs containing IP or IPv6 prefixes and OSPF routers receiving LSAs advertising IP or IPv6 prefixes SHOULD ignore them. This implies that an OSPF transport instance Link State Database should not include any of the LSAs as shown in Table 1.

OSPFv2	summary-LSAs (type 3) AS-external-LSAs (type 5)

	NSSA-LSAs (type 7)
OSPFv3	inter-area-prefix-LSAs (type 2003) AS-external-LSAs (type 0x4005) NSSA-LSAs (type 0x2007) intra-area-prefix-LSAs (type 0x2009)
OSPFv3 Extended LSA	E-inter-area-prefix-LSAs (type 0xA023) E-as-external-LSAs (type 0xC025) E-Type-7-NSSA (type 0xA027) E-intra-area-prefix-LSA (type 0xA029)

LSAs not included in OSPF transport instance

If these LSAs are erroneously advertised, they will be flooded as per standard OSPF but MUST be ignored by OSPF routers supporting this specification.

4.6. Non-routing Instance Separation

It has been suggested that an implementation could obtain the same level of separation between IP routing information and non-routing information in a single instance with slight modifications to the OSPF protocol. The authors refute this contention for the following reasons:

- o Adding internal and external mechanisms to prioritize routing information over non-routing information are much more complex than simply relegating the non-routing information to a separate instance as proposed in this specification.

- o The instance boundary offers much better separation for allocation of finite resources such as buffers, memory, processor cores, sockets, and bandwidth.

- o The instance boundary decreases the level of fate sharing for failures. Each instance may be implemented as a separate process or task.

o With non-routing information, many times not every router in the OSPF routing domain requires knowledge of every piece of non-routing information. In these cases, groups of routers which need to share information can be segregated into sparse topologies greatly reducing the amount of non-routing information any single router needs to maintain.

4.7. Non-Routing Sparse Topologies

With non-routing information, many times not every router in the OSPF routing domain requires knowledge of every piece of non-routing information. In these cases, groups of routers which need to share information can be segregated into sparse topologies. This will greatly reduce the amount of information any single router needs to maintain with the core routers possibly not requiring any non-routing information at all.

With normal OSPF, every router in an OSPF area must have every piece of topological information and every intra-area IP or IPv6 prefix. With non-routing information, only the routers needing to share a set of information need be part of the corresponding sparse topology. For directly attached routers, one only needs to configure the desired topologies on the interfaces with routers requiring the non-routing information. When the routers making up the sparse topology are not part of a uniconnected graph, two alternatives exist. The first alternative is configure tunnels to form a fully connected graph including only those routers in the sparse topology. The second alternative is use remote neighbors as described in Section 4.7.1.

4.7.1. Remote OSPF Neighbor

With sparse topologies, OSPF routers sharing non-routing information may not be directly connected. OSPF adjacencies with remote neighbors are formed exactly as they are with regular OSPF neighbors. The main difference is that a remote OSPF neighbor's address is configured and IP routing is used to deliver OSPF protocol packets to the remote neighbor. Other salient feature of the remote neighbor include:

o All OSPF packets have the remote neighbor's configured IP address

as the IP destination address.

o The adjacency is represented in the router Router-LSA as a router (type-1) link with the link data set to the remote neighbor's configured IP address.

o Similar to NBMA networks, a poll-interval is configured to determine if the remote neighbor is reachable. This value is normally much higher than the hello interval with 40 seconds RECOMMENDED as the default.

4.8. Multiple Topologies

For some applications, the information need to be flooded only to a topology which is a subset of routers of the transport instance. This allows the application specific information only to be flooded to routers that support the application. A transport instance may support multiple topologies as defined in [RFC4915]. But as pointed out in Section 4.5, a transport instance or topology SHOULD NOT install any IP or IPv6 routes.

Each topology associated with the transport instance MUST be fully connected in order for the LSAs to be successfully flooded to all routers in the topology.

5. OSPF Transport Instance Information Encoding

The format of the TLVs within the body of an LSA containing non-routing information is the same as the format used by the Traffic Engineering Extensions to OSPF [RFC3630]. The LSA payload consists of one or more nested Type/Length/Value (TLV) triplets. The format of each TLV is:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Value...                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

TLV Format

However, each unique application using the mechanisms defined in this document will have it's own unique ID. Whether to encode this ID as

the top-level TLV or make it part of the OSPF LSA ID is open for debate.

The specific TLVs and sub-TLVs relating to a given application and the corresponding IANA considerations MUST for standard applications MUST be specified in the document corresponding to that application.

5.1. OSPFv2 Transport Instance Information Encoding

Application specific information will be flooded in opaque LSAs as specified in [RFC5250].

5.2. OSPFv3 Transport Instance Information Encoding

Application specific information will be flooded in separate LSAs with separate function codes. Refer to section A.4.2.1 of [RFC5340]. for information on the LS Type encoding in OSPFv3, and section 2 of [RFC8362] for OSPFv3 extended LSA types.

6. Manageability Considerations

7. Security Considerations

The security considerations for the Transport Instance will not be different for those for OSPFv2 [RFC2328] and OSPFv3 [RFC5340].

8. IANA Considerations

No IANA actions are required.

9. Acknowledgement

10. References

10.1. 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>.

[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>.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, DOI 10.17487/RFC5250, July 2008, <https://www.rfc-editor.org/info/rfc5250>.

[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>.

[RFC6549] Lindem, A., Roy, A., and S. Mirtorabi, "OSPFv2 Multi-Instance Extensions", RFC 6549, DOI 10.17487/RFC6549, March 2012, <https://www.rfc-editor.org/info/rfc6549>.

[RFC6823] Ginsberg, L., Previdi, S., and M. Shand, "Advertising Generic Information in IS-IS", RFC 6823, DOI 10.17487/RFC6823, December 2012, <https://www.rfc-editor.org/info/rfc6823>.

[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>.

[RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", RFC 8362, DOI 10.17487/RFC8362, April 2018, <https://www.rfc-editor.org/info/rfc8362>.

10.2. Informative References

[ETSI-WP28-MEC]

Sami Kekki, etc., "MEC in 5G Networks", 2018, <https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp28_mec_in_5G_FINAL.pdf>.

[I-D.wang-lsr-igp-extensions-ifit]
Wang, Y., Zhou, T., Qin, F., Chen, H., and R. Pang, "IGP Extensions for In-situ Flow Information Telemetry (IFIT) Capability Advertisement", draft-wang-lsr-igp-extensions-ifit-01 (work in progress), July 2020.

[RFC5572] Blanchet, M. and F. Parent, "IPv6 Tunnel Broker with the Tunnel Setup Protocol (TSP)", RFC 5572, DOI 10.17487/RFC5572, February 2010, <https://www.rfc-editor.org/info/rfc5572>.

Authors' Addresses

Acee Lindem
Cisco Systems
301 Midenhall Way
CARY, NC 27513
UNITED STATES

Email: acee@cisco.com

Yingzhen Qu
Futurewei
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: yingzhen.qu@futurewei.com

Abhay Roy
Arrcus, Inc.

Email: abhay@arrcus.com

Sina Mirtorabi

Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
USA

Email: smirtora@cisco.com