

Internet Draft  
<draft-chen-ospf-transition-to-ospfv3-01.txt>  
Category: Informational

I. Chen  
A. Lindem  
Ericsson  
R. Atkinson  
Consultant  
July 2, 2014

Expires in 6 months

OSPFv3 over IPv4 for IPv6 Transition  
<draft-chen-ospf-transition-to-ospfv3-01.txt>

#### Status of this Memo

Distribution of this memo is unlimited.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on date.

#### Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

## Abstract

This document defines a mechanism to use IPv4 to transport OSPFv3 packets, in order to facilitate transition from IPv4-only to IPv6 and dual-stack within a routing domain. Using OSPFv3 over IPv4 with the existing OSPFv3 Address Family extension can simplify transition from an OSPFv2 IPv4-only routing domain to an OSPFv3 dual-stack routing domain.

## Table of Contents

1. Introduction .....	3
2. Encapsulation in IPv4 .....	4
2.1. Source Address .....	6
2.2. Destination .....	6
2.3. Operation over Virtual Link .....	6
3. IPv4-only Use Case .....	7
4. Security Considerations .....	7
5. IANA Considerations .....	8
6. References .....	8

## 1. Introduction

To facilitate transition from IPv4 [RFC791] to IPv6 [RFC2460], dual-stack or IPv6 routing protocols should be gradually deployed. Dual-stack routing protocols, such as Border Gateway Protocol [RFC4271], have an advantage during the transition, because both IPv4 and IPv6 topologies can be transported using either IPv4 or IPv6. Some IPv4-specific and IPv6-specific routing protocols share enough similarities in their protocol packet formats and protocol signaling that it is trivial to deploy an initial IPv6 routing domain by carrying the routing protocol over IPv4 initially, thereby allowing IPv6 routing domains be deployed and tested before decommissioning IPv4 and moving to an IPv6-only network.

In the case of the Open Shortest Path First (OSPF) interior gateway routing protocol (IGP), OSPFv2 [RFC2328] is the IGP deployed over IPv4, while OSPFv3 [RFC5340] is the IGP deployed over IPv6. OSPFv3 further supports multiple address families [RFC5838], including both the IPv6 unicast address family and the IPv4 unicast address family. Consequently, it is possible to deploy OSPFv3 over IPv4 without any changes either to OSPFv3 or to IPv4. During the transition to IPv6, future OSPF extension can focus on OSPFv3 and OSPFv2 can move into maintenance mode.

This document specifies how to use IPv4 packets to transport OSPFv3 packets. The mechanism takes advantage of the fact that OSPFv2 and OSPFv3 share the same IP protocol number, 89. Additionally, the OSPF packet header for both OSPFv2 and OSPFv3 places the OSPF header version (i.e., the field that distinguishes an OSPFv2 packet from an OSPFv3 packet) in the same location.

This document does not attempt to connect an IPv4 topology and an IPv6 topology that are not congruent. In normal operation, it is expected that the IPv4 topology within the OSPF domain will be congruent with the IPv6 topology of that OSPF domain. In such cases, it is expected either that all OSPFv3 packets will be transported

over IPv4 or that all OSPFv3 packets will be transported over IPv6.

If the IPv4 topology and IPv6 topology are not identical, the most likely cause is that some parts of the network deployment have not yet been upgraded to support both IPv4 and IPv6. In situations where the IPv4 deployment is a proper superset of the IPv6 deployment, it is expected that OSPFv3 packets would be transported over IPv4, until the rest of the network deployment is upgraded to support IPv6 in addition to IPv4. In situations where the IPv6 deployment is a proper superset of the IPv4 deployment, it is expected that OSPFv3 would be transported over IPv6.

Throughout this document, OSPF is used when the text applies to both OSPFv2 and OSPFv3. OSPFv2 or OSPFv3 is used when the text is specific to one version of the OSPF protocol. Similarly, IP is used when the text describes either version of the Internet protocol. IPv4 or IPv6 is used when the text is specific to a single version of the protocol.

## 2. Encapsulation in IPv4

Unlike 6to4 encapsulation [RFC3056] that tunnels IPv6 traffic through an IPv4 network, an OSPFv3 packet can be directly encapsulated within an IPv4 packet as the payload, without the IPv6 packet header, as illustrated in Figure 1. For OSPFv3 transported over IPv4, the IPv4 packet has an IPv4 protocol type of 89, denoting that the payload is an OSPF packet. The payload of the IPv4 packet consists of an OSPFv3 packet, beginning with the OSPF packet header with the OSPF version number set to 3.

An OSPFv3 packet followed by an OSPF link-local signaling (LLS) extension data block [RFC5613] encapsulated in an IPv4 packet is illustrated in Figure 2.

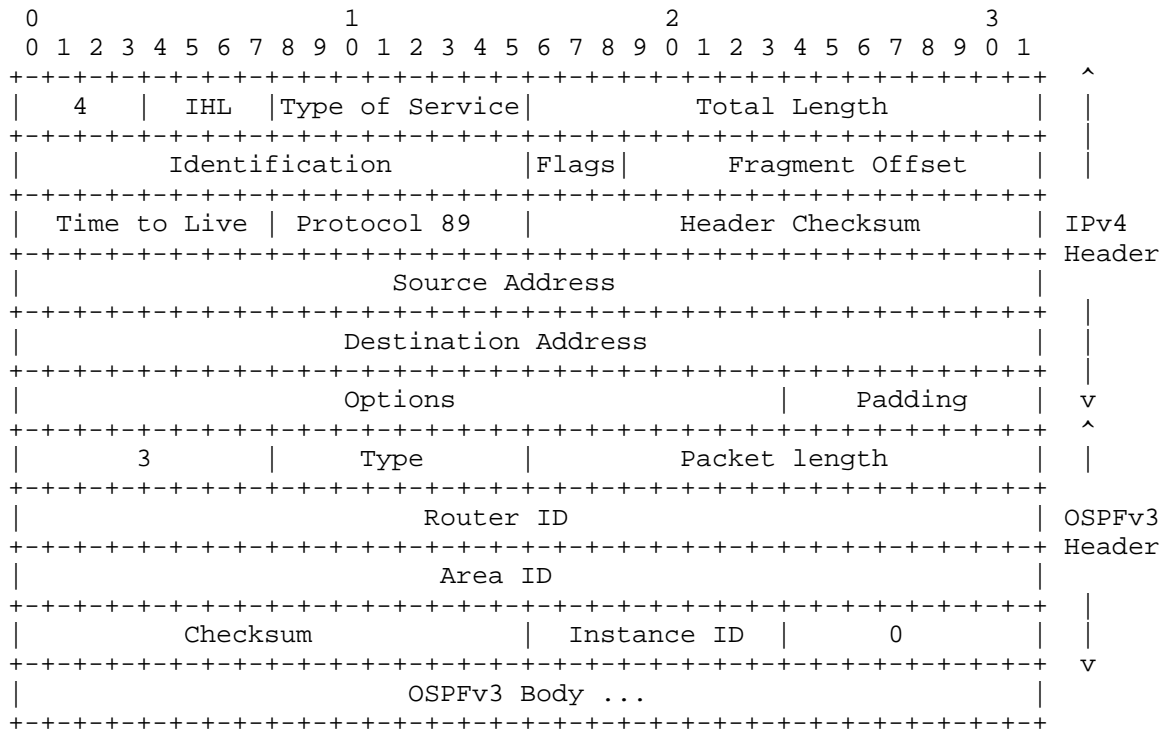


Figure 1: An IPv4 packet encapsulating an OSPFv3 packet.

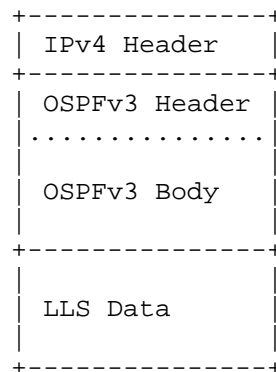


Figure 2: The IPv4 packet encapsulating an OSPFv3 packet with a trailing OSPF link-local signaling data block.

## 2.1. Source Address

For OSPFv3 over IPv4, the source address is the IPv4 interface address for the interface over which the packet is transmitted. All OSPFv3 routers on the link MUST share the same IPv4 subnet for IPv4 transport to function correctly.

## 2.2. Destination Address

As defined in OSPFv2, the IPv4 destination address of an OSPF protocol packet is either an IPv4 multicast address or the IPv4 unicast address of an OSPFv2 neighbor. Two well-known link-local multicast addresses are assigned to OSPFv2, the AllSPFRouters address (224.0.0.5) and the AllDRouters address (224.0.0.6). The multicast address used depends on the OSPF packet type, the OSPF interface type, and the OSPF router's role on multi-access networks.

Thus, for an OSPFv3 over IPv4 packet to be sent to AllSPFRouters, the destination address field in the IPv4 packet should be 224.0.0.5. For an OSPFv3 over IPv4 packet to be sent to AllDRouters, the destination address field in the IPv4 packet should be 224.0.0.6.

When an OSPF router sends a unicast OSPF packet over a connected interface, the destination of such an IP packet is the address assigned to the receiving interface. Thus, a unicast OSPFv3 packet transported in an IPv4 packet would specify the OSPFv3 neighbor's IPv4 address as the destination address.

## 2.3. Operation over Virtual Link

When an OSPF router sends an OSPF packet over a virtual link, the receiving router is a router which is not directly connected to the sending router. Thus, the destination IP address of the IP packet must be a reachable unicast IP address of the receiving router. Because IPv6 is the presumed Internet protocol and an IPv4 destination is not routable, the OSPFv3 address family extension [RFC5838] specifies that only IPv6 address family virtual links are supported.

As illustrated in Figure 1, this document specifies OSPFv3 transport over IPv4. As a result, an IPv4 packet in which the destination field is a unicast IPv4 address assigned to the virtual router is routable, and OSPFv3 virtual links in IPv4 unicast address families can be supported. Hence, the restriction in Section 2.8 of RFC 5838 [RFC5838] is removed. If IPv4 transport, as specified herein, is used for IPv6 address families, virtual

links cannot be supported. Hence, it is RECOMMENDED to use the IP transport matching the address family in OSPF routing domains requiring virtual links.

### 3. IPv4-only Use Case

OSPFv3 only requires IPv6 link-local addresses to establish a routing domain, and does not require IPv6 global-scope addresses to establish a routing domain. However, IPv6 over Ethernet [RFC2464] uses a different EtherType (0x86dd) from IPv4 (0x0800) and also from the Address Resolution Protocol (ARP) (0x0806) [RFC826] that is used with IPv4.

Some existing deployed link-layer equipment only supports IPv4 and ARP. Such equipment contains hardware filters keyed on the EtherType field of the Ethernet frame to filter which frames will be accepted into that link-layer equipment. Because IPv6 uses a different EtherType, IPv6 framing for OSPFv3 won't work with that equipment. In other cases, PPP might be used over a serial interface, but again only IPv4 over PPP might be supported over that interface. It is hoped that equipment with such limitations will be replaced eventually.

In some locations, especially locations with less communications infrastructure, satellite communications (SATCOM) is used to reduce deployment costs for data networking. SATCOM often has lower cost to deploy than running new copper or optical cables for long distances to connect remote areas. Also, in a wide range of locations including places with good communications infrastructure, Very Small Aperture Terminals (VSAT) often are used by banks and retailers to connect their stores to their main offices.

Some widely deployed VSAT equipment has either (A) Ethernet interfaces that only support Ethernet Address Resolution Protocol (ARP) and IPv4, or (B) serial interfaces that only support IPv4 and Point-to-Point Protocol (PPP) packets. Such deployments and equipment still can deploy and use OSPFv3 over IPv4 today, and then later migrate to OSPFv3 over IPv6 after equipment is upgraded or replaced. This can have lower operational costs than running OSPFv2 and then trying to make a flag-day switch to running OSPFv3. By running OSPFv3 over IPv4 now, the eventual transition to dual-stack, and then to IPv6-only can be optimized.

### 4. Security Considerations

As described in [RFC4552], OSPFv3 uses IPsec [RFC4301] for authentication and confidentiality. Consequently, an OSPFv3 packet transported within an IPv4 packet requires IPsec to provide

authentication and confidentiality. Further work such as [ipsecospf] would be required to support IPsec protection for OSPFv3 over IPv4 transport.

An optional OSPFv3 Authentication Trailer [RFC6506] also has been defined as an alternative to using IPsec. The calculation of the authentication data in the Authentication Trailer includes the source IPv6 address to protect an OSPFv3 router from Man-in-the-Middle attacks. For IPv4 encapsulation as described herein, the IPv4 source address should be placed in the first 4 octets of Apad followed by the hexadecimal value 0x878FE1F3 repeated (L-4)/4 times, where L is the length of hash measured in octet.

The processing of the optional Authentication Trailer is contained entirely within the OSPFv3 protocol. In other words, each OSPFv3 router instance is responsible for the authentication, without involvement from IPsec or any other IP layer function. Consequently, except for calculation of the value Apad, transporting OSPFv3 packets using IPv4 does not change the operation of the optional OSPFv3 Authentication Trailer.

## 5. IANA Considerations

No actions are required from IANA as result of the publication of this document.

## 6. References

### 6.1. Normative References

- [RFC791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [RFC2328] Moy, J., "OSPF Version 2", STD54, RFC 2328, April 1998.
- [RFC5838] Lindem, A., Ed., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC 5838, April 2010.

### 6.2. Informative References

- [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A



Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

- [RFC3056] Carpenter, B. and K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.
- [RFC5613] Zinin, A., Roy, A., Nguyen, L., Friedman, B., and D. Yeung, "OSPF Link-Local Signaling", RFC 5613, August 2009.
- [RFC826] Plummer, D., "Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, RFC 826, November 1982.
- [RFC2464] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC 2464, December 1998.
- [RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, June 2006.
- [RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.
- [RFC6506] Bhatia, M., Manral, V., and A. Lindem, "Supporting Authentication Trailer for OSPFv3", RFC 6506, February 2012.
- [ipseccospf] Gupta, M. and Melam, M, Work in progress, "draft-gupta-ospf-ospfv2-sec-01.txt", August 2009.

#### Authors' Addresses

I. Chen  
Ericsson  
Email: ing-wher.chen@ericsson.com

A. Lindem  
Ericsson  
Email: acee.lindem@ericsson.com

R. Atkinson  
Consultant



Open Shortest Path First IGP  
Internet-Draft  
Intended status: Standards Track  
Expires: December 28, 2014

S. Hegde  
H. Raghuveer  
H. Gredler  
Juniper Networks, Inc.  
R. Shakir  
British Telecom  
A. Smirnov  
Cisco Systems, Inc.  
Z. Li  
Huawei Technologies  
June 26, 2014

Advertising per-node administrative tags in OSPF  
draft-hegde-ospf-node-admin-tag-02

Abstract

This document describes an extension to OSPF protocol [RFC2328] to add an optional operational capability, that allows tagging and grouping of the nodes in an OSPF domain. This allows simplification, ease of management and control over route and path selection based on configured policies.

This document describes the protocol extensions to disseminate per-node admin-tags to the OSPFv2 and OSPFv3 protocol.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 <http://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 December 28, 2014.

#### Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

#### Table of Contents

1. Introduction . . . . .	2
2. Applicability . . . . .	2
3. Administrative Tag TLV . . . . .	3
4. OSPF per-node administrative tag TLV . . . . .	3
4.1. TLV format . . . . .	3
4.2. Elements of procedure . . . . .	4
5. Applications . . . . .	5
6. Security Considerations . . . . .	8
7. IANA Considerations . . . . .	9
8. Acknowledgments . . . . .	9
9. References . . . . .	9
9.1. Normative References . . . . .	9
9.2. Informative References . . . . .	9
Authors' Addresses . . . . .	10

#### 1. Introduction

This document provides mechanisms to advertise per-node administrative tags in the OSPF Router Information LSA [RFC4970]. In certain path-selection applications like for example in traffic-engineering or LFA backup selection there is a need to tag the nodes based on their roles in the network and have policies to prefer or prune a certain group of nodes.

#### 2. Applicability

For the purpose of advertising per-node administrative tags within OSPF a new TLV is proposed. Because path selection is a functional

set which applies both to TE and non-TE applications, this new TLV is carried in the Router Information LSA (RI LSA) [RFC4970]

### 3. Administrative Tag TLV

An administrative Tag is a 32-bit integer value that can be used to identify a group of nodes in the OSPF domain.

The new TLV defined will be carried within an RI LSA for OSPFV2 and OSPFV3. Router information LSA [RFC4970] can have link, area or AS level flooding scope. Choosing the flooding scope to flood the group tags are defined by the policies and is a local matter.

The TLV specifies one or more administrative tag values. An OSPF node advertises the set of groups it is part of in the OSPF domain. (for example, all PE-nodes are configured with certain tag value, all P-nodes are configured with a different tag value in a domain). The total number of admin tags that a given router can advertise at one time is restricted to 64. If more tags are needed in future, multi-instantiating of the RI LSA [RFC4970] may be required.

### 4. OSPF per-node administrative tag TLV

#### 4.1. TLV format

The format of the TLVs within the body of an RI LSA 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:

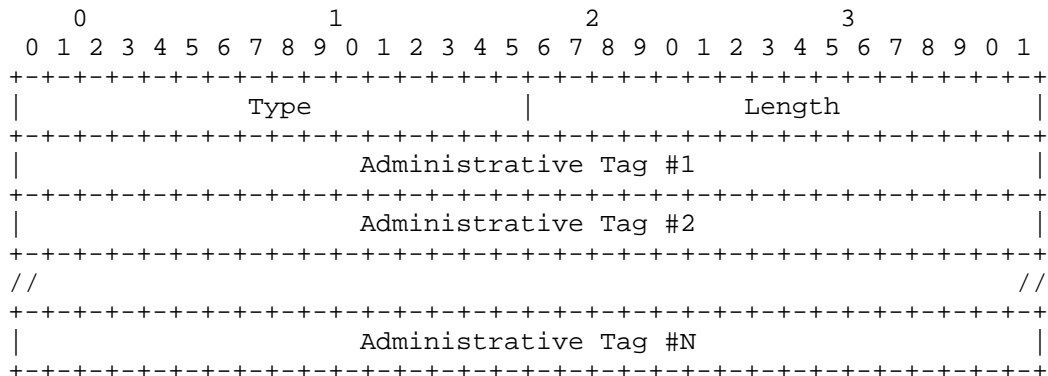


Figure 1: OSPF per-node Administrative Tag TLV

Type : TBA

Length: A 16-bit field that indicates the length of the value portion in octets and will be a multiple of 4 octets dependent on the number of tags advertised.

Value: A sequence of multiple 4 octets defining the administrative tags. The number of tags carried in this TLV is restricted to 64.

#### 4.2. Elements of procedure

Meaning of the Node administrative tags is generally opaque to OSPF. Router advertising the Node administrative tag (or tags) may be configured to do so without knowing (or even explicitly supporting) functionality implied by the tag.

Interpretation of the tag values is implementation-specific. The meaning of a Node administrative tag is defined by the network local policy and is controlled via the configuration. There are no tag values defined by this specification.

The semantics of the tag order has no meaning. That is, there is no implied meaning to the ordering of the tags that indicates a certain operation or set of operations that need to be performed based on the ordering.

Each tag SHOULD be treated as an independent identifier that MAY be used in policy to perform a policy action. Whether or not tag A precedes or succeeds tag B SHOULD not change the meaning of the tag set.

To avoid incomplete or inconsistent interpretations of the Node administrative tags the same tag value MUST NOT be advertised by a router in RI LSAs of different scopes. The same tag MAY be advertised in multiple RI LSAs of the same scope, for example, OSPF Area Border Router (ABR) may advertise the same tag in area-scope RI LSAs in multiple areas connected to the ABR.

The Node administrative tags are not meant to be extended by the future OSPF standards. The new OSPF extensions MUST NOT require use of Node administrative tags or define well-known tag values. Instead, the future OSPF extensions must define their own data signaling tailored to the needs of the feature.

Being part of the RI LSA, the Node administrative tag TLV must be reasonably small and stable. In particular, but not limited to, implementations supporting the Node administrative tags MUST NOT tie advertised tags to changes in the network topology (both within and outside the OSPF domain) or reachability of routes.

## 5. Applications

This section lists several examples of how implementations might use the Node administrative tags. These examples are given only to demonstrate generic usefulness of the router tagging mechanism. Implementation supporting this specification is not required to implement any of the use cases. It is also worth noting that in some described use cases routers configured to advertise tags help other routers in their calculations but do not themselves implement the same functionality.

### 1. Service auto-discovery

Router tagging may be used to automatically discover group of routers sharing a particular service.

For example, service provider might desire to establish full mesh of MPLS TE tunnels between all PE routers in the area of MPLS VPN network. Marking all PE routers with a tag and configuring devices with a policy to create MPLS TE tunnels to all other devices advertising this tag will automate maintenance of the full mesh. When new PE router is added to the area, all other PE devices will open TE tunnels to it without the need of reconfiguring them.

### 2. Fast-Rerouting policy

Increased deployment of Loop Free Alternates (LFA) as defined in [RFC5286] poses operation and management challenges.

[I-D.litkowski-rtgwg-lfa-manageability] proposes policies which, when implemented, will ease LFA operation concerns.

One of the proposed refinements is to be able to group the nodes in IGP domain with administrative tags and engineer the LFA based on configured policies.

(a) Administrative limitation of LFA scope

Service provider access infrastructure is frequently designed in layered approach with each layer of devices serving different purposes and thus having different hardware capabilities and configured software features. When LFA repair paths are being computed, it may be desirable to exclude devices from being considered as LFA candidates based on their layer.

For example, if the access infrastructure is divided into the Access, Distribution and Core layers it may be desirable for a Distribution device to compute LFA only via Distribution or Core devices but not via Access devices. This may be due to features enabled on Access routers; due to capacity limitations or due to the security requirements. Managing such a policy via configuration of the router computing LFA is cumbersome and error prone.

With the Node administrative tags it is possible to assign a tag to each layer and implement LFA policy of computing LFA repair paths only via neighbors which advertise the Core or Distribution tag. This requires minimal per-node configuration and network automatically adapts when new links or routers are added.

(b) LFA calculation optimization

Calculation of LFA paths may require significant resources of the router. One execution of Dijkstra algorithm is required for each neighbor eligible to become next hop of repair paths. Thus a router with a few hundreds of neighbors may need to execute the algorithm hundreds of times before the best (or even valid) repair path is found. Manually excluding from the calculation neighbors which are known to provide no valid LFA (such as single-connected routers) may significantly reduce number of Dijkstra algorithm runs.

LFA calculation policy may be configured so that routers advertising certain tag value are excluded from LFA calculation even if they are otherwise suitable.



### 3. Controlling Remote LFA tunnel termination

[I-D.ietf-rtgwg-remote-lfa] proposed method of tunneling traffic after connected link failure to extend the basic LFA coverage and algorithm to find tunnel tail-end routers fitting LFA requirement. In most cases proposed algorithm finds more than one candidate tail-end router. In real life network it may be desirable to exclude some nodes from the list of candidates based on the local policy. This may be either due to known limitations of the node (the router does not accept targeted LDP sessions required to implement Remote LFA tunneling) or due to administrative requirements (for example, it may be desirable to choose tail-end router among co-located devices).

The Node administrative tag delivers simple and scalable solution. Remote LFA can be configured with a policy to accept during the tail-end router calculation as candidates only routers advertising certain tag. Tagging routers allows to both exclude nodes not capable of serving as Remote LFA tunnel tail-ends and to define a region from which tail-end router must be selected.

### 4. Mobile backhaul network service deployment

The topology of mobile backhaul network usually adopts ring topology to save fiber resource and it is divided into the aggregate network and the access network. Cell Site Gateways(CSGs) connects the eNodeBs and RNC(Radio Network Controller) Site Gateways(RSGs) connects the RNCs. The mobile traffic is transported from CSGs to RSGs. The network takes a typical aggregate traffic model that more than one access rings will attach to one pair of aggregate site gateways(ASGs) and more than one aggregate rings will attach to one pair of RSGs.

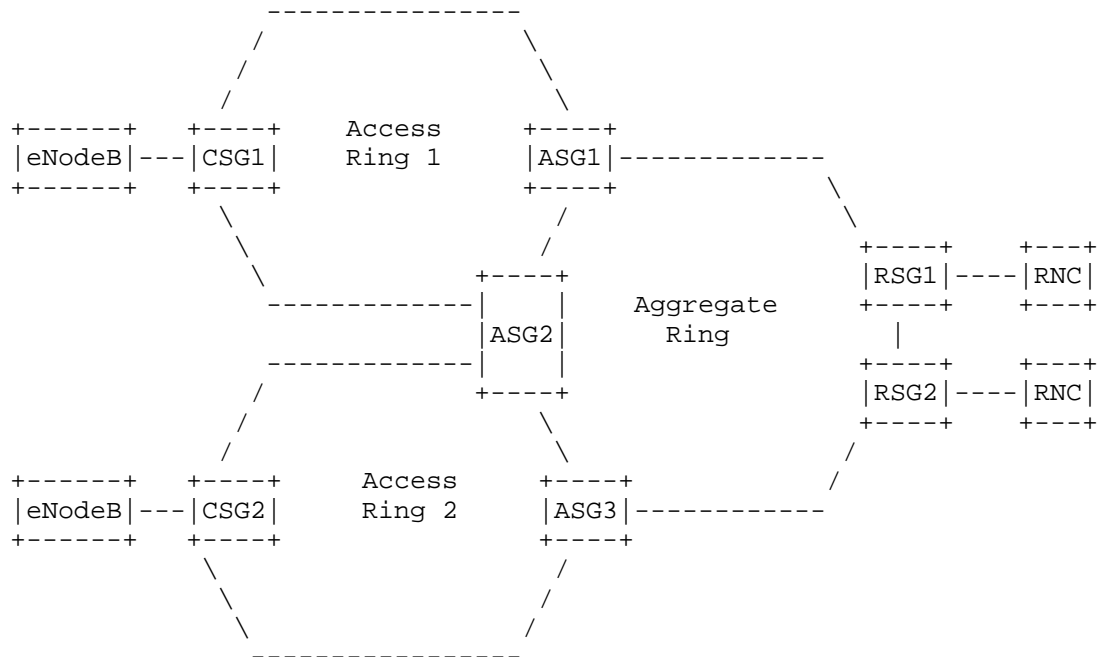


Figure 2: Mobile Backhaul Network

A typical mobile backhaul network with access rings and aggregate links is shown in figure above. The mobile backhaul networks deploy traffic engineering due to the strict Service Level Agreements(SLA). The TE paths may have additional constraints to avoid passing via different access rings or to get completely disjoint backup TE paths. The mobile backhaul networks towards the access side change frequently due to the growing mobile traffic and addition of new eNodeBs. It's complex to satisfy the requirements using cost, link color or explicit path configurations. The node administrative tag defined in this document can be effectively used to solve the problem for mobile backhaul networks. The nodes in different rings can be assigned with specific tags. TE path computation can be enhanced to consider additional constraints based on node administrative tags.

## 6. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC2328] and [RFC5340].

## 7. IANA Considerations

IANA maintains the registry for the TLVs. OSPF Administrative Tags will require one new type code for the TLV defined in this document.

## 8. Acknowledgments

Thanks to Bharath R and Pushpasis Sarakar for useful inputs. Thanks to Chris Bowers for providing useful inputs to remove ambiguity related to tag-ordering.

## 9. References

### 9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 4970, July 2007.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

### 9.2. Informative References

- [I-D.ietf-rtgwg-remote-lfa]  
Bryant, S., Filsfils, C., Previdi, S., Shand, M., and S. Ning, "Remote LFA FRR", draft-ietf-rtgwg-remote-lfa-02 (work in progress), May 2013.
- [I-D.litkowski-rtgwg-lfa-manageability]  
Litkowski, S., Decraene, B., Filsfils, C., and K. Raza, "Operational management of Loop Free Alternates", draft-litkowski-rtgwg-lfa-manageability-01 (work in progress), February 2013.
- [RFC5286] Atlas, A. and A. Zinin, "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, September 2008.

Authors' Addresses

Shraddha Hegde  
Juniper Networks, Inc.  
Embassy Business Park  
Bangalore, KA 560093  
India

Email: shraddha@juniper.net

Harish Raghuveer  
Juniper Networks, Inc.  
Embassy Business Park  
Bangalore 560093  
India

Email: hraghuveer@juniper.net

Hannes Gredler  
Juniper Networks, Inc.  
1194 N. Mathilda Ave.  
Sunnyvale, CA 94089  
US

Email: hannes@juniper.net

Rob Shakir  
British Telecom

Email: rob.shakir@bt.com

Anton Smirnov  
Cisco Systems, Inc.  
De Kleetlaan 6a  
Diegem 1831  
Belgium

Email: as@cisco.com

Li Zhenbin  
Huawei Technologies  
Huawei Bld. No.156 Beiqing Rd  
Beijing 100095  
China

Email: lizhenbin@huawei.com

Network Working Group  
Internet-Draft  
Updates: 5340 (if approved)  
Intended status: Standards Track  
Expires: August 14, 2015

A. Lindem  
Cisco Systems  
J. Arkko  
Ericsson  
February 10, 2015

OSPFv3 Auto-Configuration  
draft-ietf-ospf-ospfv3-autoconfig-15.txt

## Abstract

OSPFv3 is a candidate for deployments in environments where auto-configuration is a requirement. One such environment is the IPv6 home network where users expect to simply plug in a router and have it automatically use OSPFv3 for intra-domain routing. This document describes the necessary mechanisms for OSPFv3 to be self-configuring. This document updates RFC 5340 by relaxing the HelloInterval/RouterDeadInterval checking during OSPFv3 adjacency formation and adding hysteresis to the update of self-originated Link State Advertisements (LSAs).

## 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 <http://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 August 14, 2015.

## Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	2
1.1. Requirements notation . . . . .	3
2. OSPFv3 Default Configuration . . . . .	3
3. OSPFv3 HelloInterval/RouterDeadInterval Flexibility . . . . .	4
3.1. Wait Timer Reduction . . . . .	4
4. OSPFv3 Minimal Authentication Configuration . . . . .	5
5. OSPFv3 Router ID Selection . . . . .	5
6. OSPFv3 Adjacency Formation . . . . .	5
7. OSPFv3 Duplicate Router ID Detection and Resolution . . . . .	6
7.1. Duplicate Router ID Detection for Neighbors . . . . .	6
7.2. Duplicate Router ID Detection for Non-Neighbors . . . . .	6
7.2.1. OSPFv3 Router Auto-Configuration LSA . . . . .	7
7.2.2. Router-Hardware-Fingerprint TLV . . . . .	8
7.3. Duplicate Router ID Resolution . . . . .	9
7.4. Change to RFC 2328 Section 13.4, 'Receiving Self- Originated LSAs' . . . . .	9
8. Security Considerations . . . . .	10
9. Management Considerations . . . . .	10
10. IANA Considerations . . . . .	11
11. Acknowledgments . . . . .	11
12. References . . . . .	13
12.1. Normative References . . . . .	13
12.2. Informative References . . . . .	13
Authors' Addresses . . . . .	14

## 1. Introduction

OSPFv3 [OSPFV3] is a candidate for deployments in environments where auto-configuration is a requirement. This document describes extensions to OSPFv3 to enable it to operate in these environments. In this mode of operation, the protocol is largely unchanged from the base OSPFv3 protocol specification [OSPFV3]. Since the goals of auto-configuration and security can be conflicting, operators and network administrators should carefully consider their security requirements before deploying the solution described in this document. Refer to Section 8 for more information.

The following aspects of OSPFv3 auto-configuration are described in this document:

1. Default OSPFv3 Configuration
2. HelloInterval/RouterDeadInterval Flexibility
3. Unique OSPFv3 Router ID generation
4. OSPFv3 Adjacency Formation
5. Duplicate OSPFv3 Router ID Resolution
6. Self-Originated LSA Processing

OSPFv3 [OSPFV3] is updated by allowing OSPFv3 adjacencies to be formed between OSPFv3 routers with differing HelloIntervals or RouterDeadIntervals (refer to Section 3). Additionally, hysteresis has been added to the processing of stale self-originated LSAs to mitigate the flooding overhead created by an OSPFv3 Router with a duplicate OSPFv3 Router ID in the OSPFv3 routing domain (refer to Section 7.4. Both updates are fully backward compatible.

#### 1.1. Requirements notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC-KEYWORDS].

#### 2. OSPFv3 Default Configuration

For complete auto-configuration, OSPFv3 will need to choose suitable configuration defaults. These include:

1. Area 0 Only - All auto-configured OSPFv3 interfaces MUST be in area 0.
2. OSPFv3 SHOULD be auto-configured on all IPv6-capable interface on the router. An interface MAY be excluded if it is clear that running OSPFv3 on the interface is not required. For example, if manual configuration or another condition indicates that an interface is connected to an Internet Service Provider (ISP), there is typically no need to employ OSPFv3. In fact, [IPv6-CPE] specifically requires that IPv6 Customer Premise Equipment (CPE) routers do not initiate any dynamic routing protocol by default on the router's WAN, i.e., ISP-facing, interface. In home networking environments, an interface where no OSPFv3 neighbors are found but a DHCP IPv6 prefix can be acquired may be considered an ISP-facing interface and running OSPFv3 is unnecessary.



3. OSPFv3 interfaces will be auto-configured to an interface type corresponding to their layer-2 capability. For example, Ethernet interfaces and Wi-Fi interfaces will be auto-configured as OSPFv3 broadcast networks and Point-to-Point Protocol (PPP) interfaces will be auto-configured as OSPFv3 Point-to-Point interfaces. Most extant OSPFv3 implementations do this already. Auto-configured operation over wireless networks requiring a point-to-multipoint (P2MP) topology and dynamic metrics based on wireless feedback is not within the scope of this document. However, auto-configuration is not precluded in these environments.
4. OSPFv3 interfaces MAY use an arbitrary HelloInterval and RouterDeadInterval as specified in Section 3. Of course, an identical HelloInterval and RouterDeadInterval will still be required to form an adjacency with an OSPFv3 router not supporting auto-configuration [OSPFV3].
5. All OSPFv3 interfaces SHOULD be auto-configured to use an Interface Instance ID of 0 that corresponds to the base IPv6 unicast address family instance ID as defined in [OSPFV3-AF]. Similarly, if IPv4 unicast addresses are advertised in a separate auto-configured OSPFv3 instance, the base IPv4 unicast address family instance ID value, i.e., 64, SHOULD be auto-configured as the Interface Instance ID for all interfaces corresponding to the IPv4 unicast OSPFv3 instance [OSPFV3-AF].

### 3. OSPFv3 HelloInterval/RouterDeadInterval Flexibility

Auto-configured OSPFv3 routers will not require an identical HelloInterval and RouterDeadInterval to form adjacencies. Rather, the received HelloInterval will be ignored and the received RouterDeadInterval will be used to determine OSPFv3 liveness with the sending router. In other words, the Neighbor Inactivity Timer (Section 10 of [OSPFV2]) for each neighbor will reflect that neighbor's advertised RouterDeadInterval and MAY be different from other OSPFv3 routers on the link without impacting adjacency formation. A similar mechanism requiring additional signaling is proposed for all OSPFv2 and OSPFv3 routers [ASYNCH-HELLO].

#### 3.1. Wait Timer Reduction

In many situations, auto-configured OSPFv3 routers will be deployed in environments where back-to-back ethernet connections are utilized. When this is the case, an OSPFv3 broadcast interface will not come up until the other OSPFv3 router is connected and the routers will wait RouterDeadInterval seconds before forming an adjacency [OSPFV2]. In order to reduce this delay, an auto-configured OSPFv3 router MAY reduce the wait interval to a value no less than (HelloInterval + 1).

Reducing the setting will slightly increase the likelihood of the Designated Router (DR) flapping but is preferable to the long adjacency formation delay. Note that this value is not included in OSPFv3 Hello packets and does not impact interoperability.

#### 4. OSPFv3 Minimal Authentication Configuration

In many deployments, the requirement for OSPFv3 authentication overrides the goal of complete OSPFv3 autoconfiguration. Therefore, it is RECOMMENDED that OSPFv3 routers supporting this specification minimally offer an option to explicitly configure a single password for HMAC-SHA authentication as described in [OSPFV3-AUTH-TRAILER]. It is RECOMMENDED that the password entered as ASCII hexadecimal digits and that 32 or more digits to facilitate a password with a high degree of entropy. When configured, the password will be used on all auto-configured interfaces with the Security Association Identifier (SA ID) set to 1 and HMAC-SHA-256 used as the authentication algorithm.

#### 5. OSPFv3 Router ID Selection

An OSPFv3 router requires a unique Router ID within the OSPFv3 routing domain for correct protocol operation. Existing Router ID selection algorithms (section C.1 in [OSPFV2] and [OSPFV3]) are not viable since they are dependent on a unique IPv4 interface address which is not likely to be available in autoconfigured deployments. An OSPFv3 router implementing this specification will select a router-id that has a high probability of uniqueness. A pseudo-random number SHOULD be used for the OSPFv3 Router ID. The generation SHOULD be seeded with a variable that is likely to be unique in the applicable OSPFv3 router deployment. A good choice of seed would be some portion or hash of the Router-Hardware-Fingerprint as described in Section 7.2.2.

Since there is a possibility of a Router ID collision, duplicate Router ID detection and resolution are required as described in Section 7 and Section 7.3. OSPFv3 routers SHOULD maintain the last successfully chosen Router ID in non-volatile storage to avoid collisions subsequent to when an autoconfigured OSPFv3 router is first added to the OSPFv3 routing domain.

#### 6. OSPFv3 Adjacency Formation

Since OSPFv3 uses IPv6 link-local addresses for all protocol messages other than messages sent on virtual links (which are not applicable to auto-configuration), OSPFv3 adjacency formation can proceed as soon as a Router ID has been selected and the IPv6 link-local address has completed Duplicate Address Detection (DAD) as specified in IPv6

Stateless Address Autoconfiguration [SLAAC]. Otherwise, the only changes to the OSPFv3 base specification are supporting HelloInterval/RouterDeadInterval flexibility as described in Section 3 and duplicate Router ID detection and resolution as described in Section 7 and Section 7.3.

## 7. OSPFv3 Duplicate Router ID Detection and Resolution

There are two cases of duplicate OSPFv3 Router ID detection. One where the OSPFv3 router with the duplicate Router ID is directly connected and one where it is not. In both cases, the duplicate resolution is for one of the routers to select a new OSPFv3 Router ID.

### 7.1. Duplicate Router ID Detection for Neighbors

In this case, a duplicate Router ID is detected if any valid OSPFv3 packet is received with the same OSPFv3 Router ID but a different IPv6 link-local source address. Once this occurs, the OSPFv3 router with the numerically smaller IPv6 link-local address will need to select a new Router ID as described in Section 7.3. Note that the fact that the OSPFv3 router is a neighbor on a non-virtual interface implies that the router is directly connected. An OSPFv3 router implementing this specification should assure that the inadvertent connection of multiple router interfaces to the same physical link is not misconstrued as detection of an OSPFv3 neighbor with a duplicate Router ID.

### 7.2. Duplicate Router ID Detection for Non-Neighbors

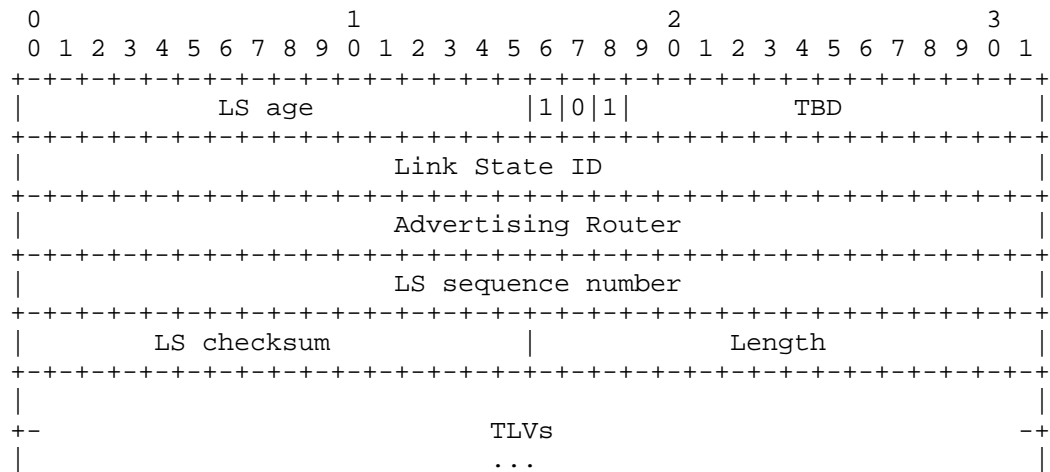
OSPFv3 routers implementing auto-configuration, as specified herein, MUST originate an Auto-Configuration (AC) Link State Advertisement (LSA) including the Router-Hardware-Fingerprint Type-Length-Value (TLV). The Router-Hardware-Fingerprint TLV contains a variable length value that has a very high probability of uniquely identifying the advertising OSPFv3 router. An OSPFv3 router implementing this specification MUST detect received Auto-Configuration LSAs with its Router ID specified in the LSA header. LSAs received with the local OSPFv3 Router's Router ID in the LSA header are perceived as self-originated (see section 4.6 of [OSPFV3]). In these received Auto-Configuration LSAs, the Router-Hardware-Fingerprint TLV is compared against the OSPFv3 Router's own router hardware fingerprint. If the fingerprints are not equal, there is a duplicate Router ID conflict and the OSPFv3 router with the numerically smaller router hardware fingerprint MUST select a new Router ID as described in Section 7.3.

This new LSA is designated for information related to OSPFv3 Auto-configuration and, in the future, could be used for other auto-

configuration information, e.g., global IPv6 prefixes. However, this is beyond the scope of this document.

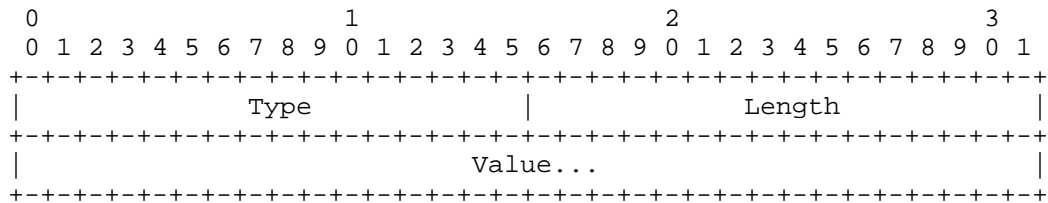
### 7.2.1. OSPFv3 Router Auto-Configuration LSA

The OSPFv3 Auto-Configuration (AC) LSA has a function code of TBD and the S2/S1 bits set to 01 indicating Area Flooding Scope. The U bit will be set indicating that the OSPFv3 AC LSA should be flooded even if it is not understood. The Link State ID (LSID) value will be a integer index used to discriminate between multiple AC LSAs originated by the same OSPFv3 router. This specification only describes the contents of an AC LSA with a Link State ID (LSID) of 0.



### OSPFv3 Auto-Configuration (AC) LSA

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



#### TLV Format

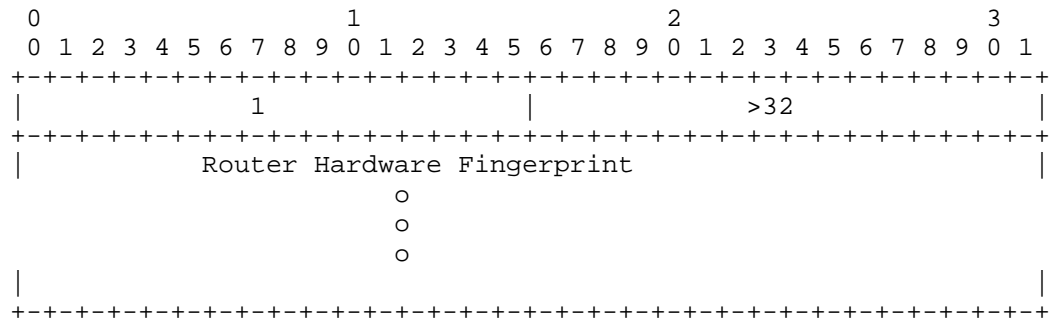
The Length field defines the length of the value portion in octets (thus a TLV with no value portion would have a length of 0). The TLV is padded to 4-octet alignment; padding is not included in the length field (so a 3-octet value would have a length of 3, but the total size of the TLV would be 8 octets). Nested TLVs are also 32-bit aligned. For example, a 1-byte value would have the length field set to 1, and 3 octets of padding would be added to the end of the value portion of the TLV. Unrecognized types are ignored.

The new LSA is designated for information related to OSPFv3 Auto-configuration and, in the future, can be used other auto-configuration information.

#### 7.2.2. Router-Hardware-Fingerprint TLV

The Router-Hardware-Fingerprint TLV is the first TLV defined for the OSPFv3 Auto-Configuration (AC) LSA. It will have type 1 and MUST be advertised in the LSID OSPFv3 AC LSA with an LSID of 0. It SHOULD occur, at most, once and the first instance of the TLV will take precedence over subsequent TLV instances. The length of the Router-Hardware-Fingerprint is variable but must be 32 octets or greater. If the Router-Hardware-Fingerprint TLV is not present as the first TLV, the AC-LSA is considered malformed and is ignored for the purposes of duplicate Router ID detection. Additionally, the event SHOULD be logged.

The contents of the hardware fingerprint MUST have an extremely high probability of uniqueness. It SHOULD be constructed from the concatenation of a number of local values that themselves have a high likelihood of uniqueness, such as MAC addresses, CPU ID, or serial numbers. It is RECOMMENDED that one or more available universal tokens (e.g., IEEE 802 48-bit MAC addresses or IEEE EUI-64 Identifiers [EUI64]) associated with the OSPFv3 router be included in the hardware fingerprint. It MUST be based on hardware attributes that will not change across hard and soft restarts.



Router-Hardware-Fingerprint TLV Format

### 7.3. Duplicate Router ID Resolution

The OSPFv3 router selected to resolve the duplicate OSPFv3 Router ID condition must select a new OSPFv3 Router ID. The OSPFv3 router SHOULD reduce the possibility of a subsequent Router ID collision by checking the Link State Database for an OSPFv3 Auto-Configuration LSA with the newly selected Router ID and a different Router-Hardware-Fingerprint. If one is detected, a new Router ID should be selected without going through the resolution process Section 7. After selecting a new Router ID, all self-originated LSAs MUST be reoriginated, and any OSPFv3 neighbor adjacencies MUST be reestablished. The OSPFv3 router retaining the Router ID causing the conflict will reoriginate or purge stale any LSAs as described in Section 13.4 [OSPFV2].

### 7.4. Change to RFC 2328 Section 13.4, 'Receiving Self-Originated LSAs'

RFC 2328 [OSPFV2], Section 13.4, describes the processing of received self-originated LSAs. If the received LSA doesn't exist, the receiving router will purge it from the OSPF routing domain. If the LSA is newer than the version in the Link State Database (LSDB), the receiving router will originate a newer version by advancing the LSA sequence number and reoriginating. Since it is possible for an auto-configured OSPFv3 router to choose a duplicate OSPFv3 Router ID, OSPFv3 routers implementing this specification should detect when multiple instances of the same self-originated LSA are purged or reoriginated since this is indicative of an OSPFv3 router with a duplicate Router ID in the OSPFv3 routing domain. When this condition is detected, the OSPFv3 router SHOULD delay self-originated LSA processing for LSAs that have recently been purged or reoriginated. This specification recommends 10 seconds as the interval defining recent self-originated LSA processing and an exponential back off of 1 to 8 seconds for the processing delay.

This additional delay should allow for the mechanisms described in Section 7 to resolve the duplicate OSPFv3 Router ID conflict.

Since this mechanism is useful in mitigating the flooding overhead associated with the inadvertent or malicious introduction of an OSPFv3 router with a duplicate Router ID into an OSPFv3 routing domain, it MAY be deployed outside of autoconfigured deployments. The detection of a self-originated LSA that is being repeated reoriginated or purged SHOULD be logged.

## 8. Security Considerations

A unique OSPFv3 Interface Instance ID is used for auto-configuration to prevent inadvertent OSPFv3 adjacency formation, see Section 2

The goals of security and complete OSPFv3 auto-configuration are somewhat contradictory. When no explicit security configuration takes place, auto-configuration implies that additional devices placed in the network are automatically adopted as a part of the network. However, auto-configuration can also be combined with password configuration (see Section 4) or future extensions for automatic pairing between devices. These mechanisms can help provide an automatically configured, securely routed network.

In deployments where different authentication algorithm, per-interface keys, or encryption is required, OSPFv3 IPsec [OSPFV3-IPSEC] or alternate OSPFv3 Authentication trailer [OSPFV3-AUTH-TRAILER] algorithms MAY be used at the expense of additional configuration. The configuration and operational description of such deployments is beyond the scope of this document. However, a deployment could always revert to explicit configuration as described in Section 9 for features such as IPsec, per-interface keys, or alternate authentication algorithms.

The introduction, either malicious or accidental, of an OSPFv3 router with a duplicate Router ID is an attack point for OSPFv3 routing domains. This is due to the fact that OSPFv3 routers will interpret LSAs advertised by the router with the same Router ID as self-originated LSAs and attempt to purge them from the routing domain. The mechanisms in Section 7.4 will mitigate the effects of duplication.

## 9. Management Considerations

It is RECOMMENDED that OSPFv3 routers supporting this specification also support explicit configuration of OSPFv3 parameters as specified in Appendix C of [OSPFV3]. This would allow explicit override of autoconfigured parameters in situations where it is required (e.g.,

if the deployment requires multiple OSPFv3 areas). This is in addition to the authentication key configuration recommended in Section 4. Additionally, it is RECOMMENDED that OSPFv3 routers supporting this specification allow autoconfiguration to be completely disabled.

Since there is a small possibility of OSPFv3 Router ID collisions, manual configuration of OSPFv3 Router IDs is RECOMMENDED in OSPFv3 routing domains where route convergence due to a router ID change is intolerable.

OSPFv3 Routers supporting this specification MUST augment mechanisms for displaying or otherwise conveying OSPFv3 operational state to indicate whether or not the OSPFv3 router was autoconfigured and whether or not its OSPFv3 interfaces have been auto-configured.

#### 10. IANA Considerations

This specification defines an OSPFv3 LSA Type for the OSPFv3 Auto-Configuration (AC) LSA, as described in Section 7.2.1. The value TBD will be allocated from the existing "OSPFv3 LSA Function Code" registry for the OSPFv3 Auto-Configuration LSA.

This specification also creates a registry for OSPFv3 Auto-Configuration (AC) LSA TLVs. This registry should be placed in the existing OSPFv3 IANA registry, and new values can be allocated via IETF Review or, under exceptional circumstances, IESG Approval. [IANA-GUIDELINES]

Three initial values are allocated:

- o 0 is marked as reserved.
- o 1 is Router-Hardware-Fingerprint TLV (Section 7.2.2).
- o 65535 is an Auto-configuration-Experiment-TLV, a common value that can be used for experimental purposes.

#### 11. Acknowledgments

This specification was inspired by the work presented in the Homenet working group meeting in October 2011 in Philadelphia, Pennsylvania. In particular, we would like to thank Fred Baker, Lorenzo Colitti, Ole Troan, Mark Townsley, and Michael Richardson.

Arthur Dimitrelis and Aidan Williams did prior work in OSPFv3 auto-configuration in the expired "Autoconfiguration of routers using a



link state routing protocol" IETF Draft. There are many similarities between the concepts and techniques in this document.

Thanks for Abhay Roy and Manav Bhatia for comments regarding duplicate router-id processing.

Thanks for Alvaro Retana and Michael Barnes for comments regarding OSPFv3 Instance ID auto-configuration.

Thanks to Faraz Shamim for review and comments.

Thanks to Mark Smith for the requirement to reduce the adjacency formation delay in the back-to-back ethernet topologies that are prevalent in home networks.

Thanks to Les Ginsberg for document review and recommendations on OSPFv3 hardware fingerprint content.

Thanks to Curtis Villamizar for document review and analysis of duplicate router-id resolution nuances.

Thanks to Uma Chunduri for comments during OSPF WG last call.

Thanks to Martin Vigoureux for Routing Area Directorate review and comments.

Thanks to Adam Montville for Security Area Directorate review and comments.

Thanks to Qin Wu for Operations & Management Area Directorate review and comments.

Thanks to Robert Sparks for General Area (GEN-ART) review and comments.

Thanks to Rama Darbha for review and comments.

Special thanks to Adrian Farrel for his in-depth review, copious comments, and suggested text.

Special thanks go to Markus Stenberg for his implementation of this specification in Bird.

Special thanks also go to David Lamparter for his implementation of this specification in Quagga.

The RFC text was produced using Marshall Rose's xml2rfc tool.

## 12. References

### 12.1. Normative References

- [OSPFV2] Moy, J., "OSPF Version 2", RFC 2328, April 1998.
- [OSPFV3] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [OSPFV3-AF] Lindem, A., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC 5838, April 2010.
- [OSPFV3-AUTH-TRAILER] Bhatia, M., Manral, V., and A. Lindem, "Supporting Authentication Trailer for OSPFv3", RFC 7166, February 2012.
- [RFC-KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.
- [SLAAC] Thomson, S., Narten, T., and J. Tatuya, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.
- [TE] Katz, D., Yeung, D., and K. Kompella, "Traffic Engineering Extensions to OSPF", RFC 3630, September 2003.

### 12.2. Informative References

- [ASYNCH-HELLO] Anand, M., Grover, H., and A. Roy, "Asymmetric OSPF Hold Timer", draft-madhukar-ospf-agr-asymmetric-01.txt (work in progress), June 2013.
- [EUI64] IEEE, "Guidelines for 64-bit Global Identifier (EUI-64) Registration Authority", IEEE Tutorial <http://standards.ieee.org/regauth/oui/tutorials/EUI64.html>, March 1997.
- [IANA-GUIDELINES] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, May 2008.

[IPv6-CPE]

Singh, H., Beebee, W., Donley, C., and B. Stark, "Basic Requirements for IPv6 Customer Edge Routers", RFC 7084, November 2013.

[OSPFV3-IPSEC]

Gupta, M. and S. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, June 2006.

Authors' Addresses

Acee Lindem  
Cisco Systems  
301 Midenhall Way  
Cary, NC 27513  
USA

Email: [acee@cisco.com](mailto:acee@cisco.com)

Jari Arkko  
Ericsson  
Jorvas, 02420  
Finland

Email: [jari.arkko@piuha.net](mailto:jari.arkko@piuha.net)

Network Working Group  
Internet-Draft  
Updates: 5340, 5838 (if approved)  
Intended status: Standards Track  
Expires: July 29, 2018

A. Lindem  
A. Roy  
Cisco Systems  
D. Goethals  
Nokia  
V. Reddy Vallem

F. Baker  
January 25, 2018

OSPFv3 LSA Extendibility  
draft-ietf-ospf-ospfv3-lsa-extend-23.txt

Abstract

OSPFv3 requires functional extension beyond what can readily be done with the fixed-format Link State Advertisement (LSA) as described in RFC 5340. Without LSA extension, attributes associated with OSPFv3 links and advertised IPv6 prefixes must be advertised in separate LSAs and correlated to the fixed-format LSAs. This document extends the LSA format by encoding the existing OSPFv3 LSA information in Type-Length-Value (TLV) tuples and allowing advertisement of additional information with additional TLVs. Backward compatibility mechanisms are also described.

This document updates RFC 5340, "OSPF for IPv6", and RFC 5838, "Support of Address Families in OSPFv3" by providing TLV-based encodings for the base OSPFv3 unicast support and OSPFv3 address family support.

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 <http://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 July 29, 2018.

## Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
1.1. Requirements notation . . . . .	4
1.2. OSPFv3 LSA Terminology . . . . .	4
2. OSPFv3 Extended LSA Types . . . . .	4
3. OSPFv3 Extended LSA TLVs . . . . .	5
3.1. Prefix Options Extensions . . . . .	6
3.1.1. N-bit Prefix Option . . . . .	6
3.2. Router-Link TLV . . . . .	7
3.3. Attached-Routers TLV . . . . .	8
3.4. Inter-Area-Prefix TLV . . . . .	10
3.5. Inter-Area-Router TLV . . . . .	11
3.6. External-Prefix TLV . . . . .	12
3.7. Intra-Area-Prefix TLV . . . . .	13
3.8. IPv6 Link-Local Address TLV . . . . .	14
3.9. IPv4 Link-Local Address TLV . . . . .	15
3.10. IPv6-Forwarding-Address Sub-TLV . . . . .	16
3.11. IPv4-Forwarding-Address Sub-TLV . . . . .	16
3.12. Route-Tag Sub-TLV . . . . .	17
4. OSPFv3 Extended LSAs . . . . .	17
4.1. OSPFv3 E-Router-LSA . . . . .	17
4.2. OSPFv3 E-Network-LSA . . . . .	19
4.3. OSPFv3 E-Inter-Area-Prefix-LSA . . . . .	20
4.4. OSPFv3 E-Inter-Area-Router-LSA . . . . .	21
4.5. OSPFv3 E-AS-External-LSA . . . . .	22
4.6. OSPFv3 E-NSSA-LSA . . . . .	23
4.7. OSPFv3 E-Link-LSA . . . . .	24
4.8. OSPFv3 E-Intra-Area-Prefix-LSA . . . . .	26
5. Malformed OSPFv3 Extended LSA Handling . . . . .	27
6. LSA Extension Backward Compatibility . . . . .	27
6.1. Full Extended LSA Migration . . . . .	27
6.2. Extended LSA Sparse-Mode Backward Compatibility . . . . .	28

6.3. LSA TLV Processing Backward Compatibility . . . . .	28
7. Security Considerations . . . . .	29
8. IANA Considerations . . . . .	29
8.1. OSPFv3 Extended-LSA TLV Registry . . . . .	29
8.2. OSPFv3 Extended-LSA sub-TLV Registry . . . . .	30
9. Contributors . . . . .	31
10. References . . . . .	31
10.1. Normative References . . . . .	31
10.2. Informative References . . . . .	31
Appendix A. Appendix A - Global Configuration Parameters . . . . .	32
Appendix B. Appendix B - Area Configuration Parameters . . . . .	32
Appendix C. Acknowledgments . . . . .	33
Authors' Addresses . . . . .	33

## 1. Introduction

OSPFv3 requires functional extension beyond what can readily be done with the fixed-format Link State Advertisement (LSA) as described in RFC 5340 [OSPFV3]. Without LSA extension, attributes associated with OSPFv3 links and advertised IPv6 prefixes must be advertised in separate LSAs and correlated to the fixed-format LSAs. This document extends the LSA format by encoding the existing OSPFv3 LSA information in Type-Length-Value (TLV) tuples and allowing advertisement of additional information with additional TLVs. Backward compatibility mechanisms are also described.

This document updates RFC 5340, "OSPF for IPv6", and RFC 5838, "Support of Address Families in OSPFv3" by providing TLV-based encodings for the base OSPFv3 support [OSPFV3] and OSPFv3 address family support [OSPFV3-AF].

A similar extension was previously proposed in support of multi-topology routing. Additional requirements for OSPFv3 LSA extension include source/destination routing, route tagging, and others.

A final requirement is to limit the changes to OSPFv3 to those necessary for TLV-based LSAs. For the most part, the semantics of existing OSPFv3 LSAs are retained for their TLV-based successor LSAs described herein. Additionally, encoding details, e.g., the representation of IPv6 prefixes as described in section A.4.1 in RFC 5340 [OSPFV3], have been retained. This requirement was included to increase the expedience of IETF adoption and deployment.

The following aspects of OSPFv3 LSA extension are described:

1. Extended LSA Types
2. Extended LSA TLVs

### 3. Extended LSA Formats

### 4. Backward Compatibility

#### 1.1. Requirements notation

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.

#### 1.2. OSPFv3 LSA Terminology

The TLV-based OSPFv3 LSAs described in this document will be referred to as Extended LSAs. The OSPFv3 fixed-format LSAs [OSPFV3] will be referred to as Legacy LSAs.

## 2. OSPFv3 Extended LSA Types

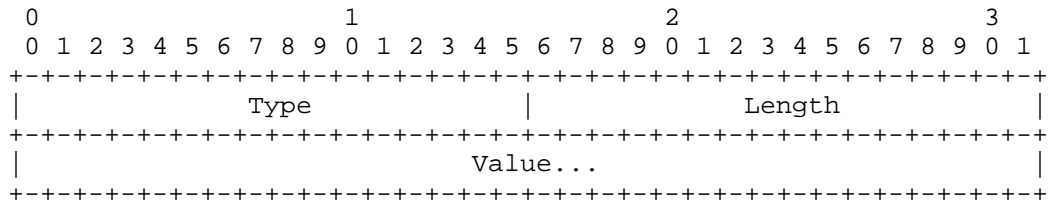
In order to provide backward compatibility, new LSA codes must be allocated. There are eight fixed-format LSAs defined in RFC 5340 [OSPFV3]. For ease of implementation and debugging, the LSA function codes are the same as the fixed-format LSAs only with 32, i.e., 0x20, added. The alternative to this mapping was to allocate a bit in the LS Type indicating the new LSA format. However, this would have used one half the LSA function code space for the migration of the eight original fixed-format LSAs. For backward compatibility, the U-bit MUST be set in LS Type so that the LSAs will be flooded by OSPFv3 routers that do not understand them.

LSA function code	LS Type	Description
33	0xA021	E-Router-LSA
34	0xA022	E-Network-LSA
35	0xA023	E-Inter-Area-Prefix-LSA
36	0xA024	E-Inter-Area-Router-LSA
37	0xC025	E-AS-External-LSA
38	N/A	Unused (Not to be allocated)
39	0xA027	E-Type-7-LSA
40	0x8028	E-Link-LSA
41	0xA029	E-Intra-Area-Prefix-LSA

#### OSPFv3 Extended LSA Types

### 3. OSPFv3 Extended LSA TLVs

The format of the TLVs within the body of the extended LSAs is the same as the format used by the Traffic Engineering Extensions to OSPF [TE]. The variable TLV section consists of one or more nested Type/Length/Value (TLV) tuples. Nested TLVs are also referred to as sub-TLVs. The format of each TLV is:



TLV Format

The Length field defines the length of the value portion in octets (thus, a TLV with no value portion would have a length of 0). The TLV is padded to 4-octet alignment; padding is not included in the length field (so a 3-octet value would have a length of 3, but the total size of the TLV would be 8 octets). Nested TLVs are also 32-bit aligned. For example, a 1-byte value would have the length field set to 1, and 3 octets of padding would be added to the end of the value portion of the TLV.

This document defines the following top-level TLV types:

- o 0 - Reserved
- o 1 - Router-Link TLV
- o 2 - Attached-Routers TLV
- o 3 - Inter-Area Prefix TLV
- o 4 - Inter-Area Router TLV
- o 5 - External Prefix TLV
- o 6 - Intra-Area Prefix TLV
- o 7 - IPv6 Link-Local Address TLV
- o 8 - IPv4 Link-Local Address TLV



Additionally, this document defines the following sub-TLV types:

- o 0 - Reserved
- o 1 - IPv6 Forwarding Address sub-TLV
- o 2 - IPv4 Forwarding Address sub-TLV
- o 3 - Route Tag sub-TLV

In general, TLVs and sub-TLVs MAY occur in any order and the specification should define whether the TLV or sub-TLV is required and the behavior when there are multiple occurrences of the TLV or sub-TLV. While this document only describes the usage of TLVs and Sub-TLVs, Sub-TLVs may be nested to any level as long as the Sub-TLVs are fully specified in the specification for the subsuming Sub-TLV.

For backward compatibility, an LSA is not considered malformed from a TLV perspective unless either a required TLV is missing or a specified TLV is less than the minimum required length. Refer to Section 6.3 for more information on TLV backward compatibility.

### 3.1. Prefix Options Extensions

The prefix options are extended from Appendix A.4.1.1 [OSPFV3]. The applicability of the LA-bit is expanded and it SHOULD be set in Inter-Area-Prefix-TLVs and MAY be set in External-Prefix-TLVs when the advertised host IPv6 address, i.e., PrefixLength = 128, is an interface address. In RFC 5340, the LA-bit is only set in Intra-Area-Prefix-LSAs (Section 4.4.3.9 in [OSPFV3]). This will allow a stable address to be advertised without having to configure a separate loopback address in every OSPFv3 area.

#### 3.1.1. N-bit Prefix Option

Additionally, the N-bit prefix option is defined. The figure below shows the position of the N-bit in the prefix options (pending IANA allocation). This corresponds to the value 0x20.

```

      0  1  2  3  4  5  6  7
+---+---+---+---+---+---+---+
|   |   | N|DN| P| x|LA|NU|
+---+---+---+---+---+---+---+

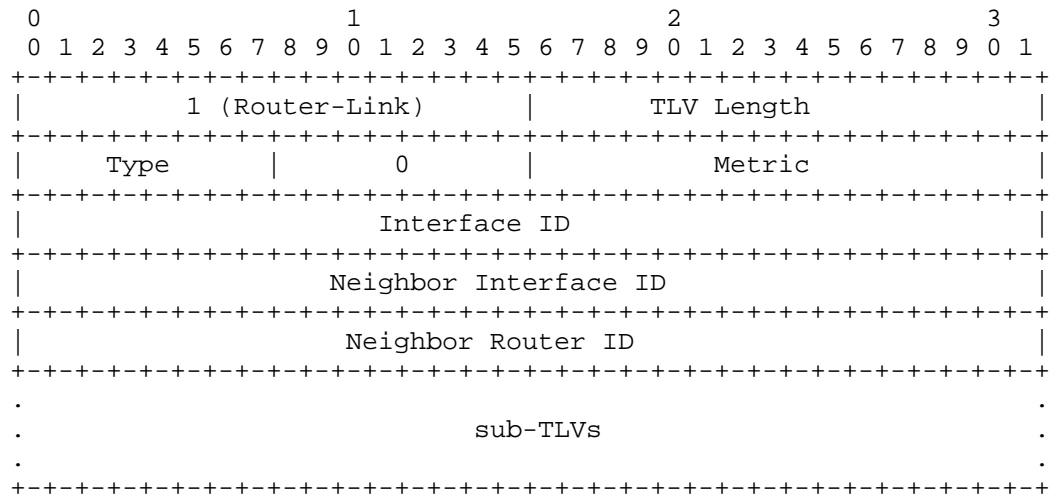
```

The Prefix Options field

The N-bit is set in PrefixOptions for a host address (PrefixLength=128) that identifies the advertising router. While it is similar to the LA-bit, there are two differences. The advertising router MAY choose NOT to set the N-bit even when the above conditions are met. If the N-bit is set and the PrefixLength is NOT 128, the N-bit MUST be ignored. Additionally, the N-bit is propagated in the PrefixOptions when an OSPFv3 Area Border Router (ABR) originates an Inter-Area-Prefix-LSA for an Intra-Area route which has the N-bit set in the PrefixOptions. Similarly, the N-bit is propagated in the PrefixOptions when an OSPFv3 NSSA ABR originates an E-AS-External-LSA corresponding to an NSSA route as described in section 3 of RFC 3101 ([NSSA]). The N-bit is added to the Inter-Area-Prefix-TLV (Section 3.4), External-Prefix-TLV (Section 3.6), and Intra-Area-Prefix-TLV (Section 3.7). The N-bit is used as hint to identify the preferred address to reach the advertising OSPFv3 router. This would be in contrast to an Anycast Address [IPV6-ADDRESS-ARCH] which could also be a local address with the LA-bit set. It is useful for applications such as identifying the prefixes corresponding to Node Segment Identifiers (SIDs) in Segment Routing [SEGMENT-ROUTING]. There may be future applications requiring selection of a prefix associated with an OSPFv3 router.

### 3.2. Router-Link TLV

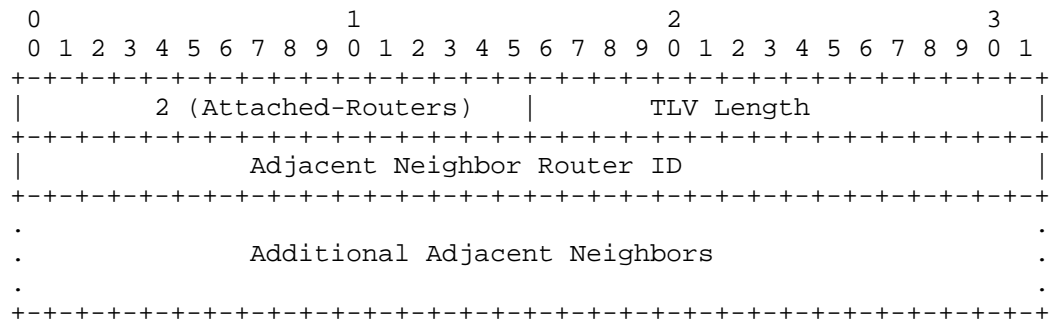
The Router-Link TLV defines a single router link and the field definitions correspond directly to links in the OSPFv3 Router-LSA, section A.4.3, [OSPFV3]. The Router-Link TLV is only applicable to the E-Router-LSA (Section 4.1). Inclusion in other Extended LSAs MUST be ignored.



Router-Link TLV

### 3.3. Attached-Routers TLV

The Attached-Routers TLV defines all the routers attached to an OSPFv3 multi-access network. The field definitions correspond directly to content of the OSPFv3 Network-LSA, section A.4.4, [OSPFV3]. The Attached-Routers TLV is only applicable to the E-Network-LSA (Section 4.2). Inclusion in other Extended LSAs MUST be ignored.



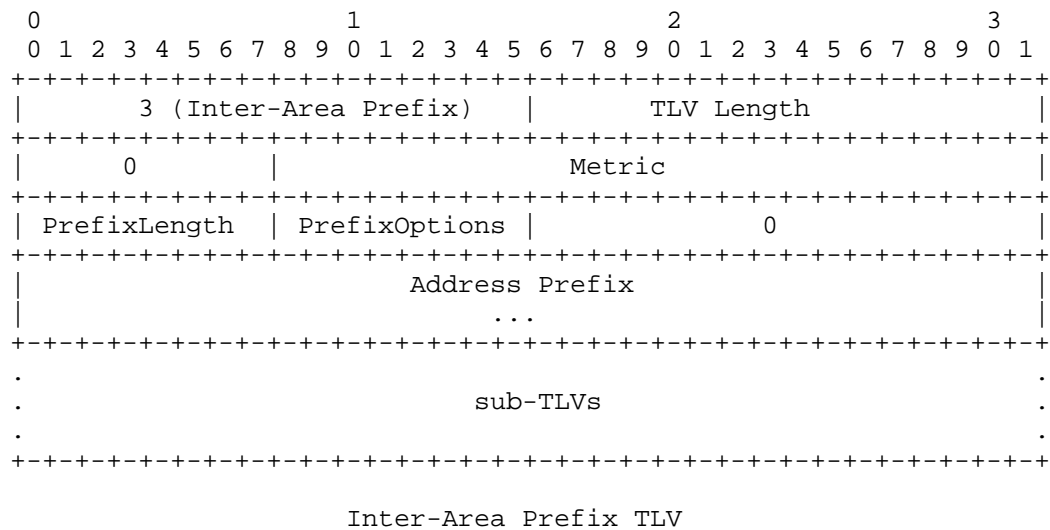
Attached-Routers TLV

There are two reasons for not having a separate TLV or sub-TLV for each adjacent neighbor. The first is to discourage using the E-Network-LSA for more than its current role of solely advertising the routers attached to a multi-access network. The router's metric as well as the attributes of individual attached routers should be

advertised in their respective E-Router-LSAs. The second reason is that there is only a single E-Network-LSA per multi-access link with the Link State ID set to the Designated Router's Interface ID and, consequently, compact encoding has been chosen to decrease the likelihood that the size of the E-Network-LSA will require IPv6 fragmentation when advertised in an OSPFv3 Link State Update packet.

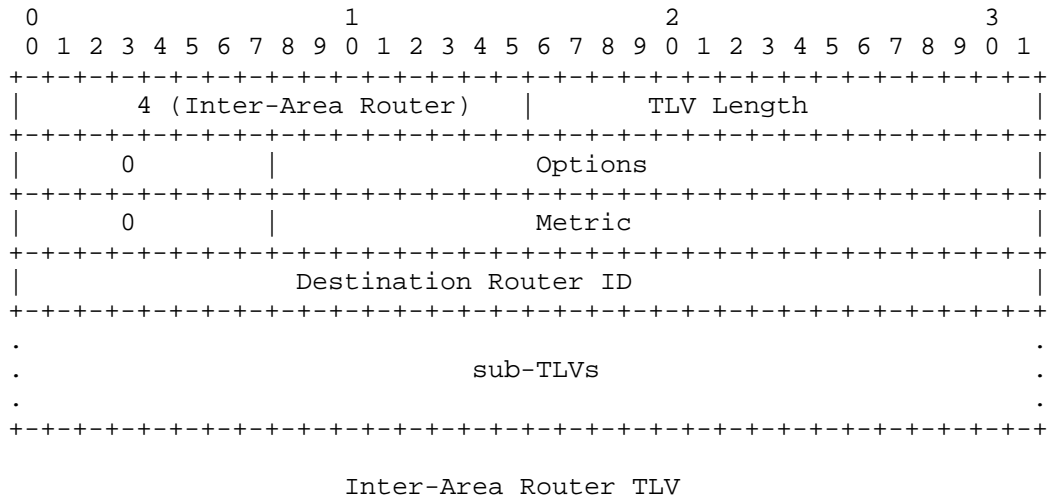
### 3.4. Inter-Area-Prefix TLV

The Inter-Area-Prefix TLV defines a single OSPFV3 inter-area prefix. The field definitions correspond directly to the content of an OSPFv3 IPv6 Prefix as defined in Section A.4.1, [OSPFV3] and an OSPFv3 Inter-Area-Prefix-LSA, as defined in section A.4.5, [OSPFV3]. Additionally, the PrefixOptions are extended as described in Section 3.1. The Inter-Area-Prefix TLV is only applicable to the E-Inter-Area-Prefix-LSA (Section 4.3). Inclusion in other Extended LSAs MUST be ignored.



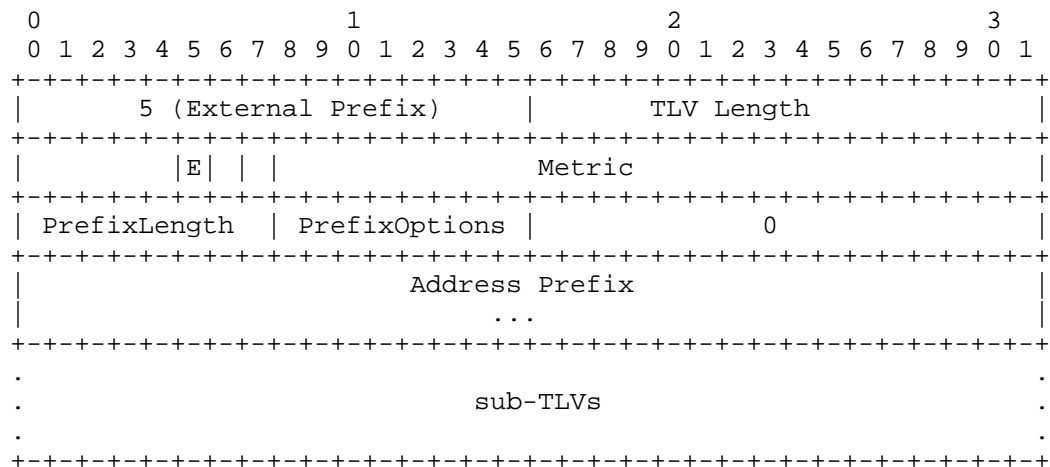
### 3.5. Inter-Area-Router TLV

The Inter-Area-Router TLV defines a single OSPFv3 Autonomous System Boundary Router (ASBR) reachable in another area. The field definitions correspond directly to the content of an OSPFv3 Inter-Area-Router-LSA, as defined in section A.4.6, [OSPFV3]. The Inter-Area-Router TLV is only applicable to the E-Inter-Area-Router-LSA (Section 4.4). Inclusion in other Extended LSAs MUST be ignored.



### 3.6. External-Prefix TLV

The External-Prefix TLV defines a single OSPFv3 external prefix. With the exception of omitted fields noted below, the field definitions correspond directly to the content of an OSPFv3 IPv6 Prefix as defined in Section A.4.1, [OSPFV3] and an OSPFv3 AS-External-LSA, as defined in section A.4.7, [OSPFV3]. The External-Prefix TLV is only applicable to the E-AS-External-LSA (Section 4.5) and the E-NSSA-LSA (Section 4.6). Additionally, the PrefixOptions are extended as described in Section 3.1. Inclusion in other Extended LSAs MUST be ignored.



External Prefix TLV

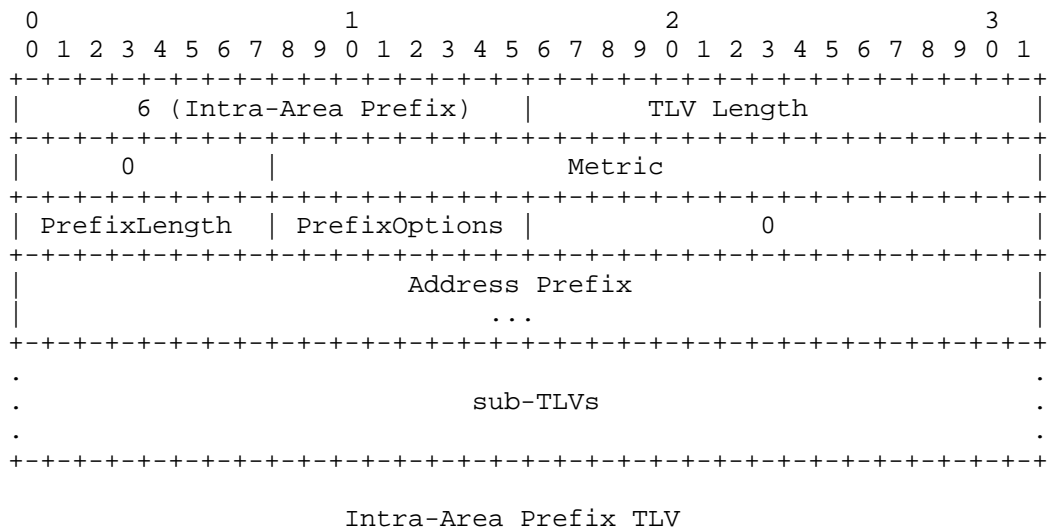
In the External-Prefix TLV, the optional IPv6/IPv4 Forwarding Address and External Route Tag are now sub-TLVs. Given the Referenced LS type and Referenced Link State ID from the AS-External-LSA have never been used or even specified, they have been omitted from the External Prefix TLV. If there were ever a requirement for a referenced LSA, it could be satisfied with a sub-TLV.

The following sub-TLVs are defined for optional inclusion in the External Prefix TLV:

- o 1 - IPv6 Forwarding Address sub-TLV (Section 3.10)
- o 2 - IPv4 Forwarding Address sub-TLV (Section 3.11)
- o 3 - Route Tag sub-TLV (Section 3.12)

### 3.7. Intra-Area-Prefix TLV

The Intra-Area-Prefix TLV defines a single OSPFv3 intra-area prefix. The field definitions correspond directly to the content of an OSPFv3 IPv6 Prefix as defined in Section A.4.1, [OSPFV3] and an OSPFv3 Link-LSA, as defined in section A.4.9, [OSPFV3]. The Intra-Area-Prefix TLV is only applicable to the E-Link-LSA (Section 4.7) and the E-Intra-Area-Prefix-LSA (Section 4.8). Additionally, the PrefixOptions are extended as described in Section 3.1. Inclusion in other Extended LSAs MUST be ignored.

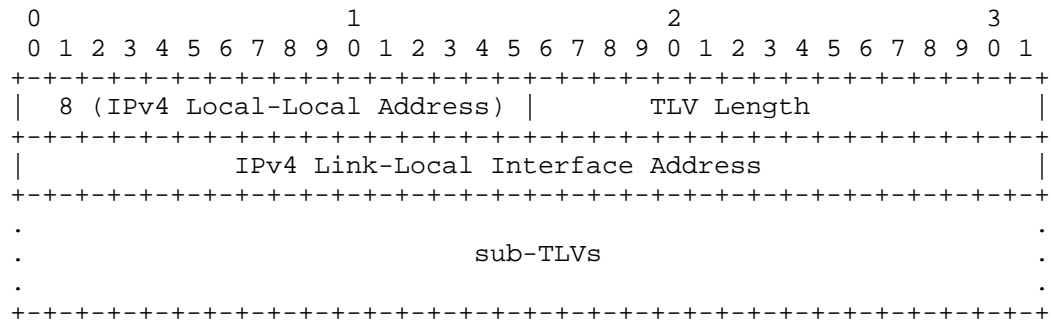






### 3.9. IPv4 Link-Local Address TLV

The IPv4 Link-Local Address TLV is to be used with IPv4 address families as defined in [OSPFV3-AF]. The IPv4 Link-Local Address TLV is only applicable to the E-Link-LSA (Section 4.7). Inclusion in other Extended LSAs MUST be ignored.

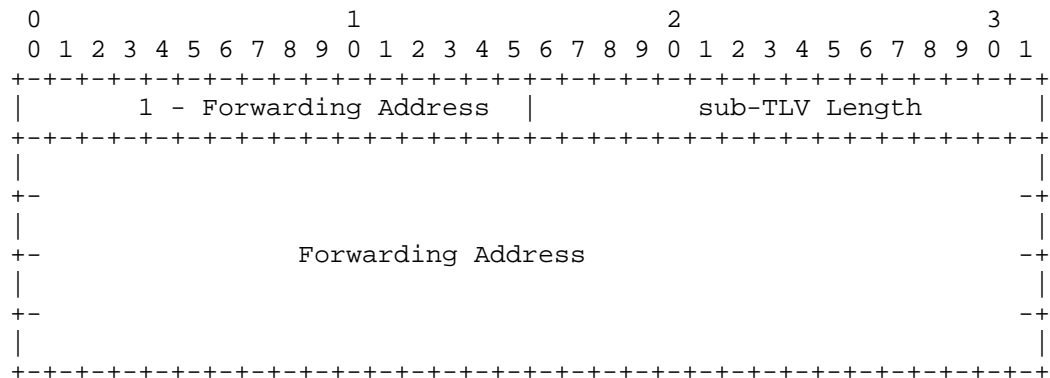


IPv4 Link-Local Address TLV

### 3.10. IPv6-Forwarding-Address Sub-TLV

The IPv6 Forwarding Address TLV has identical semantics to the optional forwarding address in section A.4.7 of [OSPFV3]. The IPv6 Forwarding Address TLV is applicable to the External-Prefix TLV (Section 3.6). Specification as a sub-TLV of other TLVs is not defined herein. The sub-TLV is optional and the first specified instance is used as the Forwarding Address as defined in [OSPFV3]. Instances subsequent to the first MUST be ignored.

The IPv6 Forwarding Address TLV is to be used with IPv6 address families as defined in [OSPFV3-AF] It MUST be ignored for other address families. The IPv6 Forwarding Address TLV length must meet minimum length (16 octets) or it will be considered malformed as described in Section 6.3.

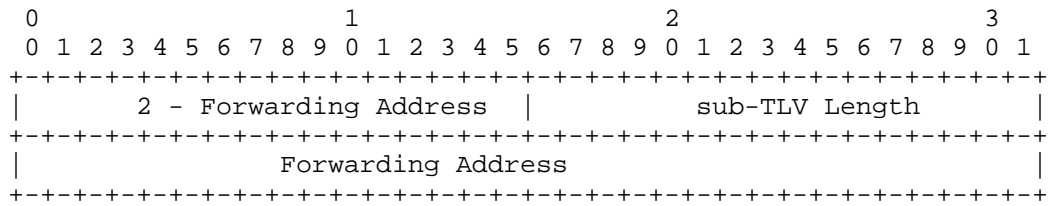


IPv6 Forwarding Address TLV

### 3.11. IPv4-Forwarding-Address Sub-TLV

The IPv4 Forwarding Address TLV has identical semantics to the optional forwarding address in section A.4.7 of [OSPFV3]. The IPv4 Forwarding Address TLV is applicable to the External-Prefix TLV (Section 3.6). Specification as a sub-TLV of other TLVs is not defined herein. The sub-TLV is optional and the first specified instance is used as the Forwarding Address as defined in [OSPFV3]. Instances subsequent to the first MUST be ignored.

The IPv4 Forwarding Address TLV is to be used with IPv4 address families as defined in [OSPFV3-AF] It MUST be ignored for other address families. The IPv4 Forwarding Address TLV length must meet minimum length (4 octets) or it will be considered malformed as described in Section 6.3.

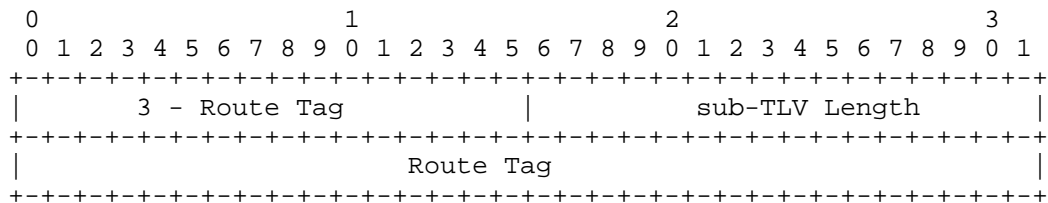


IPv4 Forwarding Address TLV

### 3.12. Route-Tag Sub-TLV

The optional Route Tag sub-TLV has identical semantics to the optional External Route Tag in section A.4.7 of [OSPFV3]. The Route Tag sub-TLV is applicable to the External-Prefix TLV (Section 3.6). Specification as a sub-TLV of other TLVs is not defined herein. The sub-TLV is optional and the first specified instance is used as the Route Tag as defined in [OSPFV3]. Instances subsequent to the first MUST be ignored.

The Route Tag TLV length must meet minimum length (4 octets) or it will be considered malformed as described in Section 6.3.



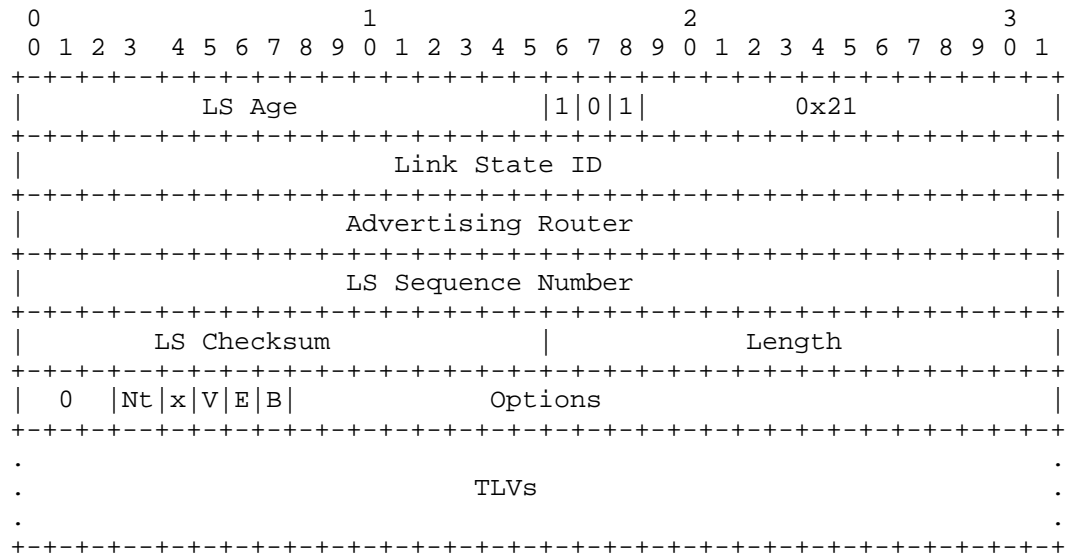
Route Tag Sub-TLV

## 4. OSPFv3 Extended LSAs

This section specifies the OSPFv3 Extended LSA formats and encoding. The Extended OSPFv3 LSAs corresponded directly to the original OSPFv3 LSAs specified in [OSPFV3].

### 4.1. OSPFv3 E-Router-LSA

The E-Router-LSA has an LS Type of 0xA021 and has the same base information content as the Router-LSA defined in section A.4.3 of [OSPFV3]. However, unlike the existing Router-LSA, it is fully extendable and represented as TLVs.

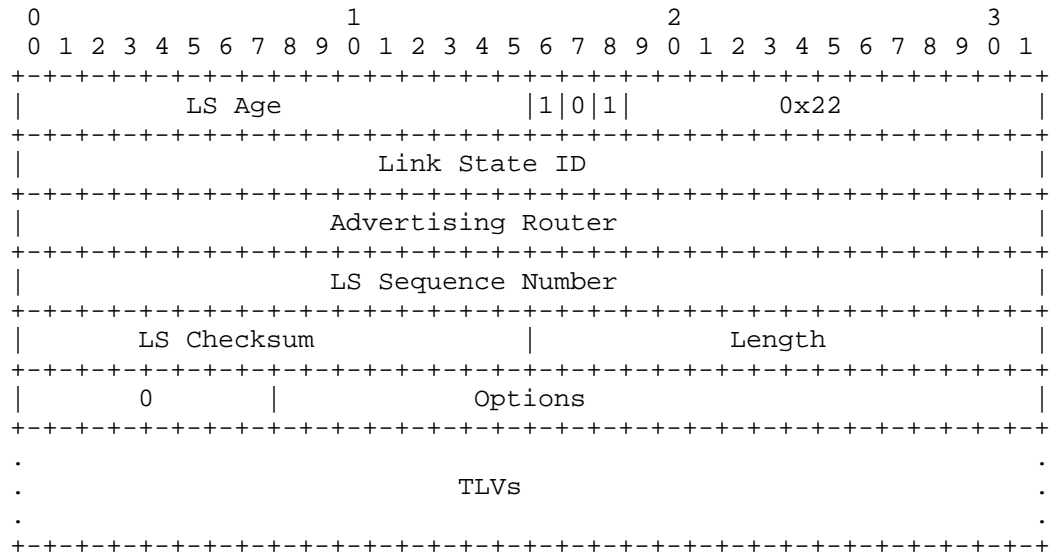


#### Extended Router-LSA

Other than having a different LS Type, all LSA Header fields are the same as defined for the Router-LSA. Initially, only the top-level Router-Link TLV Section 3.2 is applicable and an E-Router-LSA may include multiple Router-Link TLVs. Like the existing Router-LSA, the LSA length is used to determine the end of the LSA including TLVs. Depending on the implementation, it is perfectly valid for an E-Router-LSA to not contain any Router-Link TLVs. However, this would imply that the OSPFv3 router doesn't have any adjacencies in the corresponding area and is forming an adjacency or adjacencies over unnumbered link(s). Note that no E-Router-LSA stub link is advertised for an unnumbered link.

## 4.2. OSPFv3 E-Network-LSA

The E-Network-LSA has an LS Type of 0xA022 and has the same base information content as the Network-LSA defined in section A.4.4 of [OSPFV3]. However, unlike the existing Network-LSA, it is fully extendable and represented as TLVs.

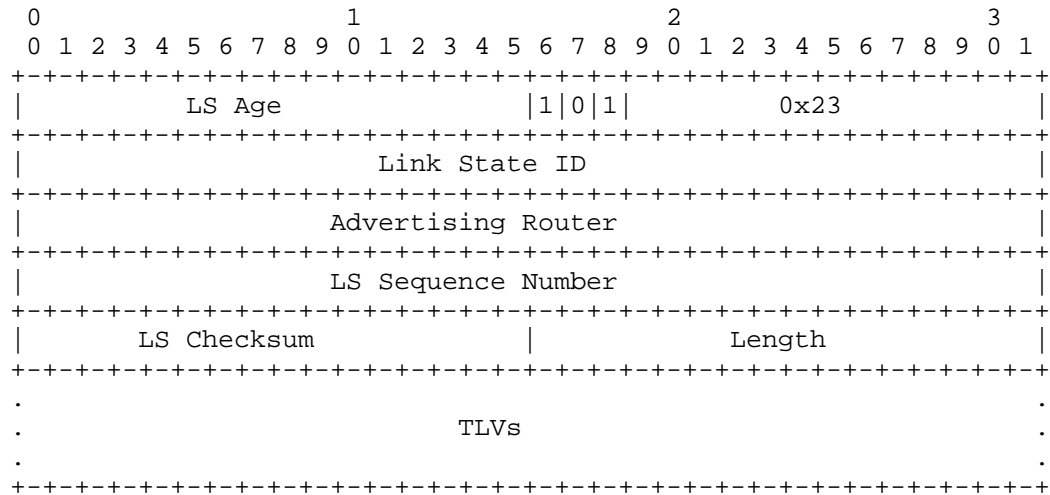


## E-Network-LSA

Other than having a different LS Type, all LSA Header fields are the same as defined for the Network-LSA. Like the existing Network-LSA, the LSA length is used to determine the end of the LSA including TLVs. Initially, only the top-level Attached-Routers TLV Section 3.3 is applicable. If the Attached-Router TLV is not included in the E-Network-LSA, it is treated as malformed as described in Section 5. Instances of the Attached-Router TLV subsequent to the first MUST be ignored.

#### 4.3. OSPFv3 E-Inter-Area-Prefix-LSA

The E-Inter-Area-Prefix-LSA has an LS Type of 0xA023 and has the same base information content as the Inter-Area-Prefix-LSA defined in section A.4.5 of [OSPFV3]. However, unlike the existing Inter-Area-Prefix-LSA, it is fully extendable and represented as TLVs.



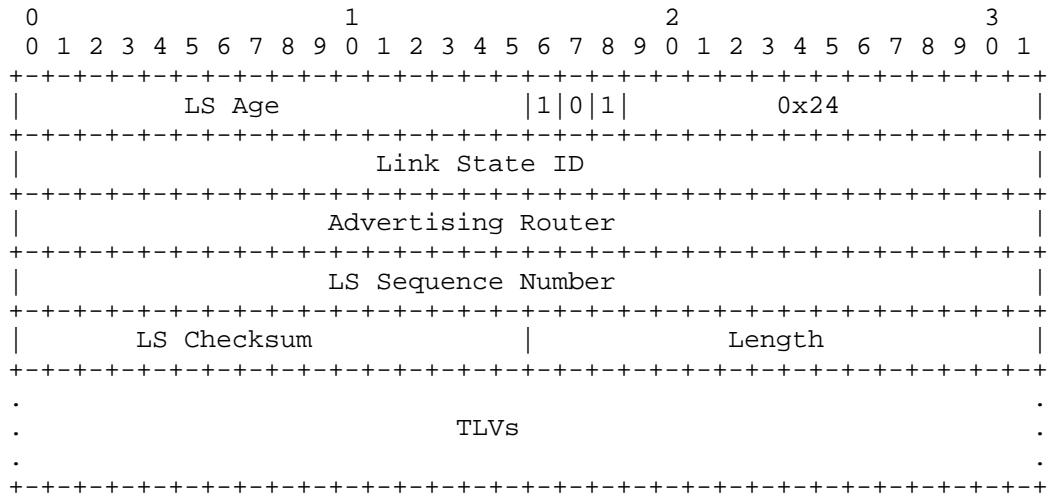
#### E-Inter-Area-Prefix-LSA

Other than having a different LS Type, all LSA Header fields are the same as defined for the Inter-Area-Prefix-LSA. In order to retain compatibility and semantics with the current OSPFv3 specification, each Inter-Area-Prefix LSA MUST contain a single Inter-Area Prefix TLV. This will facilitate migration and avoid changes to functions such as incremental SPF computation.

Like the existing Inter-Area-Prefix-LSA, the LSA length is used to determine the end of the LSA including TLV. Initially, only the top-level Inter-Area-Prefix TLV (Section 3.4) is applicable. If the Inter-Area-Prefix TLV is not included in the E-Inter-Area-Prefix-LSA, it is treated as malformed as described in Section 5. Instances of the Inter-Area-Prefix TLV subsequent to the first MUST be ignored.

#### 4.4. OSPFv3 E-Inter-Area-Router-LSA

The E-Inter-Area-Router-LSA has an LS Type of 0xA024 and has the same base information content as the Inter-Area-Router-LSA defined in section A.4.6 of [OSPFV3]. However, unlike the Inter-Area-Router-LSA, it is fully extendable and represented as TLVs.



#### E-Inter-Area-Router-LSA

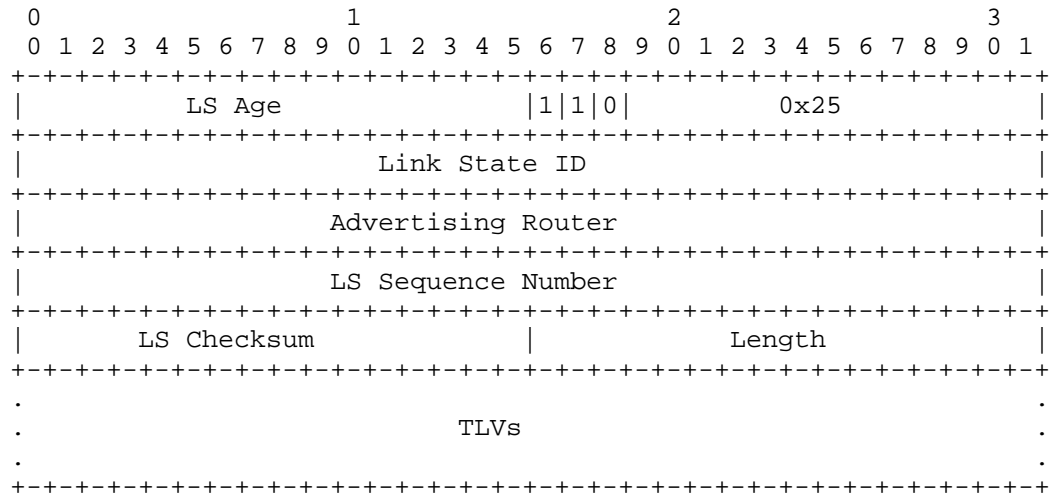
Other than having a different LS Type, all LSA Header fields are the same as defined for the Inter-Area-Router-LSA. In order to retain compatibility and semantics with the current OSPFv3 specification, each Inter-Area-Router LSA MUST contain a single Inter-Area Router TLV. This will facilitate migration and avoid changes to functions such as incremental SPF computation.

Like the existing Inter-Area-Router-LSA, the LSA length is used to determine the end of the LSA including TLV. Initially, only the top-level Inter-Area-Router TLV (Section 3.5) is applicable. If the Inter-Area-Router TLV is not included in the E-Inter-Area-Router-LSA, it is treated as malformed as described in Section 5. Instances of the Inter-Area-Router TLV subsequent to the first MUST be ignored.



## 4.5. OSPFv3 E-AS-External-LSA

The E-AS-External-LSA has an LS Type of 0xC025 and has the same base information content as the AS-External-LSA defined in section A.4.7 of [OSPFV3]. However, unlike the existing AS-External-LSA, it is fully extendable and represented as TLVs.



## E-AS-External-LSA

Other than having a different LS Type, all LSA Header fields are the same as defined for the AS-External-LSA. In order to retain compatibility and semantics with the current OSPFv3 specification, each LSA MUST contain a single External Prefix TLV. This will facilitate migration and avoid changes to OSPFv3 processes such as incremental SPF computation.

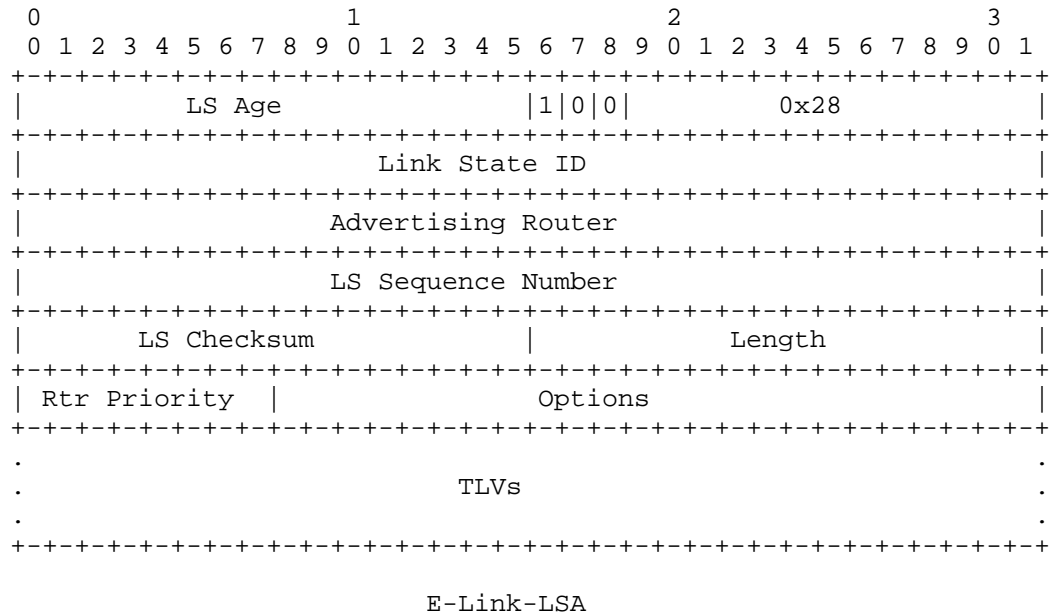
Like the existing AS-External-LSA, the LSA length is used to determine the end of the LSA including sub-TLVs. Initially, only the top-level External-Prefix TLV (Section 3.6) is applicable. If the External-Prefix TLV is not included in the E-External-AS-LSA, it is treated as malformed as described in Section 5. Instances of the External-Prefix TLV subsequent to the first MUST be ignored.

#### 4.6. OSPFv3 E-NSSA-LSA

The E-NSSA-LSA will have the same format and TLVs as the Extended AS-External-LSA Section 4.5. This is the same relationship as exists between the NSSA-LSA defined in section A.4.8 of [OSPFV3], and the AS-External-LSA. The NSSA-LSA will have type 0xA027 which implies area flooding scope. Future requirements may dictate that supported TLVs differ between the E-AS-External-LSA and the E-NSSA-LSA. However, future requirements are beyond the scope of this document.

## 4.7. OSPFv3 E-Link-LSA

The E-Link-LSA has an LS Type of 0x8028 and will have the same base information content as the Link-LSA defined in section A.4.9 of [OSPFV3]. However, unlike the existing Link-LSA, it is extendable and represented as TLVs.



Other than having a different LS Type, all LSA Header fields are the same as defined for the Link-LSA.

Only the Intra-Area-Prefix TLV (Section 3.7), IPv6 Link-Local Address TLV (Section 3.8), and IPv4 Link-Local Address TLV (Section 3.9) are applicable to the E-Link-LSA. Like the Link-LSA, the E-Link-LSA affords advertisement of multiple intra-area prefixes. Hence, multiple Intra-Area Prefix TLVs (Section 3.7) may be specified and the LSA length defines the end of the LSA including all TLVs.

A single instance of the IPv6 Link-Local Address TLV (Section 3.8) SHOULD be included in the E-Link-LSA. Instances following the first MUST be ignored. For IPv4 address families as defined in [OSPFV3-AF], this TLV MUST be ignored.

Similarly, only a single instance of the IPv4 Link-Local Address TLV (Section 3.9) SHOULD be included in the E-Link-LSA. Instances

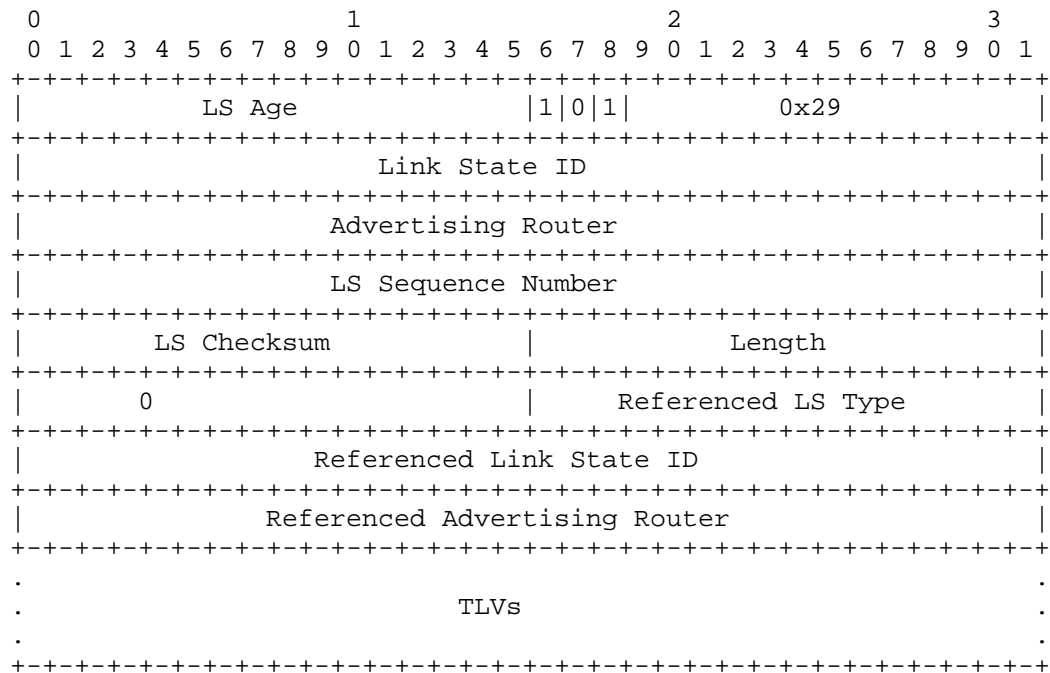
following the first MUST be ignored. For OSPFv3 IPv6 address families as defined in [OSPFV3-AF], this TLV SHOULD be ignored.

If the IPv4/IPv6 Link-Local Address TLV corresponding to the OSPFv3 Address Family is not included in the E-Link-LSA, it is treated as malformed as described in Section 5.

Future specifications may support advertisement of routing and topology information for multiple address families. However, this is beyond the scope of this document.

#### 4.8. OSPFv3 E-Intra-Area-Prefix-LSA

The E-Intra-Area-Prefix-LSA has an LS Type of 0xA029 and has the same base information content as the Intra-Area-Prefix-LSA defined in section A.4.10 of [OSPFV3] except for the Referenced LS Type. However, unlike the Intra-Area-Prefix-LSA, it is fully extendable and represented as TLVs. The Referenced LS Type MUST be either an E-Router-LSA (0xA021) or an E-Network-LSA (0xA022).



#### E-Intra-Area-Prefix-LSA

Other than having a different LS Type, all LSA Header fields are the same as defined for the Intra-Area-Prefix-LSA.

Like the Intra-Area-Prefix-LSA, the E-Intra-Area-Link-LSA affords advertisement of multiple intra-area prefixes. Hence, multiple Intra-Area Prefix TLVs may be specified and the LSA length defines the end of the LSA including all TLVs.

## 5. Malformed OSPFv3 Extended LSA Handling

Extended LSAs that have inconsistent length or other encoding errors, as described herein, MUST NOT be installed in the Link State Database, acknowledged, or flooded. Reception of malformed LSAs SHOULD be counted and/or logged for examination by the administrator of the OSPFv3 Routing Domain. Note that for the purposes of length validation, a TLV or Sub-TLV should not be considered invalid unless the length exceeds the length of the LSA or does not meet the minimum length requirements. This allows for Sub-TLVs to be added as described in Section 6.3.

Additionally, an LSA MUST be considered malformed if it does not include all of the required TLVs and Sub-TLVs.

## 6. LSA Extension Backward Compatibility

In the context of this document, backward compatibility is solely related to the capability of an OSPFv3 router to receive, process, and originate the TLV-based LSAs defined herein. Unrecognized TLVs and sub-TLVs are ignored. Backward compatibility for future OSPFv3 extensions utilizing the TLV-based LSAs is out of scope and must be covered in the documents describing those extensions. Both full and, if applicable, partial deployment SHOULD be specified for future TLV-based OSPFv3 LSA extensions.

### 6.1. Full Extended LSA Migration

If ExtendedLSASupport is enabled Appendix A, OSPFv3 Extended LSAs will be originated and used for the SPF computation. Individual OSPF Areas can be migrated separately with the Legacy AS-External LSAs being originated and used for the SPF computation. This is accomplished by enabled AreaExtendedLSASupport Appendix B.

An OSPFv3 routing domain or area may be non-disruptively migrated using separate OSPFv3 instances for the extended LSAs. Initially, the OSPFv3 instances with ExtendedLSASupport will have a lower preference, i.e., higher administrative distance, than the OSPFv3 instances originating and using the Legacy LSAs. Once the routing domain or area is fully migrated and the OSPFv3 Routing Information Bases (RIB) have been verified, the OSPFv3 instances using the extended LSAs can be given preference. When this has been completed and the routing within the OSPF routing domain or area has been verified, the original OSPFv3 instance using Legacy LSAs can be removed.

## 6.2. Extended LSA Sparse-Mode Backward Compatibility

In this mode, OSPFv3 will use the Legacy LSAs for the SPF computation and will only originate extended LSAs when LSA origination is required in support of additional functionality. Furthermore, those extended LSAs will only include the top-level TLVs (e.g., Router-Link TLVs or Inter-Area TLVs) which require further specification for that new functionality. However, if a top-level TLV is advertised, it MUST include required Sub-TLVs or it will be considered malformed as described in Section 5. Hence, this mode of compatibility is known as "sparse-mode". The advantage of sparse-mode is that functionality utilizing the OSPFv3 extended LSAs can be added to an existing OSPFv3 routing domain without the requirement for migration. In essence, this compatibility mode is very much like the approach taken for OSPFv2 [OSPF-PREFIX-LINK]. As with all the compatibility modes, backward compatibility for the functions utilizing the extended LSAs must be described in the IETF documents describing those functions.

## 6.3. LSA TLV Processing Backward Compatibility

This section defines the general rules for processing LSA TLVs. To ensure compatibility of future TLV-based LSA extensions, all implementations MUST adhere to these rules:

1. Unrecognized TLVs and sub-TLVs are ignored when parsing or processing Extended-LSAs.
2. Whether or not partial deployment of a given TLV is supported MUST be specified.
3. If partial deployment is not supported, mechanisms to ensure the corresponding feature are not deployed MUST be specified in the document defining the new TLV or sub-TLV.
4. If partial deployment is supported, backward compatibility and partial deployment MUST be specified in the document defining the new TLV or sub-TLV.
5. If a TLV or Sub-TLV is recognized but the length is less than the minimum, then the LSA should be considered malformed and it SHOULD NOT be acknowledged. Additionally, the occurrence SHOULD be logged with enough information to identify the LSA by type, originator, and sequence number and the TLV or Sub-TLV in error. Ideally, the log entry would include the hexadecimal or binary representation of the LSA including the malformed TLV or Sub-TLV.
6. Documents specifying future TLVs or Sub-TLVs MUST specify the requirements for usage of those TLVs or Sub-TLVs.

7. Future TLV or Sub-TLVs must be optional. However, there may be requirements for Sub-TLVs if an optional TLV is specified.

## 7. Security Considerations

In general, extendible OSPFv3 LSAs are subject to the same security concerns as those described in RFC 5340 [OSPFV3]. Additionally, implementations must assure that malformed TLV and sub-TLV permutations do not result in errors that cause hard OSPFv3 failures.

If there were ever a requirement to digitally sign OSPFv3 LSAs as described for OSPFv2 LSAs in RFC 2154 [OSPF-DIGITAL-SIGNATURE], the mechanisms described herein would greatly simplify the extension.

## 8. IANA Considerations

This specification defines nine OSPFv3 Extended LSA types as described in Section 2. These are added to the existing OSPFv3 LSA Function Codes registry.

The specification defines a new code point for the N-bit in the OSPFv3 Prefix-Options registry. The value 0x20 is suggested.

This specification also creates two registries OSPFv3 Extended-LSAs TLVs and sub-TLVs. The TLV and sub-TLV code-points in these registries are common to all Extended-LSAs and their respective definitions must define where they are applicable.

### 8.1. OSPFv3 Extended-LSA TLV Registry

The OSPFv3 Extended-LSA TLV registry defines top-level TLVs for Extended-LSAs and should be placed in the existing OSPFv3 IANA registry.

Nine values are allocated by this specification:

- o 0 - Reserved
- o 1 - Router-Link TLV
- o 2 - Attached-Routers TLV
- o 3 - Inter-Area Prefix TLV
- o 4 - Inter-Area Router TLV
- o 5 - External Prefix TLV



- o 6 - Intra-Area Prefix TLV
- o 7 - IPv6 Link-Local Address TLV
- o 8 - IPv4 Link-Local Address TLV

Types in the range 9-32767 are allocated via IETF Consensus or IESG Approval.

Types in the range 32768-33023 are for experimental use; these will not be registered with IANA, and MUST NOT be mentioned by RFCs.

Types in the range 33024-45055 are to be assigned on a First-Come-First-Serve (FCFS) basis.

Types in the range 45056-65535 are not to be assigned at this time. Before any assignments can be made in the 33024-65535 range, there MUST be an IETF specification that specifies IANA Considerations that covers the range being assigned.

## 8.2. OSPFv3 Extended-LSA sub-TLV Registry

The OSPFv3 Extended-LSA sub-TLV registry defines sub-TLVs at any level of nesting for Extended-LSAs and should be placed in the existing OSPFv3 IANA registry.

Four values are allocated by this specification:

- o 0 - Reserved
- o 1 - IPv6 Forwarding Address sub-TLV
- o 2 - IPv4 Forwarding Address sub-TLV
- o 3 - Route Tag sub-TLV

Types in the range 4-32767 are allocated via IETF Consensus or IESG Approval.

Types in the range 32768-33023 are for experimental use; these will not be registered with IANA, and MUST NOT be mentioned by RFCs.

Types in the range 33024-45055 are to be assigned on a First-Come-First-Serve (FCFS) basis.

Types in the range 45056-65535 are not to be assigned at this time. Before any assignments can be made in the 33024-65535 range, there

MUST be an IETF specification that specifies IANA Considerations that covers the range being assigned.

## 9. Contributors

### Contributors' Addresses

Sina Mirtorabi  
Cisco Systems  
170 Tasman Drive  
San Jose, CA 95134  
USA  
Email: sina@cisco.com

## 10. References

### 10.1. Normative References

- [NSSA] Murphy, P., "The OSPF Not-So-Stubby Area (NSSA) Option", RFC 3101, January 2003.
- [OSPFV3] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [OSPFV3-AF]  
Lindem, A., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC 5838, April 2010.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", RFC 8174, May 2017.
- [TE] Katz, D., Yeung, D., and K. Kompella, "Traffic Engineering Extensions to OSPF", RFC 3630, September 2003.

### 10.2. Informative References

- [IPV6-ADDRESS-ARCH]  
Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

## [MT-OSPFV3]

Mirtorabi, S. and A. Roy, "Multi-topology routing in OSPFv3 (MT-OSPFV3)", draft-ietf-ospf-mt-ospfv3-04.txt (work in progress), January 2008.

## [OSPF-DIGITAL-SIGNATURE]

Murphy, S., Badger, M., and B. Wellington, "OSPF with Digital Signatures", RFC 2154, June 1997.

## [OSPF-PREFIX-LINK]

Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPF Prefix/Link Attributes", RFC 7684, December 2015.

## [SEGMENT-ROUTING]

Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPFv3 Extensions for Segment Routing", draft-ietf-ospf-ospfv3-segment-routing-extensions-10.txt (work in progress), July 2016.

## Appendix A. Appendix A - Global Configuration Parameters

The global configurable parameter `ExtendedLSASupport` is added to the OSPFv3 protocol. If `ExtendedLSASupport` is enabled, the OSPFv3 Router will originate OSPFv3 Extended LSAs and use the LSAs for the SPF computation. If `ExtendedLSASupport` is not enabled, a subset of OSPFv3 Extended LSAs may still be originated and used for other functions as described in Section 6.2.

## Appendix B. Appendix B - Area Configuration Parameters

The area configurable parameter `AreaExtendedLSASupport` is added to the OSPFv3 protocol. If `AreaExtendedLSASupport` is enabled, the OSPFv3 Router will originate link and area OSPFv3 Extended LSAs and use the LSAs for the SPF computation. Legacy AS-Scoped LSAs will still be originated and used for the AS External LSA computation. If `AreaExtendedLSASupport` is not enabled a subset of OSPFv3 link and area Extended LSAs may still be originated and used for other functions as described in Section 6.2.

For regular areas, i.e., areas where AS scoped LSAs are flooded, disabling `AreaExtendedLSASupport` for a regular OSPFv3 area (not a Stub or NSSA area) when `ExtendedLSASupport` is enabled is contradictory and SHOULD be prohibited by the implementation.

## Appendix C. Acknowledgments

OSPFv3 TLV-based LSAs were first proposed in "Multi-topology routing in OSPFv3 (MT-OSPFv3)" [MT-OSPFV3].

Thanks for Peter Psenak for significant contributions to the backward compatibility mechanisms.

Thanks go to Michael Barnes, Mike Dubrovsky, Anton Smirnov, and Tony Przygienda for review of the draft versions and discussions of backward compatibility.

Thanks to Alan Davey for review and comments including the suggestion to separate the extended LSA TLV definitions from the extended LSAs definitions.

Thanks to David Lamparter for review and suggestions on backward compatibility.

Thanks to Karsten Thomann, Chris Bowers, Meng Zhang, and Nagendra Kumar for review and editorial comments.

Thanks to Alia Atlas for substantive Routing Area Director (AD) comments prior to IETF last call.

Thanks to Alvaro Retana and Suresh Krishna for substantive comments during IESG Review.

Thanks to Mehmet Ersue for OPS Directorate review.

The RFC text was produced using Marshall Rose's xml2rfc tool.

## Authors' Addresses

Acee Lindem  
Cisco Systems  
301 Midenhall Way  
Cary, NC 27513  
USA

Email: [acee@cisco.com](mailto:acee@cisco.com)

Abhay Roy  
Cisco Systems  
170 Tasman Drive  
San Jose, CA 95134  
USA

Email: akr@cisco.com

Dirk Goethals  
Nokia  
Copernicuslaan 50  
Antwerp 2018  
Belgium

Email: dirk.goethals@nokia.com

Veerendranatha Reddy Vallem  
Bangalore  
India

Email: vallem.veerendra@gmail.com

Fred Baker  
Santa Barbara, California 93117  
USA

Email: FredBaker.IETF@gmail.com

OSPF Working Group  
Internet-Draft  
Updates: 2328, 5709  
(if approved)  
Intended status: Standards Track  
Expires: May 11, 2015

M. Bhatia  
Ionos Networks  
S. Hartman  
Painless Security  
D. Zhang  
Huawei Technologies co., LTD.  
A. Lindem, Ed.  
Cisco  
November 7, 2014

Security Extension for OSPFv2 when using Manual Key Management  
draft-ietf-ospf-security-extension-manual-keying-11

Abstract

The current OSPFv2 cryptographic authentication mechanism as defined in RFC 2328 and RFC 5709 is vulnerable to both inter-session and intra-session replay attacks when using manual keying. Additionally, the existing cryptographic authentication mechanism does not cover the IP header. This omission can be exploited to carry out various types of attacks.

This document defines changes to the authentication sequence number mechanism that will protect OSPFv2 from both inter-session and intra-session replay attacks when using manual keys for securing OSPFv2 protocol packets. Additionally, we also describe some changes in the cryptographic hash computation that will eliminate attacks resulting from OSPFv2 not protecting the IP header.

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 <http://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 May 11, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	3
1.1. Requirements Section . . . . .	3
1.2. Acknowledgments . . . . .	4
2. Replay Protection using Extended Sequence Numbers . . . . .	4
3. OSPF Packet Extensions . . . . .	5
4. OSPF Packet Key Selection . . . . .	6
4.1. Key Selection for Unicast OSPF Packet Transmission . . . . .	7
4.2. Key Selection for Multicast OSPF Packet Transmission . . . . .	8
4.3. Key Selection for OSPF Packet Reception . . . . .	8
5. Securing the IP header . . . . .	9
6. Mitigating Cross-Protocol Attacks . . . . .	9
7. Backward Compatibility . . . . .	10
8. Security Considerations . . . . .	10
9. IANA Considerations . . . . .	11
10. References . . . . .	12
10.1. Normative References . . . . .	12
10.2. Informative References . . . . .	12
Authors' Addresses . . . . .	13

## 1. Introduction

The OSPFv2 cryptographic authentication mechanism as described in [RFC2328] uses per-packet sequence numbers to provide protection against replay attacks. The sequence numbers increase monotonically so that attempts to replay stale packets can be thwarted. The sequence number values are maintained as a part of neighbor adjacency state. Therefore, if an adjacency is taken down, the associated sequence numbers get reinitialized and neighbor adjacency formation starts over again. Additionally, the cryptographic authentication mechanism does not specify how to deal with the rollover of a sequence number when its value wraps. These omissions can be exploited by attackers to implement various replay attacks ([RFC6039]). In order to address these issues, we define extensions to the authentication sequence number mechanism.

The cryptographic authentication as described in [RFC2328] and later updated in [RFC5709] does not include the IP header. This omission can be exploited to launch several attacks as the source address in the IP header is not protected. The OSPF specification, for broadcast and NBMA (Non-Broadcast Multi-Access Networks), requires implementations to use the source address in the IP header to determine the neighbor from which the packet was received. Changing the IP source address of a packet to a conflicting IP address can be exploited to produce a number of denial of service attacks [RFC6039]. If the packet is interpreted as coming from a different neighbor, the received sequence number state for that neighbor may be incorrectly updated. This attack may disrupt communication with a legitimate neighbor. Hello packets may be reflected to cause a neighbor to appear to have one-way communication. Additionally, Database Description packets may be reflected in cases where the per-packet sequence numbers are sufficiently divergent in order to disrupt an adjacency [RFC6863]. This is referred to as the IP layer issue in [RFC6862].

[RFC2328] states that implementations MUST offer keyed MD5 authentication. It is likely that this will be deprecated in favor of the stronger algorithms described in [RFC5709] and required in [RFC6094].

This document defines a few simple changes to the cryptographic authentication mechanism, as currently described in [RFC5709], to prevent such IP layer attacks.

### 1.1. Requirements Section

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this



document are to be interpreted as described in RFC2119 [RFC2119].

When used in lowercase, these words convey their typical use in common language, and are not to be interpreted as described in RFC2119 [RFC2119].

## 1.2. Acknowledgments

Thanks to Ran Atkinson for help in the analysis of RFC 6506 errata leading to clarifications in this document.

Thanks to Gabi Nakibly for pointing out a possible attack on p2p links.

Thanks to Suresh Krishnan for comments made during the Gen-Art review. In particular, thanks for pointing out an ambiguity in the initialization of Apad.

Thanks to Shaun Cooley for the security directorate review.

Thanks to Adrian Farrel for comments during the IESG last call.

## 2. Replay Protection using Extended Sequence Numbers

In order to provide replay protection against both inter-session and intra-session replay attacks, the OSPFv2 sequence number is expanded to 64-bits with the least significant 32-bit value containing a strictly increasing sequence number and the most significant 32-bit value containing the boot count. OSPFv2 implementations are required to retain the boot count in non-volatile storage for the deployment life the OSPF router. The requirement to preserve the boot count is also placed on SNMP agents by the SNMPv3 security architecture (refer to `snmpEngineBoots` in section 2.2 of [RFC2574]).

Since there is no room in the OSPFv2 packet for a 64-bit sequence number, it will occupy the 8 octets following the OSPFv2 packet and MUST be included when calculating the OSPFv2 packet digest. These additional 8 octets are not included in the OSPFv2 packet header length but are included in the OSPFv2 header Authentication Data length and the IPv4 packet header length.

The lower order 32-bit sequence number MUST be incremented for every OSPF packet sent by the OSPF router. Upon reception, the sequence number MUST be greater than the sequence number in the last OSPF packet of that type accepted from the sending OSPF neighbor. Otherwise, the OSPF packet is considered a replayed packet and dropped. OSPF packets of different types may arrive out of order if

they are prioritized as recommended in [RFC4222].

OSPF routers implementing this specification MUST use available mechanisms to preserve the sequence number's strictly increasing property for the deployed life of the OSPFv2 router (including cold restarts). This is achieved by maintaining a boot count in non-volatile storage and incrementing it each time the OSPF router loses its prior sequence number state. The SNMPv3 `snmpEngineBoots` variable [RFC2574] MAY be used for this purpose. However, maintaining a separate boot count solely for OSPF sequence numbers has the advantage of decoupling SNMP reinitialization and OSPF reinitialization. Also, in the rare event that the lower order 32-bit sequence number wraps, the boot count can be incremented to preserve the strictly increasing property of the aggregate sequence number. Hence, a separate OSPF boot count is RECOMMENDED.

### 3. OSPF Packet Extensions

The OSPF packet header includes an authentication type field, and 64-bits of data for use by the appropriate authentication scheme (determined by the type field). Authentication types 0, 1 and 2 are defined [RFC2328]. This section defines Authentication type TBD (3 is recommended).

When using this authentication scheme, the 64-bit Authentication field in the OSPF packet header as defined in section D.3 of [RFC2328] and [RFC6549] is changed as shown below. The sequence number is removed and the Key ID is extended to 32 bits and moved to the former position of the sequence number.

Additionally, the 64-bit sequence number is moved to the first 64-bits following the OSPFv2 packet and is protected by the authentication digest. These additional 64 bits or 8 octets are included in the IP header length but not the OSPF header packet length.

Finally, the 0 field at the start of the OSPFv2 header authentication is extended from 16 bits to 24 bits.

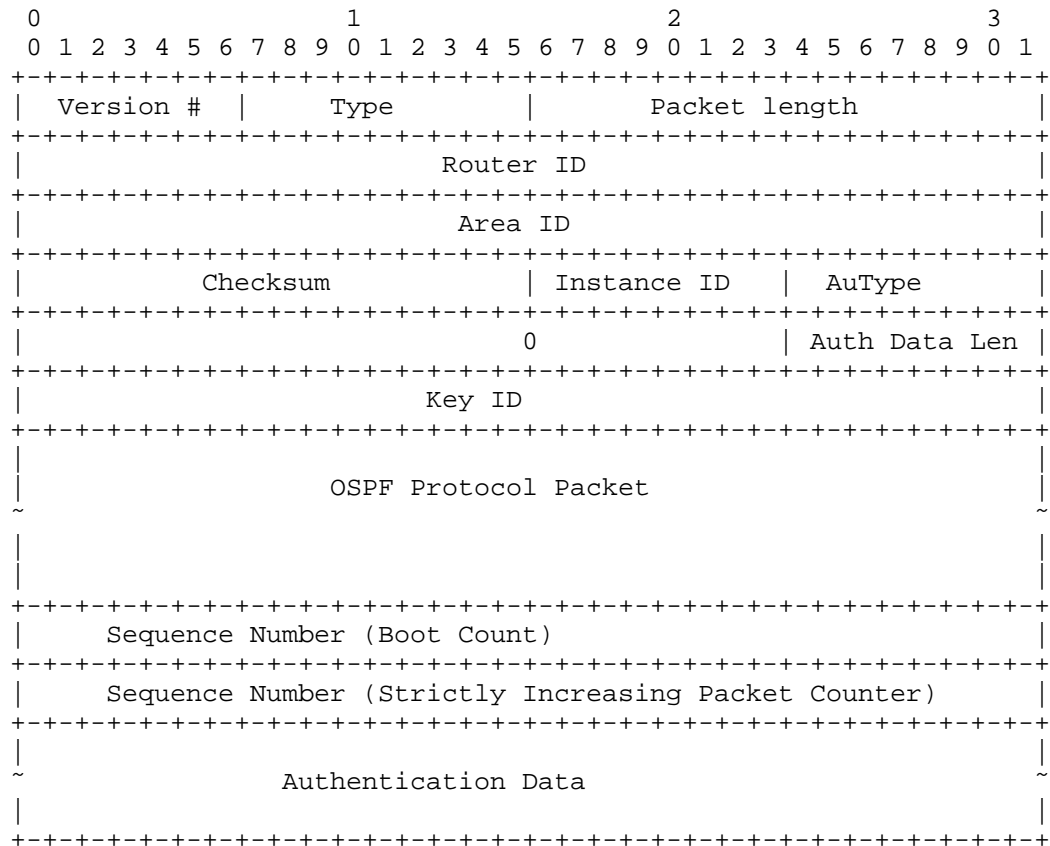


Figure 1 - Extended Sequence Number Packet Extensions

#### 4. OSPF Packet Key Selection

This section describes how this security solution selects long-lived keys from key tables. [RFC7210]. In this context, we are selecting the key and corresponding Security Association (SA) as defined in section 3.2 of [RFC5709]. Generally, a key used for OSPFv2 packet authentication should satisfy the following requirements:

- o For packet transmission, the key validity interval as defined by SendLifetimeStart and SendLifetimeEnd must include the current time.
- o For packet reception, the key validity interval as defined by AcceptLifetimeStart and AcceptLifetimeEnd must include the current

time.

- o The key must be valid for the desired security algorithm.

In the remainder of this section, additional requirements for keys are enumerated for different scenarios.

#### 4.1. Key Selection for Unicast OSPF Packet Transmission

Assume that a router R1 tries to send a unicast OSPF packet from its interface I1 to the interface I2 of a remote router R2 using security protocol P via interface I at time T. First, consider the circumstances where R1 and R2 are not connected with a virtual link. R1 then needs to select a long long-lived symmetric key from its key table. Because the key should be shared by both R1 and R2 to protect the communication between I1 and I2, the key should satisfy the following requirements:

- o The Peers field is unused. OSPF authentication is interface based.
- o The Interfaces field includes the local IP address of the interface for numbered interfaces or the MIB-II [RFC1213] ifIndex for unnumbered interfaces.
- o The Direction field is either "out" or "both".
- o If multiple keys match the Interfaces field, the key with the most recent SendLifetimeStart time will be selected. This will facilitate graceful key rollover.
- o The Key ID field in the OSPFv2 header (refer to figure 1) will be set to the selected key's LocalKeyName.

When R1 and R2 are connected to a virtual link, the Interfaces field must identify the virtual endpoint rather than the virtual link. Since there may be virtual links to the same router, the transit area ID must be part of the identifier. Hence, the key should satisfy the following requirements:

- o The Peers field is unused. OSPF authentication is interface based.
- o The Interfaces field includes both the virtual endpoint's OSPF router ID and the transit area ID for the virtual link.
- o The Direction field is either "out" or "both".

- o If multiple keys match the Interfaces field, the key with the most recent SendLifetimeStart time will be selected. This will facilitate graceful key rollover.
- o The Key ID field in the OSPFv2 header (refer to figure 1) will be set to the selected key's LocalKeyName.

#### 4.2. Key Selection for Multicast OSPF Packet Transmission

If a router R1 sends an OSPF packet from its interface I1 to a multicast address (i.e., AllSPFRouters or AllDRouters), it needs to select a key according to the following requirements:

- o The Peers field is unused. OSPF authentication is interface based.
- o The Interfaces field includes the local IP address of the interface for numbered interfaces or the MIB-II [RFC1213] ifIndex for unnumbered interfaces.
- o The Direction field is either "out" or "both".
- o If multiple keys match the Interfaces field, the key with the most recent SendLifetimeStart time will be selected. This will facilitate graceful key rollover.
- o The Key ID field in the OSPFv2 header (refer to figure 1) will be set to the selected key's LocalKeyName.

#### 4.3. Key Selection for OSPF Packet Reception

When Cryptographic Authentication is used, the ID of the authentication key is included in the authentication field of the OSPF packet header. Using this Key ID, it is straight forward for a receiver to locate the corresponding key. The simple requirements are:

- o The interface on which the key was received is associated with the key's interface.
- o The Key ID obtained from the OSPFv2 packet header corresponds to the neighbor's PeerKeyName. Since OSPFv2 keys are symmetric, the LocalKeyName and PeerKeyName for OSPFv2 keys will be identical. Hence, the Key ID will be used to select the correct local key.
- o The Direction field is either "in" or "both".

## 5. Securing the IP header

This document updates the definition of the Apad constant, as it is defined in [RFC5709], to include the IP source address from the IP header of the OSPFv2 protocol packet. The overall cryptographic authentication process defined in [RFC5709] remains unchanged. To reduce the potential for confusion, this section minimizes the repetition of text from RFC 5709 [RFC5709]. The changes are:

RFC 5709, Section 3.3, describes how the cryptographic authentication must be computed. In RFC 5709, the First-Hash includes the OSPF packet and Authentication Trailer. With this specification, the 64-bit sequence number will be included in the First-Hash along with the Authentication Trailer and OSPF packet.

RFC 5709, Section 3.3 also requires the OSPFv2 packet's Authentication Trailer (which is the appendage described in RFC 2328, Section D.4.3, Page 233, items (6)(a) and (6)(d)) to be filled with the value Apad. Apad is a hexadecimal constant with the value 0x878FE1F3 repeated (L/4) times, where L is the length of the hash being used and is measured in octets rather than bits.

OSPF routers sending OSPF packets must initialize the first 4 octets of Apad to the value of the IP source address that would be used when sending the OSPFv2 packet. The remainder of Apad will contain the value 0x878FE1F3 repeated (L - 4)/4 times, where L is the length of the hash, measured in octets. The basic idea is to incorporate the IP source address from the IP header in the cryptographic authentication computation so that any change of IP source address in a replayed packet can be detected.

When an OSPF packet is received, implementations MUST initialize Apad as the IP source address from the IP Header of the incoming OSPFv2 packet, repeated L/4 times, instead of the constant that's currently defined in [RFC5709]. Besides changing the value of Apad, this document does not introduce any other changes to the authentication mechanism described in [RFC5709]. This would prevent all attacks where a rogue OSPF router changes the IP source address of an OSPFv2 packet and replays it on the same multi-access interface or another interface since the IP source address is now included in the cryptographic hash computation and modification would result in the OSPFv2 packet being dropped due to an authentication failure.

## 6. Mitigating Cross-Protocol Attacks

In order to prevent cross-protocol replay attacks for protocols sharing common keys, the two octet OSPFv2 Cryptographic Protocol ID

is appended to the authentication key prior to use. Refer to IANA Considerations (Section 9).

[RFC5709], Section 3.3 describes the mechanism to prepare the key used in the hash computation. This document updates the sub section "PREPARATION OF KEY" as follows:

The OSPFv2 Cryptographic Protocol ID is appended to the Authentication Key (K) yielding a Protocol-Specific Authentication Key (Ks). In this application, Ko is always L octets long. While [RFC2104] supports a key that is up to B octets long, this application uses L as the Ks length consistent with [RFC4822], [RFC5310], and [RFC5709]. According to [FIPS-198], Section 3, keys greater than L octets do not significantly increase the function strength. Ks is computed as follows:

If the Protocol-Specific Authentication Key (Ks) is L octets long, then Ko is equal to Ks. If the Protocol-Specific Authentication Key (Ks) is more than L octets long, then Ko is set to H(Ks). If the Protocol-Specific Authentication Key (Ks) is less than L octets long, then Ko is set to the Protocol-Specific Authentication Key (Ks) with zeros appended to the end of the Protocol-Specific Authentication Key (Ks) such that Ko is L octets long.

Once the cryptographic key (Ko) used with the hash algorithm is derived the rest of the authentication mechanism described in [RFC5709] remains unchanged other than one detail that was unspecified. When XORing Ko and Ipad or Opad, Ko MUST be padded with zeros to the length of Ipad or Opad. It is expected that RFC 5709 [RFC5709] implementations perform this padding implicitly.

## 7. Backward Compatibility

This security extension uses a new authentication type, AuType in the OSPFv2 header (refer to figure 1). When an OSPFv2 packet is received and the AuType doesn't match the configured authentication type for the interface, the OSPFv2 packet will be dropped as specified in RFC 2328 [RFC2328]. This guarantees backward compatible behavior consistent with any other authentication type mismatch.

## 8. Security Considerations

This document rectifies the manual key management procedure that currently exists within OSPFv2, as part of the Phase 1 of the KARP Working Group. Therefore, only the OSPFv2 manual key management mechanism is considered. Any solution that takes advantage of the

automatic key management mechanism is beyond the scope of this document.

The described sequence number extension offers most of the benefits of more complicated mechanisms without their attendant challenges. There are, however, a couple drawbacks to this approach. First, it requires the OSPF implementation to be able to save its boot count in non-volatile storage. If the non-volatile storage is ever repaired or upgraded such that the contents are lost or the OSPFv2 router is replaced, the authentication keys **MUST** be changed to prevent replay attacks.

Second, if a router is taken out of service completely (either intentionally or due to a persistent failure), the potential exists for reestablishment of an OSPFv2 adjacency by replaying the entire OSPFv2 session establishment. However, this scenario is extremely unlikely, since it would imply an identical OSPFv2 adjacency formation packet exchange. Without adjacency formation, the replay of OSPFv2 hello packets alone for an OSPFv2 router that has been taken out of service should not result in any serious attack as the only consequence is superfluous processing. Of course, this attack could also be thwarted by changing the relevant manual keys.

This document also provides a solution to prevent certain denial of service attacks that can be launched by changing the source address in the IP header of an OSPFv2 protocol packet.

Using a single crypto sequence number can leave the router vulnerable to a replay attack where it uses the same source IP address on two different point-to-point unnumbered links. In such environments where an attacker can actively tap the point-to-point links, its recommended that the user employs different keys on each of those unnumbered IP interfaces.

## 9. IANA Considerations

This document requests a new code point from the "OSPF Shortest Path First (OSPF) Authentication Codes" registry:

- o 3 - Cryptographic Authentication with Extended Sequence Numbers.

This document also requests a new code point from the "Authentication Cryptographic Protocol ID" registry defined under "Keying and Authentication for Routing Protocols (KARP) Parameters":

- o TBD (3 Suggested) - OSPFv2.



## 10. References

### 10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC5709] Bhatia, M., Manral, V., Fanto, M., White, R., Barnes, M., Li, T., and R. Atkinson, "OSPFv2 HMAC-SHA Cryptographic Authentication", RFC 5709, October 2009.

### 10.2. Informative References

- [FIPS-198] US National Institute of Standards & Technology, "The Keyed-Hash Message Authentication Code (HMAC)", FIPS PUB 198 , March 2002.
- [RFC1213] McCloghrie, K. and M. Rose, "Management Information Base for Network Management of TCP/IP-based internets:MIB-II", STD 17, RFC 1213, March 1991.
- [RFC2104] Krawczyk, H., Bellare, M., and R. Canetti, "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, February 1997.
- [RFC2574] Blumenthal, U. and B. Wijnen, "User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)", RFC 2574, April 1999.
- [RFC4222] Choudhury, G., "Prioritized Treatment of Specific OSPF Version 2 Packets and Congestion Avoidance", BCP 112, RFC 4222, October 2005.
- [RFC4822] Atkinson, R. and M. Fanto, "RIPv2 Cryptographic Authentication", RFC 4822, February 2007.
- [RFC5310] Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, February 2009.
- [RFC6039] Manral, V., Bhatia, M., Jaeggli, J., and R. White, "Issues with Existing Cryptographic Protection Methods for Routing Protocols", RFC 6039, October 2010.
- [RFC6094] Bhatia, M. and V. Manral, "Summary of Cryptographic

Authentication Algorithm Implementation Requirements for Routing Protocols", RFC 6094, February 2011.

- [RFC6549] Lindem, A., Roy, A., and S. Mirtorabi, "OSPFv2 Multi-Instance Extensions", RFC 6549, March 2012.
- [RFC6862] Lebovitz, G., Bhatia, M., and B. Weis, "Keying and Authentication for Routing Protocols (KARP) Overview, Threats, and Requirements", RFC 6862, March 2013.
- [RFC6863] Hartman, S. and D. Zhang, "Analysis of OSPF Security According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6863, March 2013.
- [RFC7210] Housley, R., Polk, T., Hartman, S., and D. Zhang, "Database of Long-Lived Symmetric Cryptographic Keys", RFC 7210, April 2014.

#### Authors' Addresses

Manav Bhatia  
Ionos Networks  
Bangalore,  
India

Email: manav@ionosnetworks.com

Sam Hartman  
Painless Security

Email: hartmans@painless-security.com

Dacheng Zhang  
Huawei Technologies co., LTD.  
Beijing,  
China

Email: zhangdacheng@huawei.com  
URI:

Acee Lindem (editor)  
Cisco  
USA

Email: [acee@cisco.com](mailto:acee@cisco.com)



OSPF  
Internet-Draft  
Updates: 5786 (if approved)  
Intended status: Standards Track  
Expires: July 14, 2017

A. Smirnov  
A. Retana  
M. Barnes  
Cisco Systems, Inc.  
January 10, 2017

OSPF Routing with Cross-Address Family MPLS Traffic Engineering Tunnels  
draft-smirnov-ospf-xaf-te-07

Abstract

When using Traffic Engineering (TE) in a dual-stack IPv4/IPv6 network the Multiprotocol Label Switching (MPLS) TE Label Switched Paths (LSP) infrastructure may be duplicated, even if the destination IPv4 and IPv6 addresses belong to the same remote router. In order to achieve an integrated MPLS TE LSP infrastructure, OSPF routes must be computed over MPLS TE tunnels created using information propagated in another OSPF instance. This is solved by advertising cross-address family (X-AF) OSPF TE information.

This document describes an update to RFC5786 that allows for the easy identification of a router's local X-AF IP addresses.

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 <http://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 July 14, 2017.

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents

(<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	2
2. Requirements Language . . . . .	3
3. Operation . . . . .	3
4. Backward Compatibility . . . . .	5
5. Security Considerations . . . . .	6
6. IANA Considerations . . . . .	6
7. Acknowledgements . . . . .	6
8. References . . . . .	6
8.1. Normative References . . . . .	6
8.2. Informative References . . . . .	7
Authors' Addresses . . . . .	7

## 1. Introduction

TE Extensions to OSPFv2 [RFC3630] and to OSPFv3 [RFC5329] have been described to support intra-area TE in IPv4 and IPv6 networks, respectively. In both cases the TE database provides a tight coupling between the routed protocol and TE signaling information in it. In other words, any use of the TE link state database is limited to IPv4 for OSPFv2 [RFC2328] and IPv6 for OSPFv3 [RFC5340].

In a dual stack network it may be desirable to set up common MPLS TE LSPs to carry traffic destined to addresses from different address families on a router. The use of common LSPs eases potential scalability and management concerns by halving the number of LSPs in the network. Besides, it allows operators to group traffic based on business characteristics and/or applications or class of service, not constrained by the network protocol which carries it.

For example, an LSP created based on MPLS TE information propagated by OSPFv2 instance can be defined to carry both IPv4 and IPv6 traffic, instead of having both OSPFv2 and OSPFv3 to provision a separate LSP for each address family. Even if in some cases the address family-specific traffic is to be separated, the calculation from a common database may prove operationally beneficial.

A requirement when creating a common MPLS TE infrastructure is the ability to reliably map the X-AF family addresses to the

corresponding advertising tail-end router. This mapping is a challenge because the LSAs containing the routing information are carried in one OSPF instance while the TE calculation may be done using a TE database from a different instance.

A simple solution to this problem is to rely on the Router ID to identify a node in the corresponding OSPFv2 and OSPFv3 databases. This solution would mandate both instances on the same router to be configured with the same Router ID. However, relying on the correctness of the configuration puts additional burden on network management and adds cost to the operation of the network. The network becomes even more difficult to manage if OSPFv2 and OSPFv3 topologies do not match exactly, for example if area borders are drawn differently in the two protocols. Also, if the routing processes do fall out of sync (having different Router IDs, even if for local administrative reasons), there is no defined way for other routers to discover such misalignment and to take any corrective measures (such as to avoid routing through affected TE tunnels or issuing warning to network management). The use of misaligned router IDs may result in delivering the traffic to the wrong tail-end router, which could lead to suboptimal routing or even traffic loops.

This document describes an update to [RFC5786] that allows for the easy identification of a router's local X-AF IP addresses. Routers using the Node Attribute TLV [RFC5786] can include non-TE enabled interface addresses in their OSPF TE advertisements, and also use the same sub-TLVs to carry X-AF information, facilitating the mapping mentioned above.

The method described in this document can also be used to compute X-AF mapping of egress LSR for sub-LSPs of a Point-to-Multipoint LSP (see [RFC4461]). Considerations of using Point-to-Multipoint MPLS TE for X-AF traffic forwarding is outside the scope of this specification.

## 2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

## 3. Operation

[RFC5786] defined the Node IPv4 Local Address and Node IPv6 Local Address sub-TLVs of the Node Attribute TLV for a router to advertise additional local IPv4 and IPv6 addresses. To solve the problem outlined in [RFC5786] OSPFv2 would advertise and use only IPv4 addresses and OSPFv3 would advertise and use only IPv6 addresses.

This document updates [RFC5786] so that a router can also announce one or more local X-AF addresses using the corresponding Local Address sub-TLV. In other words, to implement the X-AF routing technique proposed in this document, OSPFv2 will advertise the Node IPv6 Local Address sub-TLV and OSPFv3 will advertise the Node IPv4 Local Address sub-TLV, possibly in addition to advertising other IP addresses as documented by [RFC5786].

A node that implements X-AF routing SHOULD advertise in the corresponding Node Local Address sub-TLV all X-AF IP addresses local to the router that can be used by Constrained SPF (CSPF) to calculate MPLS TE LSPs. In general, OSPF SHOULD advertise the IP address listed in the Router Address TLV of the X-AF instance maintaining MPLS TE database plus any additional local addresses advertised by the X-AF OSPF instance in its Node Local Address sub-TLV. Implementation MAY advertise other local X-AF addresses.

If the Node Attribute TLV carries both the Node IPv4 Local Address sub-TLV and the Node IPv6 Local Address sub-TLV, then the X-AF component must be considered for the consolidated calculation of MPLS TE LSPs. Both instances may carry the required information, it is left to local configuration to determine which database is used.

On Area Border Routers (ABR), each advertised X-AF IP address MUST be advertised into at most one area. If OSPFv2 and OSPFv3 area borders match (i.e. for each interface area number for OSPFv2 and OSPFv3 instances is numerically equal), then the X-AF addresses MUST be advertised into the same area in both instances. This allows other ABRs connected to the same set of areas to know with which area to associate MPLS TE tunnels.

During the X-AF routing calculation, X-AF IP addresses are used to map locally created LSPs to tail-end routers in the LSDB. The mapping algorithm can be described as:

Walk the list of all MPLS TE tunnels for which the computing router is a head-end. For each MPLS TE tunnel T:

1. If T's destination IP address is from the same address family as the computing OSPF instance, then the tunnel must have been signaled based on MPLS TE information propagated in the same OSPF instance. Process the tunnel as per [RFC3630] or [RFC5329].
2. Otherwise it is a X-AF MPLS TE tunnel. Note tunnel's destination IP address.
3. Walk the X-AF IP addresses in the LSDBs of all connected areas. If a matching IP address is found, advertised by router R in area



A, then mark the tunnel T as belonging to area A and terminating on tail-end router R. Assign an intra-area SPF cost to reach router R within area A as the IGP cost of tunnel T.

After completing this calculation, each TE tunnel is associated with an area and tail-end router in terms of the routing LSDB of the computing OSPF instance and has a metric.

Note that for clarity of description the mapping algorithm is specified as a single calculation. Actual implementations for the efficiency may choose to support equivalent mapping functionality without implementing the algorithm exactly as it is described.

As an example lets consider a router in dual-stack network running OSPFv2 and OSPFv3 for IPv4 and IPv6 routing correspondingly. Suppose OSPFv2 instance is used to propagate MPLS TE information and the router is configured to accept TE LSPs terminating at local addresses 198.51.100.1 and 198.51.100.2. Then the router will advertise into OSPFv2 instance IPv4 address 198.51.100.1 in the Router Address TLV, additional local IPv4 address 198.51.100.2 in the Node IPv4 Local Address sub-TLV, plus other Traffic Engineering TLVs as required by [RFC3630]. If OSPFv3 instance in the network is enabled for X-AF TE routing (that is, to use for IPv6 routing MPLS TE LSPs computed by OSPFv2), then the OSPFv3 instance of the router will advertise the Node IPv4 Local Address sub-TLV listing local IPv4 addresses 198.51.100.1 and 198.51.100.2. Other routers in the OSPFv3 network will use this information to reliably identify this router as egress LSR for MPLS TE LSPs terminating at either 198.51.100.1 or 198.51.100.2.

#### 4. Backward Compatibility

Node Attribute TLV and Node Local Address sub-TLVs and their usage are defined in [RFC5786] and updated by [RFC6827]. Way of using these TLVs as specified in this document is fully backward compatible with previous standard documents.

An implementation processing Node Attribute TLV MUST interpret its content as follows:

- o If the Node Attribute TLV contains Local TE Router ID sub-TLV then this Node Attribute TLV MUST be treated as carrying routing information for ASON (Automatically Switched Optical Network) and processed as specified in [RFC6827].
- o Otherwise Node Attribute TLV contains one or more instance(s) of Node IPv4 Local Address and/or Node IPv6 Local Address sub-TLVs.

Meaning of each Local Address sub-TLV has to be identified separately.

- \* If Node Local Address sub-TLV belongs to the same address family as instance of OSPF protocol advertising it then address carried in the sub-TLV MUST be treated as described in [RFC5786].
- \* Otherwise the address is used for X-AF tunnel tail-end mapping as defined by this document.

## 5. Security Considerations

This document introduces no new security concerns. Security considerations of using Node Attribute TLV are discussed in [RFC5786].

## 6. IANA Considerations

This document has no IANA actions.

## 7. Acknowledgements

The authors would like to thank Peter Psenak and Eric Osborne for early discussions and Acee Lindem for discussing compatibility with ASON extensions.

We would also like to thank the authors of RFC5786 for laying down the foundation for this work.

## 8. References

### 8.1. 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, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC5786] Aggarwal, R. and K. Kompella, "Advertising a Router's Local Addresses in OSPF Traffic Engineering (TE) Extensions", RFC 5786, DOI 10.17487/RFC5786, March 2010, <<http://www.rfc-editor.org/info/rfc5786>>.

## 8.2. Informative References

- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <<http://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, <<http://www.rfc-editor.org/info/rfc3630>>.
- [RFC4461] Yasukawa, S., Ed., "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)", RFC 4461, DOI 10.17487/RFC4461, April 2006, <<http://www.rfc-editor.org/info/rfc4461>>.
- [RFC5329] Ishiguro, K., Manral, V., Davey, A., and A. Lindem, Ed., "Traffic Engineering Extensions to OSPF Version 3", RFC 5329, DOI 10.17487/RFC5329, September 2008, <<http://www.rfc-editor.org/info/rfc5329>>.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <<http://www.rfc-editor.org/info/rfc5340>>.
- [RFC6827] Malis, A., Ed., Lindem, A., Ed., and D. Papadimitriou, Ed., "Automatically Switched Optical Network (ASON) Routing for OSPFv2 Protocols", RFC 6827, DOI 10.17487/RFC6827, January 2013, <<http://www.rfc-editor.org/info/rfc6827>>.

## Authors' Addresses

Anton Smirnov  
Cisco Systems, Inc.  
De kleetlaan 6a  
Diegem 1831  
Belgium

Email: [as@cisco.com](mailto:as@cisco.com)

Alvaro Retana  
Cisco Systems, Inc.  
7025 Kit Creek Rd.  
Research Triangle Park, NC 27709  
USA

Email: aretana@cisco.com

Michael Barnes  
Cisco Systems, Inc.  
510 McCarthy Blvd.  
Milpitas, CA 95035  
USA

Email: mjbarnes@cisco.com

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: April 13, 2015

X. Xu  
Huawei  
S. Kini  
Ericsson  
S. Sivabalan  
C. Filsfils  
Cisco  
S. Litkowski  
Orange  
October 10, 2014

Signaling Entropy Label Capability Using OSPF  
draft-xu-ospf-mpls-elc-01

Abstract

Multi Protocol Label Switching (MPLS) has defined a mechanism to load balance traffic flows using Entropy Labels (EL). An ingress LSR cannot insert ELs for packets going into a given tunnel unless an egress LSR has indicated via signaling that it can process ELs on that tunnel. This draft defines a mechanism to signal that capability using OSPF. This mechanism is useful when the label advertisement is also done via OSPF.

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 <http://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 April 13, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

1. Introduction . . . . .	2
1.1. Requirements Language . . . . .	3
2. Terminology . . . . .	3
3. Advertising ELC Using OSPF . . . . .	3
4. Advertising RLSDC Using OSPF . . . . .	3
5. Acknowledgements . . . . .	3
6. IANA Considerations . . . . .	4
7. Security Considerations . . . . .	4
8. References . . . . .	4
8.1. Normative References . . . . .	4
8.2. Informative References . . . . .	4
Authors' Addresses . . . . .	5

## 1. Introduction

Multi Protocol Label Switching (MPLS) has defined a method in [RFC6790] to load balance traffic flows using Entropy Labels (EL). An ingress LSR cannot insert ELs for packets going into a given tunnel unless an egress LSR has indicated that it can process ELs on that tunnel. [RFC6790] defines the signaling of this capability (a.k.a Entropy Label Capability - ELC) via signaling protocols. Recently, mechanisms are being defined to signal labels via link state Interior Gateway Protocols (IGP) such as OSPF [I-D.ietf-ospf-segment-routing-extensions]. In such scenario the signaling mechanisms defined in [RFC6790] are inadequate. This draft defines a mechanism to signal the ELC using OSPF. This mechanism is useful when the label advertisement is also done via OSPF. In addition, in the cases where stacked LSPs are used for whatever reasons (e.g., SPRING-MPLS [I-D.gredler-spring-mpls] [I-D.filsfils-spring-segment-routing-mpls]), it would be useful for ingress LSRs to know each LSR's capability of reading the maximum label stack depth. This capability, referred to as Readable Label Stack Depth Capability (RLSDC) can be used by ingress LSRs to determine whether it's necessary to insert an EL for a given LSP tunnel in the case where there has already been at least one EL in the label stack [I-D.kini-mpls-spring-entropy-label]. Of course,

even it has been determined that it's necessary to insert an EL for a given LSP tunnel, if the egress LSR of that LSP tunnel has not yet indicated that it can process ELs for that tunnel, the ingress LSR MUST NOT include an entropy label for that tunnel as well.

### 1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

## 2. Terminology

This memo makes use of the terms defined in [RFC6790] and [RFC4970].

## 3. Advertising ELC Using OSPF

The OSPF Router Information (RI) Opaque LSA defined in [RFC4970] is used by OSPF routers to announce their capabilities. A new TLV within the body of this LSA, called ELC TLV is defined to advertise the capability of the router to process the ELs. It is formatted as described in Section 2.1 of [RFC4970]. This TLV is applicable to both OSPFv2 and OSPFv3. The Type for the ELC TLV needs to be assigned by IANA and it has a Length of zero. The scope of the advertisement depends on the application but it is recommended that it SHOULD be AS-scoped.

## 4. Advertising RLSDC Using OSPF

A new TLV within the body of the OSPF RI LSA, called RLSDC TLV is defined to advertise the capability of the router to read the maximum label stack depth. It is formatted as described in Section 2.1 of [RFC4970] with a Type code to be assigned by IANA and a Length of one. The Value field is set to the maximum readable label stack depth in the range between 1 to 255. The scope of the advertisement depends on the application but it is RECOMMENDED that it SHOULD be domain-wide. If a router has multiple linecards with different capabilities of reading the maximum label stack depth, the router MUST advertise the smallest one in the RLSDC TLV. This TLV is applicable to both OSPFv2 and OSPFv3.

## 5. Acknowledgements

The authors would like to thank Yimin Shen and George Swallow for their comments.

## 6. IANA Considerations

This memo includes a request to IANA to allocate two TLV types from the OSPF RI TLVs registry.

## 7. Security Considerations

This document does not introduce any new security risk.

## 8. References

### 8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 4970, July 2007.

### 8.2. Informative References

- [I-D.filsfils-spring-segment-routing-mpls]  
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing with MPLS data plane", draft-filsfils-spring-segment-routing-mpls-03 (work in progress), August 2014.
- [I-D.gredler-spring-mpls]  
Gredler, H., Rekhter, Y., Jalil, L., Kini, S., and X. Xu, "Supporting Source/Explicitly Routed Tunnels via Stacked LSPs", draft-gredler-spring-mpls-06 (work in progress), May 2014.
- [I-D.ietf-ospf-segment-routing-extensions]  
Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", draft-ietf-ospf-segment-routing-extensions-02 (work in progress), August 2014.
- [I-D.kini-mpls-spring-entropy-label]  
Kini, S., Kompella, K., Sivabalan, S., Litkowski, S., Shakir, R., Xu, X., Henderickx, W., and J. Tantsura, "Entropy labels for source routed stacked tunnels", draft-kini-mpls-spring-entropy-label-01 (work in progress), September 2014.



[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and  
L. Yong, "The Use of Entropy Labels in MPLS Forwarding",  
RFC 6790, November 2012.

Authors' Addresses

Xiaohu Xu  
Huawei

Email: xuxiaohu@huawei.com

Sriganesh Kini  
Ericsson

Email: sriganesh.kini@ericsson.com

Siva Sivabalan  
Cisco

Email: msiva@cisco.com

Clarence Filsfils  
Cisco

Email: cfilsfil@cisco.com

Stephane Litkowski  
Orange

Email: stephane.litkowski@orange.com

Open Shortest Path First  
Internet-Draft  
Updates: 2328, 5340 (if approved)  
Intended status: Standards Track  
Expires: April 17, 2015

Z. Zhang  
L. Wang  
Juniper Networks, Inc.  
A. Lindem  
Cisco Systems  
D. Dubois  
General Dynamics C4S  
V. Julka  
T. McMillan  
L3 Communications, Linkabit  
October 14, 2014

OSPF Two-part Metric  
draft-zzhang-ospf-two-part-metric-05.txt

Abstract

This document specifies an optional extension to the OSPF protocol, to represent the metric on a multi-access network as two parts: the metric from a router to the network, and the metric from the network to the router. The router to router metric would be the sum of the two.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC2119.

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 <http://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 April 17, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

#### Table of Contents

1. Introduction . . . . .	3
2. Proposed Enhancement . . . . .	3
3. Specifications . . . . .	4
3.1. Router Interface Parameters . . . . .	4
3.2. Advertising Network-to-Router metric in OSPFv2 . . . . .	5
3.3. Advertising Network-to-Router metric in OSPFv3 . . . . .	5
3.4. SPF Calculation . . . . .	5
3.5. Backward Compatibility . . . . .	6
4. IANA Considerations . . . . .	6
5. Security Considerations . . . . .	6
6. Acknowledgements . . . . .	6
7. References . . . . .	7
7.1. Normative References . . . . .	7
7.2. Informative References . . . . .	8

## 1. Introduction

For a broadcast network, a Network-LSA is advertised to list all routers on the network, and each router on the network includes a link in its Router-LSA to describe its connection to the network. The link in the Router-LSA includes a metric but the listed routers in the Network LSA do not include a metric. This is based on the assumption that from a particular router, all others on the same network can be reached with the same metric.

With some broadcast networks, different routers can be reached with different metrics. RFC 6845 extends the OSPF protocol with a hybrid interface type for that kind of broadcast network, where no Network LSA is advertised and Router-LSAs simply include p2p links to all routers on the same network with individual metrics. Broadcast capability is still utilized to optimize database synchronization and adjacency maintenance.

That works well for broadcast networks where the metric between different pair of routers are really independent. For example, VPLS networks.

With certain types of broadcast networks, further optimization can be made to reduce the size of the Router-LSAs and number of updates.

Consider a satellite radio network with fixed and mobile ground terminals. All communication goes through the satellite. When the mobile terminals move about, their communication capability may change. When OSPF runs over the radio network (routers being or in tandem with the terminals), RFC 6845 hybrid interface can be used, but with the following drawbacks.

Consider that one terminal/router moves into an area where its communication capability degrades significantly. Through the radio control protocol, all other routers determine that the metric to this particular router changed and they all need to update their Router-LSAs accordingly. The router in question also determines that its metric to reach all others also changed and it also needs to update its Router-LSA. Consider that there could be many terminals and many of them can be moving fast and frequently, the number/frequency of updates of those large Router-LSAs could inhibit network scaling.

## 2. Proposed Enhancement

Notice that in the above scenario, when one terminal's communication capability changes, its metric to all other terminals and the metric from all other terminals to it will all change in a similar fashion. Given this, the above problem can be easily addressed by breaking the

metric into two parts: the metric to the satellite and the metric from the satellite. The metric from terminal R1 to R2 would be the sum of the metric from R1 to the satellite and the metric from the satellite to R2.

Now instead of using the RFC 6845 hybrid interface type, the network is just treated as a regular broadcast network. A router on the network no longer lists individual metrics to each neighbor in its Router-LSA. Instead, each router advertises the metric from the network to itself in addition to the normal metric for the network. With the normal Router-to-Network and additional Network-to-Router metrics advertised for each router, individual router-to-router metric can be calculated.

With the proposed enhancement, the size of Router-LSA will be significantly reduced. In addition, when a router's communication capability changes, only that router needs to update its Router-LSA.

Note that while the example uses the satellite as the relay point at the radio level (layer-2), at layer-3, the satellite does not participate in packet forwarding. In fact, the satellite does not need to be running any layer-3 protocol. Therefore for generality, the metric is abstracted as to/from the "network" rather than specifically to/from the "satellite".

### 3. Specifications

The following protocol specifications are added to or modified from the base OSPF protocol. If an area contains one or more two-part metric networks, then all routers in the area must support the extensions specified herein. This is ensured by procedures described in Section 3.5.

#### 3.1. Router Interface Parameters

The "Router interface parameters" have the following additions:

- o Two-part metric: TRUE if the interface connects to a multi-access network that uses two-part metric. All routers connected to the same network SHOULD have the same configuration for their corresponding interfaces.
- o Interface input cost: Link state metric from the two-part-metric network to this router. Defaulted to "Interface output cost" but not valid for normal networks using a single metric. May be configured or dynamically adjusted to a value different from the "Interface output cost".

### 3.2. Advertising Network-to-Router metric in OSPFv2

For OSPFv2, the Network-to-Router metric is encoded in an OSPF Extended Link TLV Sub-TLV [ietf-ospf-lsa-extend], defined in this document as the Network-to-Router Metric Sub-TLV. The type of the Sub-TLV is TBD. The length of the Sub-TLV is 4 (for the value part only). The value part of the Sub-TLV is defined as follows:

```

      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
+-----+-----+-----+-----+-----+-----+-----+-----+
|               MT               | 0 |               MT   metric   |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Multiple such Sub-TLVs can exist in a single OSPF Extended Link TLV, one for each topology. The OSPF Extended Link TLV identifies the transit link to the network, and is part of an OSPFv2 Extended-Link Opaque LSA. The Sub-TLV MUST ONLY appear in Extended-Link TLVs for Link Type 2 (link to transit network), and MUST be ignored if received for other link types.

### 3.3. Advertising Network-to-Router metric in OSPFv3

For OSPFv3, the same Network-to-Router Metric Sub-TLV definition is used, though it is part of the Router-Link TLV of E-Router-LSA [ietf-ospf-ospfv3-lsa-extend]. Currently OSPFv3 Multi-Topology is not defined so the only valid value for the MT field is 0 and only one such Sub-TLV SHOULD be included in the Router-Link TLV. Received Sub-TLVs with non-zero MT field MUST be ignored.

Similarly, the Sub-TLV MUST ONLY appear in Router-Link TLVs for Link Type 2 (connection to a transit network) and MUST be ignored if received for other link types.

### 3.4. SPF Calculation

During the first stage of shortest-path tree calculation for an area, when a vertex V corresponding to a Network-LSA is added to the shortest-path tree and its adjacent vertex W (joined by a link in V's corresponding Network LSA), the cost from V to W, which is W's network-to-router cost, is determined as follows:

- o For OSPFv2, if vertex W has a corresponding Extended-Link Opaque LSA with an Extended Link TLV for the link from W to V, and the Extended Link TLV has a Network-to-Router Metric Sub-TLV for the corresponding topology, then the cost from V to W is the metric in the Sub-TLV. Otherwise, the cost is 0.

- o For OSPFv3, if vertex W has a corresponding E-Router-LSA with a Router-Link TLV for the link from W to V, and the Router-Link TLV has a Network-to-Router Metric Sub-TLV, then the cost from V to W is the metric in the Sub-TLV. If not, the cost is 0.

### 3.5. Backward Compatibility

Due to the change of procedures in the SPF calculation, all routers in an area that includes one or more two-part metric networks must support the changes specified in this document. To ensure that, if an area is provisioned to support two-part metric networks, all routers supporting this capability must advertise a Router Information (RI) LSA with a Router Functional Capabilities TLV [acee-ospf-rfc4970bis] that includes the following Router Functional Capability Bit:

Bit	Capabilities
0	OSPF Two-part Metric [TPM]

Upon detecting the presence of a reachable Router-LSA without a companion RI LSA that has the bit set, all routers MUST disable the two-part metric functionalities and take the following actions:

- o If this router currently advertises network-to-router costs, remove the Network-to-Router Metric Sub-TLVs. This may lead to removal of parent TLVs and even withdrawal of the parent LSAs.
- o Recalculate routes w/o considering any network-to-router costs.

### 4. IANA Considerations

This document requests IANA to assign a new bit in the Router Functional Capabilities TLV to indicate the capability of supporting two-part metric, a new Sub-TLV in the OSPF Extended-Link TLV Sub-TLV Registry, and a new Sub-TLV in the The OSPFv3 Extend-LSA Sub-TLV registry.

### 5. Security Considerations

This document does not introduce new security risks.

### 6. Acknowledgements

The authors would like to thank Abhay Roy, Hannes Gredler, Peter Psenak and Eric Wu for their comments and suggestions.

### 7. References

## 7.1. Normative References

- [I-D.acee-ospf-rfc4970bis] Lindem, A., Shen, N., Vasseur, J., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", draft-acee-ospf-rfc4970bis-00 (work in progress), July 2014.
- [I-D.ietf-ospf-lsa-extend] Psenak, P., Previdi, S., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 LSA Extendibility", draft-ietf-ospf-lsa-extend-00 (work in progress), August 2014.
- [I-D.ietf-ospf-ospfv3-lsa-extend] Lindem, A., Mirtorabi, S., Roy, A., and F. Baker, "OSPFv3 LSA Extendibility", draft-ietf-ospf-ospfv3-lsa-extend-04 (work in progress), September 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.
- [RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, June 2007.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [RFC5613] Zinin, A., Roy, A., Nguyen, L., Friedman, B., and D. Yeung, "OSPF Link-Local Signaling", RFC 5613, August 2009.



## 7.2. Informative References

[RFC6845]

Sheth, N., Wang, L., and J. Zhang,  
"OSPF Hybrid Broadcast and Point-  
to-Multipoint Interface Type",  
RFC 6845, January 2013.

## Authors' Addresses

Jeffrey Zhang  
Juniper Networks, Inc.  
10 Technology Park Drive  
Westford, MA 01886

EMail: zzhang@juniper.net

Lili Wang  
Juniper Networks, Inc.  
10 Technology Park Drive  
Westford, MA 01886

EMail: liliw@juniper.net

Acee Lindem  
Cisco Systems  
301 Midenhall Way  
Cary, NC 27513

EMail: acee@cisco.com

David Dubois  
General Dynamics C4S  
400 John Quincy Adams Road  
Taunton, MA 02780

EMail: dave.dubois@gdc4s.com

Vibhor Julka  
L3 Communications, Linkabit  
9890 Towne Centre Drive  
San Diego, CA 92121

EMail: vibhor.julka@l-3Com.com

Tom McMillan  
L3 Communications, Linkabit  
9890 Towne Centre Drive  
San Diego, CA 92121

EMail: tom.mcmillan@l-3com.com

