

KARP Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 10, 2011

M. Bhatia
Alcatel-Lucent
October 7, 2010

Mechanism to protect OSPFv2 authentication from IP Layer Issues
draft-bhatia-karp-ospf-ip-layer-protection-00

Abstract

The IP header is not covered by the MAC in the cryptographic authentication scheme as described in RFC 2328 and RFC 5709, and an attack can be made to exploit this omission. This draft proposes a simple change in how the authentication is computed to eliminate most of such attacks.

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 10, 2011.

Copyright Notice

Copyright (c) 2010 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.

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

1. Introduction

The OSPFv2 [RFC2328] cryptographic authentication as described in [RFC2328] and later updated in [RFC5709] does not include the IP header. This can be exploited to launch several attacks as the source address in the IP header is no longer protected. The OSPF specification, in certain cases, requires the implementations to look at the source address carried in the IP header to determine the neighbor the packet was received from. Changing the source address of a packet can thus, confuse the receiver which can be exploited to produce a number of denial of service attacks.

[I-D.ietf-opsec-routing-protocols-crypto-issues]. If the packet is interpreted as coming from a different neighbor, the sequence number received from the neighbor may be updated. This may disrupt communication with the legitimate neighbor. Hello packets may be reflected to cause a neighbor to appear to have one-way communication. Old Database descriptions may be reflected in cases where the per-packet sequence numbers are sufficiently divergent in order to disrupt an adjacency [I-D.hartman-ospf-analysis].

[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] in future deployments.

This draft proposes a simple change in the cryptographic authentication mechanism, as currently described in [RFC5709], to prevent such IP layer attacks.

2. Cryptographic Authentication Mechanism in OSPFv2

The overall cryptographic authentication process defined in [RFC5709] remains unchanged. To reduce the potential for confusion, this section minimises the repetition of text from RFC 5709 and is incorporated here by reference [RFC5709].

RFC 5709, Section 3.3, describes how the cryptographic authentication must be computed. It requires 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 where Apad is a hexadecimal constant value 0x878FE1F3 repeated (L/4) times, where L is the length of the hash being used and is measured in octets rather than bits.

3. Proposed Enhancement

This document updates the definition of Apad which is currently a constant defined in [RFC5709] to the source address that's carried in the IP header of the OSPFv2 protocol packet. Routers at the sending side must initialize Apad to a value of the source address that would be used when sending out the OSPFv2 packet, repeated L/4 times, where L is the length of the hash, measured in octets. The basic idea is to incorporate the source address from the IP header in the cryptographic authentication computation so that any change there can be detected.

At the receiving end implementations MUST initialize Apad as the source address that exists in the IP Header of the incoming OSPFv2 protocol 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 source address of the protocol packet and reflects it on some other interface as the authentication check would fail and all such packets would get rejected.

4. Security Considerations

This document enhances the security of OSPFv2 and provides a solution to prevent certain denial of service attacks that can be launched by changing the source address of the OSPFv2 protocol packet. The proposal defined does not introduce any new security concerns.

5. IANA Considerations

This document has no actions for IANA.

6. Acknowledgements

7. References

7.1. 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.

7.2. Informative References

- [I-D.hartman-ospf-analysis]
Hartman, S. and D. Zhang, "Analysis of OSPF Security According to KARP Design Guide",
draft-hartman-ospf-analysis-01 (work in progress),
June 2010.
- [I-D.ietf-opsec-routing-protocols-crypto-issues]
Jaeggli, J., Hares, S., Bhatia, M., Manral, V., and R. White, "Issues with existing Cryptographic Protection Methods for Routing Protocols",
draft-ietf-opsec-routing-protocols-crypto-issues-07 (work in progress), August 2010.

Author's Address

Manav Bhatia
Alcatel-Lucent
Bangalore,
India

Phone:
Email: manav.bhatia@alcatel-lucent.com

OSPF Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2011

M. Bhatia
Alcatel-Lucent
V. Manral
IP Infusion
A. Lindem
Ericsson
October 15, 2010

Supporting Authentication Trailer for OSPFv3
draft-bhatia-manral-auth-trailer-ospfv3-01

Abstract

Currently OSPFv3 uses IPsec for authenticating protocol packets. However, there are some environments, e.g., Mobile Ad-hoc Networks (MANETs), where IPsec is difficult to configure and maintain, and this mechanism cannot be used. This draft proposes an alternative mechanism that can be used so that OSPFv3 does not depend upon IPsec for authentication.

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 18, 2011.

Copyright Notice

Copyright (c) 2010 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	4
2. Proposed Solution	6
2.1. AT-Bit in Options Field	6
2.2. Basic Operation	7
3. OSPFv3 Security Association	8
4. Authentication Procedure	10
4.1. Authentication Trailer	10
4.2. Cryptographic Authentication Procedure	11
4.3. Cryptographic Aspects	11
4.4. Message Verification	13
5. Security Considerations	14
6. IANA Considerations	15
7. References	16
7.1. Normative References	16
7.2. Informative References	16
Authors' Addresses	18

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. Introduction

Unlike Open Shortest Path First version 2 (OSPFv2) [RFC2328], OSPF for IPv6 (OSPFv3) [RFC5340], does not include the AuType and Authentication fields in its headers for authenticating protocol packets. Instead, OSPFv3 relies on the IPv6 Authentication Header (AH)[RFC4302] and IPv6 Encapsulating Security Payload (ESP) [RFC4303] to provide integrity, authentication, and/or confidentiality.

[RFC4522] describes how IPv6 AH and ESP extension headers can be used to provide authentication and/or confidentiality to OSPFv3.

However, there are some environments, e.g., Mobile Ad-hoc Networks (MANETs), where IPsec is difficult to configure and maintain, and this mechanism cannot be used. There is also an issue with IPsec not being available on some platforms or it requiring an additional license.

[RFC4522] discusses, at length, the reasoning behind using manually configured keys, rather than some automated key management protocol such as IKEv2 [RFC5996]. The primary problem is the lack of suitable key management mechanism, as OSPF adjacencies are formed on a one-to-many basis and most key management mechanisms are designed for a one-to-one communication model. This forces the system administrator to use manually configured security associations (SAs) and cryptographic keys to provide the authentication and, if desired, confidentiality services.

Regarding replay protection [RFC4522] states that:

"As it is not possible as per the current standards to provide complete replay protection while using manual keying, the proposed solution will not provide protection against replay attacks."

Since there is no replay protection provided there are a number of vulnerabilities in OSPFv3 which have been discussed in [I-D.ietf-opsec-routing-protocols-crypto-issues].

Since there is no deterministic way to differentiate between encrypted and unencrypted ESP packets by simply examining the packet, it could become tricky for some implementations to prioritize certain OSPFv3 packets (Hellos for example) over the others.

This draft proposes a new mechanism that works similar to OSPFv2 [RFC5709] for providing authentication to the OSPFv3 packets and attempts to solve the problems described above for OSPFv3.

Additionally this document describes how HMAC-SHA authentication can

be used for OSPFv3.

By definition, HMAC ([RFC2104] , [FIPS-198]) requires a cryptographic hash function. This document proposes to use any one of SHA-1, SHA-256, SHA-384, or SHA-512 [FIPS-180-3] to authenticate the OSPFv3 packets.

It is believed that [RFC2104] is mathematically identical to [FIPS-198] and it is also believed that algorithms in [RFC4634] are mathematically identical to [FIPS-180-3].

2. Proposed Solution

To perform non-IPsec cryptographic authentication, OSPFv3 routers append a special data block, henceforth referred to as, the authentication trailer to the end of the OSPFv3 packets. The length of the authentication trailer is not included into the length of the OSPFv3 packet, but is included in the IPv6 payload length.

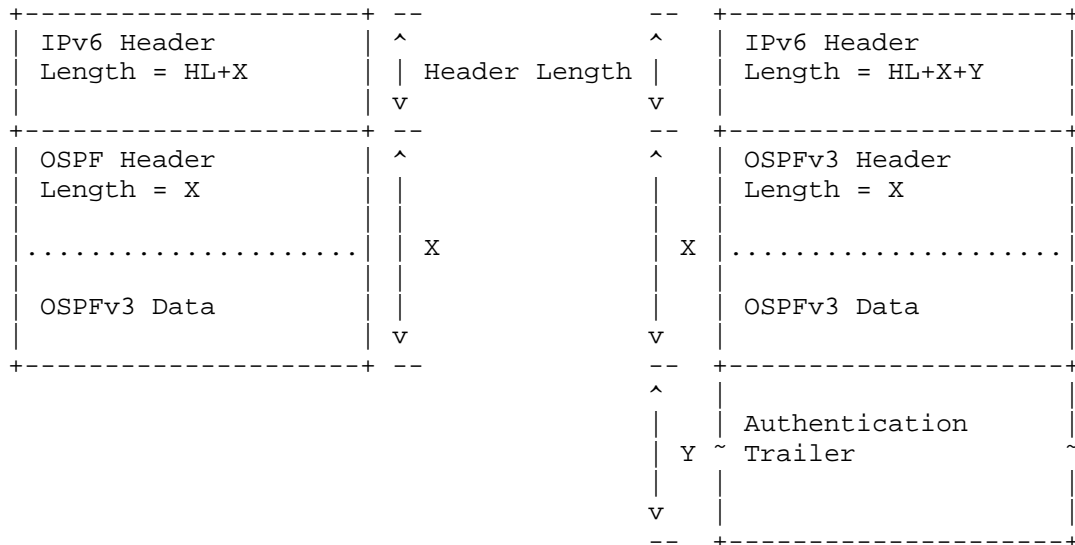


Figure 1: Authentication Trailer in OSPFv3

For the sake of consistency and simplicity the authentication trailer in the OSPFv3 packets MUST be inserted before the link local signalling (LLS) [RFC5613] block, if it exists. This is inline with the authentication mechanism that currently exists for OSPFv2.

2.1. AT-Bit in Options Field

A new AT-bit (AT stands for Authentication Trailer) is introduced into the OSPFv3 Options field. OSPFv3 routers MUST set the AT-bit in OSPFv3 Hello and Database Description packets to indicate that the OSPFv3 router will include the authentication trailer in all OSPFv3 packets on the link. In other words, the authentication trailer is only examined if the AT-bit is set.

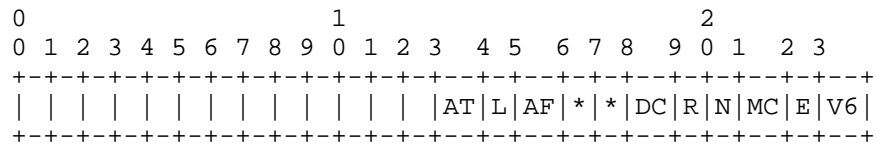


Figure 2: OSPFv3 Options Field

The AT-bit must be set in all OSPFv3 protocol packets that contain an authentication trailer.

2.2. Basic Operation

The procedure followed for computing the Authentication Trailer is exactly the same as described in [RFC5709] and [RFC2328].

The way the authentication data is carried in the Authentication Trailer is very similar to how its done in case of [RFC2328]. The only difference between this mechanism and OSPFv2's authentication mechanism is that for OSPFv3 some additional authentication information in addition to the message digest, is appended to the protocol packet.

3. OSPFv3 Security Association

An OSPFv3 Security Association (SA) contains a set of parameters shared between any two legitimate OSPFv3 speakers.

Parameters associated with an OSPFv3 SA:

- o Key Identifier (Key ID)

This is a 32-bit unsigned integer used to uniquely identify an OSPFv3 SA, as manually configured by the network operator.

The receiver determines the active SA by looking at the Key ID field in the incoming protocol packet.

The sender based on the active configuration, selects an SA to use and puts the correct Key ID value associated with the SA in the OSPFv3 protocol packet. If multiple valid and active OSPFv3 SAs exist for a given interface, the sender may use any of those SAs to protect the packet.

Using Key IDs makes changing keys while maintaining protocol operation convenient. Each key ID specifies two independent parts, the authentication protocol and the authentication key, as explained below.

Normally, an implementation would allow the network operator to configure a set of keys in a key chain, with each key in the chain having fixed lifetime. The actual operation of these mechanisms is outside the scope of this document.

Note that each key ID can indicate a key with a different authentication protocol. This allows the introduction of new authentication mechanisms without disrupting existing OSPFv3 adjacencies.

- o Authentication Algorithm

This signifies the authentication algorithm to be used with the OSPFv3 SA. This information is never sent in cleartext over the wire. Because this information is not sent on the wire, the implementer chooses an implementation specific representation for this information.

At present, the following values are possible:

HMAC-SHA-1, HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512.

- o Authentication Key

This value denotes the cryptographic authentication key associated with the OSPFv3 SA. The length of this key is variable and depends upon the authentication algorithm specified by the OSPFv3 SA.

4. Authentication Procedure

4.1. Authentication Trailer

The authentication trailer that is appended to the OSPFv3 protocol packet is described below:

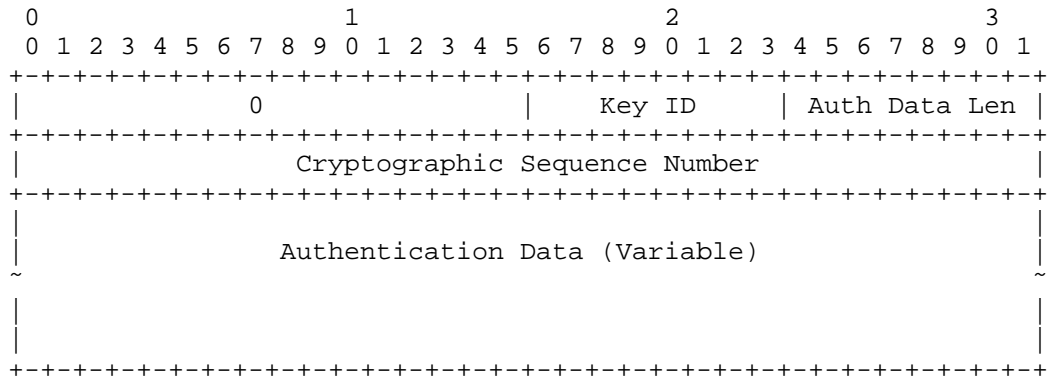


Figure 3: Authentication Trailer Format

The idea is to keep the fields as similar as possible with OSPFv2 so that most of the source code can be reused for authenticating the OSPFv3 protocol packets.

The various fields in the Authentication trailer are:

- o Reserved

16-bit reserved field. The value MUST be initialized to zero by the sender, and MUST be ignored by the receiver.

- o Key ID (Identifier)

32-bit field that identifies the algorithm and the secret key used to create the message digest appended to the OSPFv3 protocol packet. Key Identifiers are unique per-interface.

- o Cryptographic Sequence Number

32-bit non-decreasing sequence number that is used to guard against replay attacks.

- o Authentication Data

Variable data that is carrying the digest of the protocol packet.

4.2. Cryptographic Authentication Procedure

As noted earlier the algorithms used to generate and verify the message digest are specified implicitly by the secret key. This specification discusses the computation of OSPFv3 Cryptographic Authentication data when any of the NIST SHS family of algorithms is used in the Hashed Message Authentication Code (HMAC) mode.

The currently valid algorithms (including mode) for OSPFv3 Cryptographic Authentication include:

HMAC-SHA-1, HMAC-SHA-256, HMAC-SHA-384 and HMAC-SHA-512

Of the above, implementations of this specification MUST include support for at least HMAC-SHA-1 and SHOULD include support for HMAC-SHA-256 and MAY also include support for HMAC-SHA-384 and HMAC-SHA-512.

4.3. Cryptographic Aspects

In the algorithm description below, the following nomenclature, which is consistent with [FIPS-198], is used:

H is the specific hashing algorithm (e.g. SHA-256).

K is the Authentication Key for the OSPFv3 security association.

Ko is the cryptographic key used with the hash algorithm.

B is the block size of H, measured in octets rather than bits.

Note that B is the internal block size, not the hash size.

For SHA-1 and SHA-256: B == 64

For SHA-384 and SHA-512: B == 128

L is the length of the hash, measured in octets rather than bits.

XOR is the exclusive-or operation.

Opad is the hexadecimal value 0x5c repeated B times.

Ipad is the hexadecimal value 0x36 repeated B times.

Apad is the hexadecimal value 0x878FE1F3 repeated (L/4) times.

Implementation Note:

This definition of Apad means that Apad is always the same length as the hash output.

1. Preparation of the Key

In this application, Ko is always L octets long.

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

2. First Hash

First, the OSPFv3 packet's Authentication Trailer (which is very similar to the appendage described in RFC 2328, Section D.4.3, Page 233, items(6)(a) and (6)(d)) is filled with the value Apad.

Then, a First-Hash, also known as the inner hash, is computed as follows:

$$\text{First-Hash} = H(\text{Ko XOR Ipad} \parallel (\text{OSPFv3 Packet}))$$

Implementation Notes:

Note that the First-Hash above includes the Authentication Trailer containing the Apad value, as well as the OSPFv3 packet, as per RFC 2328, Section D.4.3.

The definition of Apad (above) ensures it is always the same length as the hash output. This is consistent with RFC 2328. The "(OSPFv3 Packet)" mentioned in the First-Hash (above) does include the OSPF Authentication Trailer.

The digest length for SHA-1 is 20 bytes; for SHA-256, 32 bytes; for SHA-384, 48 bytes; and for SHA-512, 64 bytes.

3. Second Hash

Then a second hash, also known as the outer hash, is computed as follows:

Second-Hash = H(Ko XOR Opad || First-Hash)

4. Result

The resulting Second-Hash becomes the authentication data that is sent in the Authentication Trailer of the OSPFv3 packet. The length of the authentication data is always identical to the message digest size of the specific hash function H that is being used.

This also means that the use of hash functions with larger output sizes will also increase the size of the OSPFv3 packet as transmitted on the wire.

Implementation Note:

RFC 2328, Appendix D specifies that the Authentication Trailer is not counted in the OSPF packet's own Length field, but is included in the packet's IP Length field. Similar to this, the Authentication Trailer is not included in OSPFv3's own Length field, but is included in IPv6's payload length.

4.4. Message Verification

A router would determine that OSPFv3 is using an Authentication trailer by examining the AT-bit in the Options field in the OSPFv3 header for Hello and Database Description packets. The specification in the Hello and Database description options indicates that other OSPFv3 packets will include the authentication trailer.

Authentication algorithm dependent processing needs to be performed, using the algorithm specified by the appropriate OSPFv3 SA for the received packet.

Before an implementation performs any processing it needs to save the values of the Authentication data field from the Authentication Trailer appended to the OSPFv3 packet.

It should then set the Authentication data field with Apad before the authentication data is computed. The calculated data is compared with the received authentication data in the Authentication trailer and the packet MUST be discarded if the two do not match. In such a case, an error event SHOULD be logged.

An implementation MAY have a transition mode where it includes the Authentication Trailer in the packets but does not verify this information. This is provided as a transition aid for networks in the process of migrating to the mechanism described in this draft.

5. Security Considerations

The document proposes extensions to OSPFv3 which would make it more secure than what it is today. It does not provide confidentiality as a routing protocol contains information that does not need to be kept secret. It does, however, provide means to authenticate the sender of the packets which is of interest to us.

It should be noted that authentication method described in this document is not being used to authenticate the specific originator of a packet, but is rather being used to confirm that the packet has indeed been issued by a router which had access to the password.

The mechanism described here is not perfect and does not need to be perfect. Instead, this mechanism represents a significant increase in the work function of an adversary attacking the OSPFv3 protocol, while not causing undue implementation, deployment, or operational complexity.

6. IANA Considerations

IANA is requested to allocate AT-bit in the OSPFv3 "Options Registry"

7. References

7.1. 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.

7.2. Informative References

- [FIPS-180-3]
US National Institute of Standards & Technology, "Secure Hash Standard (SHS)", FIPS PUB 180-3 , October 2008.
- [FIPS-198]
US National Institute of Standards & Technology, "The Keyed-Hash Message Authentication Code (HMAC)", FIPS PUB 198 , March 2002.
- [I-D.hartman-ospf-analysis]
Hartman, S. and D. Zhang, "Analysis of OSPF Security According to KARP Design Guide", draft-hartman-ospf-analysis-01 (work in progress), June 2010.
- [I-D.ietf-opsec-routing-protocols-crypto-issues]
Jaeggli, J., Hares, S., Bhatia, M., Manral, V., and R. White, "Issues with existing Cryptographic Protection Methods for Routing Protocols", draft-ietf-opsec-routing-protocols-crypto-issues-07 (work in progress), August 2010.
- [RFC2104] Krawczyk, H., Bellare, M., and R. Canetti, "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, February 1997.
- [RFC4302] Kent, S., "IP Authentication Header", RFC 4302, December 2005.
- [RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.
- [RFC4522] Legg, S., "Lightweight Directory Access Protocol (LDAP):

The Binary Encoding Option", RFC 4522, June 2006.

- [RFC4634] Eastlake, D. and T. Hansen, "US Secure Hash Algorithms (SHA and HMAC-SHA)", RFC 4634, July 2006.
- [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.
- [RFC5996] Kaufman, C., Hoffman, P., Nir, Y., and P. Eronen, "Internet Key Exchange Protocol Version 2 (IKEv2)", RFC 5996, September 2010.

Authors' Addresses

Manav Bhatia
Alcatel-Lucent
Bangalore,
India

Phone:
Email: manav.bhatia@alcatel-lucent.com

Vishwas
IP Infusion
USA

Phone:
Email: vishwas@ipinfusion.com

Acee Lindem
Ericsson
102 Carric Bend Court
Cary, NC 27519
USA

Phone:
Email: acee.lindem@ericsson.com

OSPF Working Group
Internet Draft
Intended status: Standard Track
Expires: November 22, 2010

Dimitri Papadimitriou
Alcatel-Lucent
May 23, 2010

Phased OSPF Link-State Database Synchronization
draft-dimitri-ospf-phased-db-sync-00.txt

Status of this Memo

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

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

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 November 22, 2010.

Copyright Notice

Copyright (c) 2010 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.

Abstract

Opaque Link-State Advertisements (LSA) extend the topological link state of Open Shortest Path First (OSPF). The information contained in Opaque LSA may be used directly by OSPF or indirectly by some application wishing to distribute information throughout the OSPF domain. However the Link-State Database (LSDB) synchronization process is kept unified, i.e., there is no messaging or processing allowing to order the exchanges in the link state database synchronization process. We call this ordering, phasing of logically segmented LSDB into Opaque and non-Opaque. The motivation is to prevent delaying reaching Full state whereas synchronizing over the entire LSDB would delay full adjacency establishment.

Table of Contents

1. Introduction.....	3
2. Conventions used in this document.....	4
3. Link-State Database (LSDB) Synchronization.....	4
3.1. LSDB Synchronization: General Description.....	4
3.2. Application of LSDB Synchronization to Opaque LSA.....	5
4. Phased Link-State Database (LSDB) Synchronization.....	6
4.1. Phased Link-State Database (LSDB) Synchronization Process.....	6
4.2. Transition from LSDB to Opaque LSDB Synchronization....	8
4.3. Capability Negotiation.....	8
4.3.1. Option field in Hello Packets.....	9
4.3.2. Link Local Signaling (LLS).....	9
5. Backward Compatibility.....	9
6. Security Considerations.....	10
7. IANA Considerations.....	10
8. References.....	10

8.1. Normative References.....	10
8.2. Informative References.....	10
9. Acknowledgments.....	10

1. Introduction

Open Shortest Path First (OSPF) link-state routing protocol supports a class of Link-State Advertisements (LSA) called Opaque LSAs that provide a generalized mechanism to allow extensibility of OSPF [RFC2328]. The information field contained in Opaque LSAs is often indirectly used by some application wishing to distribute information throughout the OSPF domain. Standard OSPF Link-State Database (LSDB) flooding mechanisms are used to distribute Opaque LSAs to all or some limited portion of the OSPF topology [RFC2370]. Nevertheless, OSPF mandates full synchronization of the LSDB before a routing adjacency is declared in state Full (when LSDB synchronization is completed, the neighbor is in state Full and the two routers are fully adjacent).

The reliable and effective LSDB synchronization but also link-state flooding mechanism provided by OSPF has thus been re-used by many distributed network applications that rely on OSPF to exchange non IP routing information. By non-IP routing information, we mean any information that is not directly or indirectly related to the forwarding of IP datagrams. Another case that often leads to a delayed synchronization process is when the number of entries is not bound by the number of links. This observation also leads us to think that AS-external LSAs in particular are good candidate for the approach proposed here and the mechanism can thus be seen as a complement to [RFC1765].

The proposed mechanism phases the LSDB synchronization process by first exchanging IP routing LSAs (Router, Network, Summary, AS-external, and Not-so-stubby area LSAs) and then, the Opaque LSAs as defined in [RFC2370]. The purpose is to prevent delaying the establishment of fully adjacent routers - at this point the adjacency is listed in LSAs - even if the "Opaque part" of the LSDB is not synchronized. We note here that in most cases the application itself makes use of the IP adjacencies for application specific message exchanges and thus the applications would not be slow down by this process. In this sense, the present document reverts back to [RFC2328] the LSDB synchronization process as extended by [RFC2370] (that covers LSDB including both non-Opaque and Opaque LSAs). Phasing is achieved by logically segmenting the LSDB synchronization process: add "on top of" the LSDB synchronization process

described in [RFC2328], a synchronization process dedicated to Opaque LSAs. This phasing prevents delaying establishment of full adjacency between two routers (Full state) resulting from the time needed to synchronize Opaque LSAs. This condition occurs in particular when the number of Opaque LSAs >> non-Opaque LSAs. The fundamental aspect of the proposed approach consists thus of considering the state Full as the invariant state for reaching full adjacency.

The purpose of the proposed Opaque LSDB Synchronization process is to devise a less drastic alternative to the current approach developed at OSPF WG that mandates complete separation of OSPF instances when Opaque LSA are decoupled from IP Routing [OSPF-TP]. The latter does not actually solve the so-called "Opaque overload" problem because it separates IP-related from non-IP related routing information instead of Opaque from non-Opaque LSAs. The proposed approach here is to avoid duplicating OSPF instances while keeping Opaque LSA messaging and processing as part of a single OSPF instance.

2. Conventions used in this document

In examples, "C:" and "S:" indicate lines sent by the client and server respectively.

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.

3. Link-State Database (LSDB) Synchronization

3.1. LSDB Synchronization: General Description

In link-state routing, it is very important for all routers' Link-State Databases (LSDB) to stay synchronized. OSPF simplifies this by requiring only adjacent routers to remain synchronized. The synchronization process begins as soon as the routers attempt to bring up the routing adjacency.

Each router describes its database by sending a sequence of Database Description (DD) packets to its neighbor. Each Database Description packet describes a set of LSAs belonging to the router's database. When the neighbor sees an LSA that is more recent than its own database copy, it makes a note that this newer LSA should be requested. This sending and receiving of Database Description packets is called the "Database Exchange

Process". During this process, the two routers form a master/slave relationship. Each Database Description packet has a sequence number. Database Description packets sent by the master (polls) are acknowledged by the slave through echoing of the sequence number. Both polls and their responses contain summaries of link state data. The master is the only one allowed to retransmit Database Description packets. It does so only at fixed intervals, the length of which is the configured per-interface constant RxmtInterval.

3.2. Application of LSDB Synchronization to Opaque LSA

Per [RFC2370]: an opaque-capable router learns of its neighbor's opaque capability at the beginning of the "Database Exchange Process" (see Section 10.6 of [RFC2328], receiving Database Description packets from a neighbor in state ExStart). A neighbor is opaque-capable if and only if it sets the O-bit in the Options field of its Database Description packets; the O-bit is not set in packets other than Database Description packets. Then, in the next step of the Database Exchange process, Opaque LSAs are included in the Database summary list that is sent to the neighbor if and only if the neighbor is opaque capable. When flooding Opaque LSAs to adjacent neighbors, an opaque-capable router looks at the neighbor's opaque capability. Opaque LSAs are only flooded to opaque-capable neighbors, i.e., Opaque LSAs are only placed on the link-state retransmission lists of opaque-capable neighbors. In case non-opaque-capable neighbor inadvertently receives Opaque LSAs, the non-opaque-capable router will then simply discard the LSA receiving LSAs having unknown LS types).

Hence, [RFC2370] does not modify the state machine as defined in Section 10.3 of [RFC2328] except for the action associated with State: ExStart, Event: NegotiationDone which is where the Database summary list is built in order to incorporate the Opaque LSA in OSPF (see Figure 1).

4. Phased Link-State Database (LSDB) Synchronization

Compared to the [RFC2370] processing, the Phase Link-State Database (LSDB) synchronization modifies the LSDB exchange process as follows: Opaque LSAs are included in the LSDB summary list that is sent to the neighbor, if and only if

- i) The neighbor is Opaque capable (see Section 4 and Appendix A of [RFC2370])

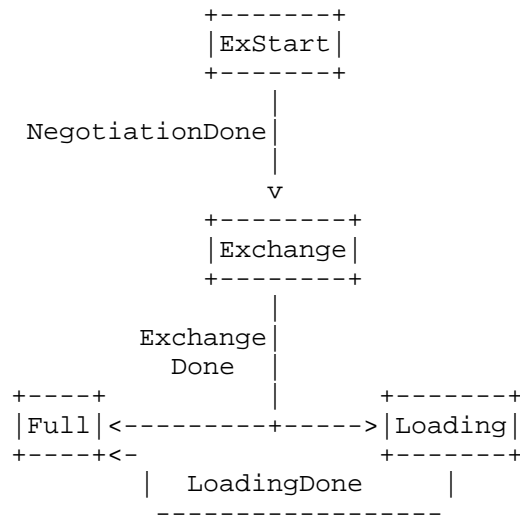


Fig.1 Neighbor state changes (Database Exchange)

ii) The neighbor has fully exchanged the area LSDB that consists of the router-LSAs (Type 1), network-LSAs (Type 2), and summary-LSAs (Type 3, 4) contained in the area structure, along with the AS-external-LSAs (Type 5) contained in the global structure, and Not-So-Stubby Area (NSSA) LSAs (Type 7) [RFC3101], i.e., the "Full" state has been reached.

iii) Both local and neighbor router supports the phased LSDB synchronization (see Section 4.3).

4.1. Phased Link-State Database (LSDB) Synchronization Process

The process is depicted in Fig.2, the ExStart, Exchange, Loading and Full states are defined per [RFC2328]. Note that in Full State, the router can perform all subsequent operations per [RFC 2328] including, the computation of the shortest-path tree for an area, and the computation of the AS external routes, as described in Section 16 of [RFC2328]. Events NegotiationDone, ExchangeDone and LoadingDone are used as defined per [RFC2328].

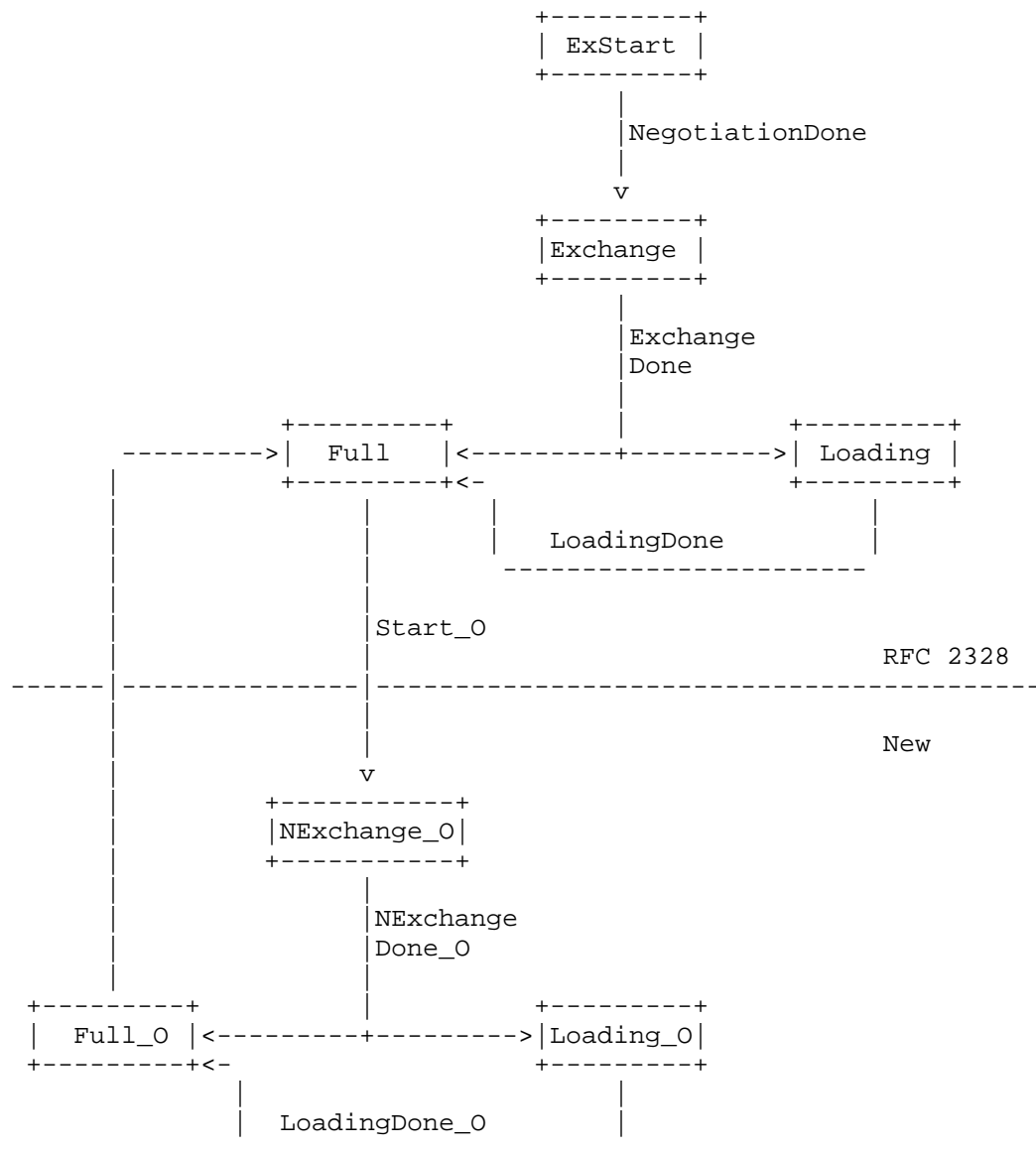


Fig.2 Modified Neighbor state changes (Database Exchange)

4.2. Transition from LSDB to Opaque LSDB Synchronization

In case Full state is not reached due to e.g. corruption or Fletcher checksum error, the exchange process restarts (go back to ExStart).

In case Full state is reached, the process continues as follows (note that the master remains the master as negotiated during the ExStart step):

- Start_O (Event): LSDB contain Opaque_LSA's AND capability as described in Section 4.3 successfully negotiated. This ensures backward compatibility with [RFC2370].

- NExchange_O (State): the router lists the contents of its Opaque area LSDB in the neighbor Database summary list. The Opaque area LSDB consists of Type 9 and Type 10 Opaque LSAs along with Type 11 Opaque LSAs contained in the global structure. The router sends the Database Description (DD) packets for these Opaque LSAs to the neighbor. Each DD Packet has a DD sequence number, and is explicitly acknowledged. Only one DD Packet is allowed outstanding at any one time. In this state, LS Request Packets may also be sent asking for the neighbor's more recent Opaque LSAs.

- NExchangeDone_O (Event): both routers have successfully transmitted a full sequence of DD packets. Each router now knows what parts of its Opaque LSDB are out of date.

- Loading_O (State): LS Request packets are sent to the neighbor asking for the more recent Opaque LSAs that have been discovered (but not yet received) in the NExchange_O state.

- LoadingDone_O (Event): LS Updates have been received for all out-of-date portions of the Opaque LSDB. This is indicated by the Link state request list becoming empty after the Database Exchange process has completed.

- Full_O (State): neighboring nodes have completed Opaque LSA exchange.

4.3. Capability Negotiation

Negotiating Phased LSDB synchronization can be performed by inserting a Phased LSDB Flag in:

i) Option field of Hello Packets and DD Packets

ii) Data block of Link Local Signaling (LLS) and DD Packets

4.3.1. Option field in Hello Packets

The Options field (8-bits) enables OSPF routers to support (or not support) optional capabilities, and to communicate their capability level to other OSPF routers. Through this mechanism routers of differing capabilities can be mixed within an OSPF routing domain. When used in Hello packets, the Options field allows a router to accept a neighbor at the condition of adaptation to neighbors's capability (due to the initial mismatch): if either of the neighbors does not support or does not recognize the capability the synchronization is not phased. In this case, routers encountering the unrecognized Option bits in received Hello Packets ignore the capability and process the packet normally.

Then, by exchanging this capability in Database Description (DD) packets a router can sequence its exchange of LSAs (starting by non-Opaque LSAs and then Opaque LSAs): if both neighbors are capable of phased synchronization, they may still decide to use or not.

This alternative is perfectly valid but requires usage of one bit of the Option field that is a very sparse resource.

4.3.2. Link Local Signaling (LLS)

To Link-local signaling (LLS), OSPF routers add a special data block to the end of OSPF packets (or right after the authentication data block when cryptographic authentication is used). The LLS block is attached to OSPF Hello packets. The drawback of this alternative is that the delivery of LLS data in Hello packets is not guaranteed.

To circumvent this problem, the solution consists in piggy bagging the Phased DB Flag in the Database Description packets.

5. Backward Compatibility

The proposed synchronization process is backward compatible since the mechanism extends the current process if and only if the mechanism is locally (see Section 4.2) and remotely supported (see Section 4.3). If either of these conditions is not met LSDB synchronization falls back to the linear process currently

specific per [RFC2370]. Note also that the proposed mechanism does not modify the LSDB process as specified in [RFC2328]. However, it does not prevent that routers may be required to support both methods. This could be the case typically for ABR's.

6. Security Considerations

TBD

7. IANA Considerations

TBD

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.
- [RFC2234] Crocker, D. and Overell, P.(Editors), "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, Internet Mail Consortium and Demon Internet Ltd., November 1997.
- [RFC2328] J. Moy, "OSPF version 2", RFC 2328, Internet Engineering Task Force, April 1998.
- [RFC2370] R. Coltun, "The OSPF Opaque LSA Option", RFC 2370, July 1998.

8.2. Informative References

- [OSPF-TP] A.Lindem, et al. "OSPF Transport Instance Extensions", IETF Draft, Work in Progress, draft-ietf-ospf-transport-instance-04.txt, April 2010.
- [RFC3101] P. Murphy, "The OSPF Not-So-Stubby Area (NSSA) Option", RFC 3101, Internet Engineering Task Force, January 2003.

9. Acknowledgments

Authors would like to thank Marc Lasserre, Devendra Raut, Andrew Lange, and Manav Bhatia for their comments.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Dimitri Papadimitriou
Alcatel-Lucent Bell
Copernicuslaan 50
2018 Antwerpen
Belgium
Phone: +32 3 2408491
Email: dimitri.papadimitriou@alcatel-lucent.com

Open Shortest Path First IGP
Working Group
Internet-Draft
Updates: 2328 (if approved)
Intended status: Standards Track
Expires: April 16, 2011

P. Jakma
DCS, Uni. of Glasgow
M. Bhatia
Alcatel-Lucent
October 13, 2010

Stronger, Automatic Integrity Checks for OSPF Packets
draft-jakma-ospf-integrity-00

Abstract

This document describes an extension to OSPFv2 and OSPFv3 to allow a stronger integrity check to be applied to the protocol packets, than the default OSPF checksum, which is known to be weak.

The extension allows OSPF speakers to negotiate the use of a CRC integrity check, as a new psuedo-authentication type.

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 16, 2011.

Copyright Notice

Copyright (c) 2010 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. Requirements Language	3
2. Introduction	3
3. Stronger Checksum mechanism for OSPFv2	3
3.1. Null Authentication Data	4
4. Stronger Checksum mechanism for OSPFv3	4
4.1. EC-Bit in Options Field	4
4.2. Extended Checksum Data Block	5
5. Generation	5
6. Verification	6
7. Stronger Integrity Algorithm Types	7
7.1. CRC32	7
7.2. MD5-Digest	7
8. IANA Considerations	7
9. Security Considerations	7
10. References	8
10.1. Normative References	8
10.2. Informative References	8
Authors' Addresses	8

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 [RFC2119].

2. Introduction

The integrity of Open Shortest Path First versions 2 (OSPFv2)[RFC2328] and 3 (OSPFv3)[RFC5340] packets is protected either through the standard internet protocol checksum, or through some cryptographic integrity scheme within OSPF, or, more rarely, through IPSec. This provides a check against errors that can not be caught by the link-layer integrity checks, e.g. errors in lower layers of the software stack or in hardware of the host.

The internet protocol checksum is known to have weaknesses[partridge]. In particular it can not detect re-ordered words and certain patterns of bit flips. If stronger integrity checks are desired, the only option is to use cryptographic HMACs, either with MD5 (all conforming [RFC2328] implementations) or, if supported, the stronger algorithms specified by [RFC5709]. There are some disadvantages though to using the existing support for cryptographic HMACs purely for integrity checking. The algorithms require more computation, which may be noticable on less powerful and/or energy-sensitive platforms. Additionally, the need to configure key material is an additional administrative burden.

This documents extends OSPF to allow for the automatic and backward compatible use of stronger integrity checks. Backward compatibility implies the default null authentication type must be used and extended.

3. Stronger Checksum mechanism for OSPFv2

The null authentication mode of OSPFv2 is extended to make use of the authentication data field of the OSPFv2 packet header. Where previously this field was ignored for null authentication, now an OPTIONAL "Null Authentication Data" structure is recognised there.

Implementations MUST provide a means to disable this extension, in case there are non-conforming RFC2328 implementations. Implementations MAY wish to generate a CRC32 checksum by default via this extension, and SHOULD attempt to verify any received, regardless of whether they generate the same or not.

3.1. Null Authentication Data

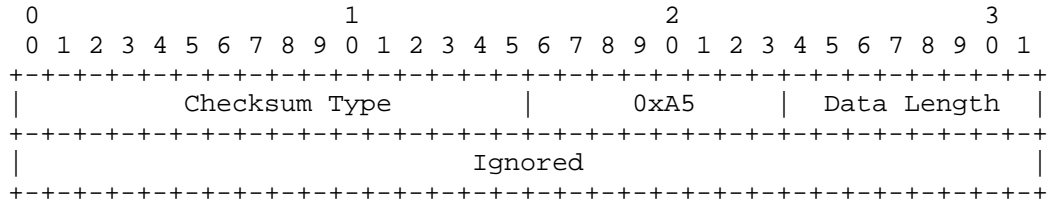


Figure 1: Null Authentication Data

The authentication data field in the standard OSPFv2 packet header is redefined as shown above, when null authentication is used. The new field definitions are as follows:

Checksum Type:

This field indicates the new checksum algorithm that the routers must use and is described in detail in the later sections.

Magic:

This field is set to 0xA5. This magic, in combination with the OSPF and IP packet lengths, signals the use of this extension.

Data Length:

The length in 4-octet words of the extended checksum data block appended to the OSPFv2 packet.

4. Stronger Checksum mechanism for OSPFv3

OSPFv3 uses IPSec for protection and does not carry any authentication information in its headers. Thus it is not possible to overload the Null Authentication type as was done in case of OSPFv2.

4.1. EC-Bit in Options Field

A new EC-bit (EC stands for Extended Checksum) is introduced into the OSPFv3 Options field. Routers MUST set the EC-bit in all OSPFv3 packets to indicate that the packet is carrying the new extended checksum data.

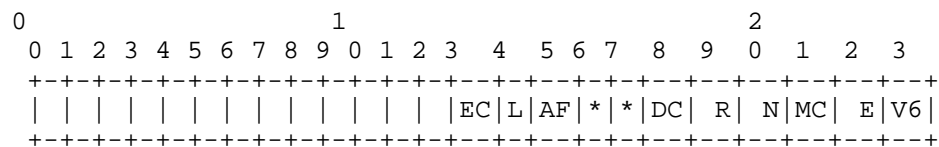


Figure 2: OSPFv3 Options Field

4.2. Extended Checksum Data Block

The data block for carrying extended checksum in OSPFv3 is formatted as described below.

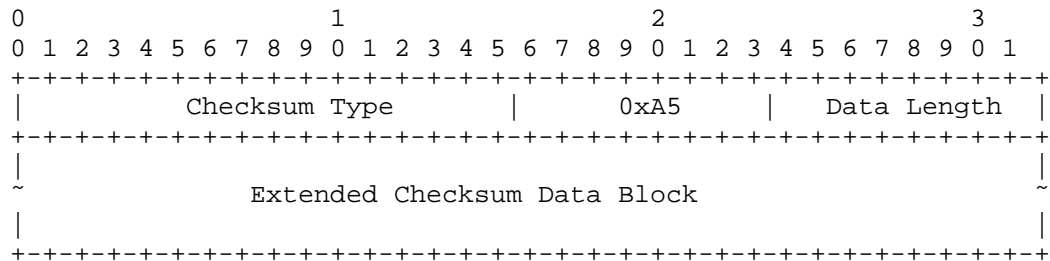


Figure 3: OSPFv3 Options Field

The Checksum Type is of two octets and indicates the new checksum algorithm that the routers must use. This is described in detail in the later sections. The next field is a reserved magic field set to 0xA5. The Data length field is of two octets and carries the size of the entire extended checksum data block that has been appended to the OSPFv3 payload, specified in units of 4-octet words. The Extended Checksum Data Block carries the checksum data that the receivers will use to verify the integrity of the OSPFv3 protocol payload.

5. Generation

The same steps are followed as for D.4.1 of [RFC2328]. Additionally, a 2nd integrity check algorithm is also computed over the packet data, with at least the same amount of zero padding, to produce an "extended checksum", which is appended to the OSPFv2 packet. Its size is accounted for in the Null Authentication Data "data length" field and in the IP length, but not in the OSPFv2 packet header, in a similar fashion to the standard OSPFv2 cryptographic authentication mechanism.

The "Checksum Type" and "Data Length" fields are set to the appropriate values for the 2nd integrity check algorithm.

In case of OSPFv3 the entire extended checksum block is appended to the OSPFv3 packet, with its size accounted for in the IPv6 payload length, but not in the OSPFv3 packet header.

Implementations MUST append the extended checksum data, that is carried as part of the OSPF protocol payload, before the link local signaling (LLS) [RFC5613] block (if it exists).

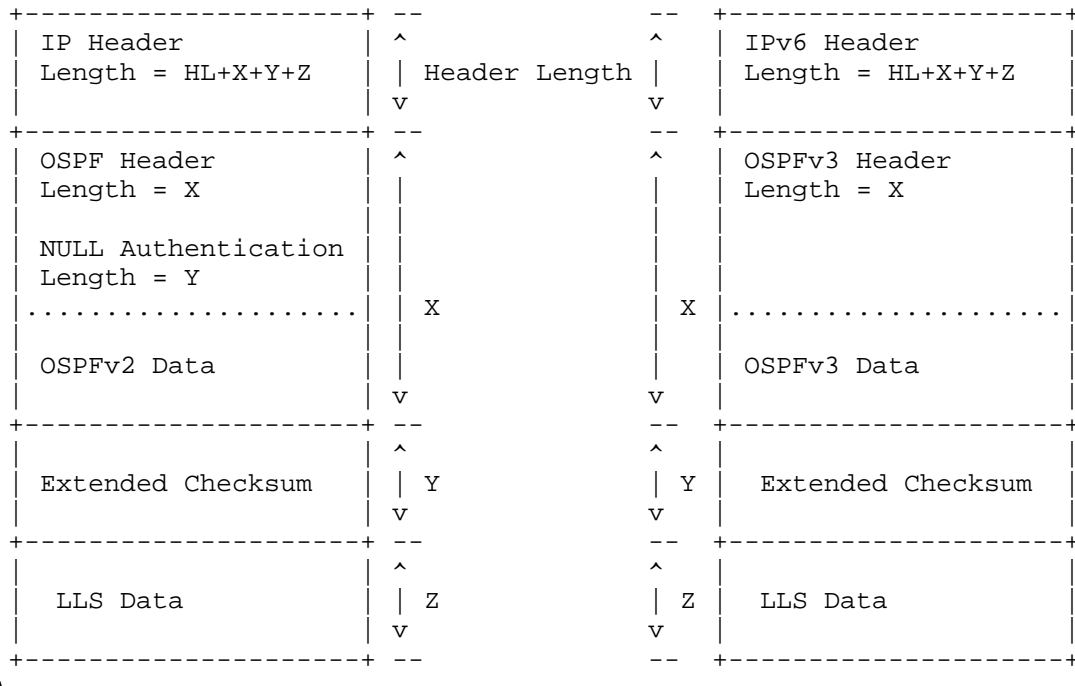


Figure 4: Extended Checksum Block in OSPFv2 and OSPFv3

6. Verification

The packet data is padded out, as required by [RFC2328].

In case of OSPFv2, the Null Authentication Data "0xA5" magic field is examined. If it does not match, then verification proceeds as per D.5.1 of [RFC2328]. If it matches, then the IP length in the header MUST be verified. An incoming packet will only contain a valid extended checksum if the length in the IP header = length in OSPF header + "data length" in the NULL Authentication header + data length in the LLS [RFC5613] block (if it exists). Implementations can trivially determine if an LLS block is being carried by inspecting the "L" bit in the OSPF Options field in the HELLOs and DDs. Implementations MUST proceed with regular checksum if these numbers don't match. If they do then the IP checksum field of the OSPF header MUST be ignored. Instead the stronger integrity

algorithm specified by the "Checksum Type" field is used, and verified against the corresponding checksum. The packet MUST be discarded if the computed checksum does not match with what's carried in the OSPF packet.

In case of OSPFv3, the presence of the EC-bit in the OSPFv3 Options field will indicate that a new checksum algorithm is being used. Routers MUST parse the packet till the end of the OSPFv3 payload till it reaches the start of the extended checksum data block. The processing that follows next is similar to the way its done for OSPFv2 as explained earlier.

7. Stronger Integrity Algorithm Types

7.1. CRC32

The CRC32 algorithm, as used with IEEE 802.3 and defined by [hammond] is used to calculate its 4-byte digest. The length set in the Null Authentication Data thus will be 1.

7.2. MD5-Digest

The MD5 algorithm, as per 5ref17 of [RFC2328] is used in plain digest mode (i.e. solely over the data, unlike the HMAC mode used by cryptographic authentication) to calculate its 8-byte digest. The length set in the Null Authentication Data thus will be 2.

8. IANA Considerations

OSPFv2 Null Authentication Checksum Types are maintained by the IANA. Extensions to OSPFv2 that require a new Checksum Type must be reviewed by a designated expert from the routing area.

This document assigns OSPF Null Authentication Checksum Types 1 and 2, for CRC32 and MD5-Digest respectively.

IANA is also requested to allocate EC-bit in the OSPFv3 "Options Registry"

9. Security Considerations

This extension does not raise any new security concerns. It only is used where operators have chosen not to configure cryptographic security mechanisms.

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.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [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.
- [hammond] Hammond, J., Brown, J., and S. Lui, "Development of a Transmission Error Model and an Error Control Model", Technical Report Georgia Institute of Technology, May 1975, <<http://handle.dtic.mil/100.2/ADA013939>>.

10.2. Informative References

- [RFC5613] Zinin, A., Roy, A., Nguyen, L., Friedman, B., and D. Yeung, "OSPF Link-Local Signaling", RFC 5613, August 2009.
- [partridge] Stone, J., Greenwald, M., Partridge, C., and J. Hughes, "Performance of checksums and CRC's over real data", IEEE/ACM Trans. Netw. vol 6, num 5, pages 529-543, 1998, <<http://dx.doi.org/10.1109/90.731187>>.

Authors' Addresses

Paul Jakma
School of Computing Science, University of Glasgow
Lilybank Gardens
Glasgow G12 8QQ
Scotland

Email: paulj@dcs.gla.ac.uk

Manav Bhatia
Alcatel-Lucent
Bangalore,
India

Phone:
Email: manav.bhatia@alcatel-lucent.com

OSPF Working Group
Internet-Draft
Intended Status: Standards Track
Expires: April 2011

S. Kini
W. Lu
A. Tian
Ericsson
October 18, 2010

OSPF Fast Notifications
draft-kini-ospf-fast-notification-00.txt

Status of this Memo

Distribution of this memo is unlimited.

This Internet-Draft is submitted to IETF 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 April 21, 2011.

Copyright Notice

Copyright (c) 2010 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

Several applications could use a mechanism to quickly notify one or more routers about control-protocol events. Current mechanisms to convey such information to routers multiple hops away involves hop-by-hop protocol-specific control plane processing as well as hop-by-hop control plane forwarding. The delay due to control planes involvement in processing/forwarding, adversely affects the application's goal (e.g. fast convergence). This document describes a framework to use data plane forwarding to convey control protocol information multiple hops away. It also defines some sample applications within this framework.

Table of Contents

1. Introduction	4
2. Conventions used in this document	5
3. Scope	5
4. Requirements	5
5. Architecture	5
5. Security Considerations	6
6. IANA Considerations	6
7. References	6
7.1. Normative References	6
8. Acknowledgements	6
Appendix A: OSPF fast convergence on link-down using FN	7
A.1. OSPF procedural changes	7
A.2. FN service using spanning tree	7
Authors' Addresses	9

1. Introduction

There are several applications that could use a mechanism to quickly notify one or more routers in a network about a specific control-protocol event. If the destination router(s) is more than one hop away then the message has to be forwarded by the intermediate routers. This forwarding typically does not exclusively happen via the forwarding plane.

Some applications establish adjacent neighbor relationship with single hop neighbors. Information that needs to be conveyed multiple hops away is first conveyed to adjacent neighbors that are a single hop away. Each neighbor then performs application specific processing and forwards information further. The delay in receiving the information at a router is gated by the processing and forwarding speed of the control plane at each hop along a path from the originating router.

A typical example of an application that sends information to directly connected adjacent neighbors is a link-state routing interior gateway protocol (IGP) such as [OSPF]. When conveying a Link State Advertisement (LSA) to all routers in the area, OSPF's flooding algorithm transmits the LSA to its single hop away adjacent neighbor. The received LSA undergoes processing according to OSPF's processing rules and is then forwarded to OSPF neighbors further away from the router originating the LSA. As explained earlier the delay in receiving a LSA at a router is gated by the processing and forwarding speed of the control plane at each hop along a path from the originating OSPF router.

Some applications need to send information to routers that are multiple hops away even though they do not have adjacency relationship with directly connected neighbors. In such cases the forwarding of application messages depends on the forwarding plane being setup by an underlying protocol that has established adjacent neighbor relationship with routers that are a single hop away. In scenarios where the data plane forwarding is changing due to the underlying protocol, the applications message forwarding speed and reliability is gated by the speed and mechanisms of the underlying protocols hop-by-hop message processing and forwarding by control-plane.

A typical example of an application that could use a mechanism to send information to non-directly connected neighbors is IP FastReroute (IP-FRR). It could use a forwarding mechanism that has been setup by an underlying protocol to trigger (on failure) a non-directly connected neighbor, to switch traffic to an alternate path. To reliably deliver the applications message, the forwarding

mechanism has to be resilient against failures and the changed topology.

2. Conventions used in this document

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. Scope

This document describes a framework for quickly delivering notifications from one router to one or more routers using data plane as the main forwarding mechanism. It also defines some solutions under this framework to address the needs of some specific applications.

4. Requirements

Fast notifications (henceforth referred to as FN) must be designed as a set of services that can satisfy the requirements of different control-protocol applications. FN should avoid introducing new protocols and should re-use existing, commonly used protocols as much as possible.

Deploying FN must not introduce new encapsulation requirements for routers unless those encapsulations are already available in the data plane for those applications. Notifying multiple routers should use multicast whenever possible.

5. Architecture

A choice of protocol to realize FN must be based on the set of commonly deployed protocols. These protocols must preferably have applicability in a wide set of network architectures such as IP-routing, L3VPN, L2VPN etc. Also, the knowledge of the network topology would be particularly useful for path computation purposes. The logical candidate for these requirements would be a link-state interior gateway protocol (IGP) such as OSPF or IS-IS.

To implement a specific FN service, a router must convey its capability to the set of routers that setup forwarding to one or more routers in that set for specific packets in a way such that data-plane forwarding of notifications. It must also convey its share of the information that is needed to implement that FN service.

To convey this information via OSPF, an opaque LSA is used. An Opaque type field "FN" is defined. The type specific ID indicates a

particular FN service. The content of the LSA is a variable list of TLVs that include information required to implement that FN service. Different FN services will have different sets of TLVs. A specific instance of a FN service and how an application might use it is specified in Appendix A.

5. Security Considerations

Security considerations of the application also apply when FN service is used by the application. If additional security considerations arise due to the way in which FN is used by the application, then those should be resolved in the document that explains how an application uses FN.

6. IANA Considerations

IANA needs to allocate a OSPF opaque type field for FN. Within that LSID values for different FN services will have to be allocated. Also a TLV type field will have to be allocated.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [OSPF] Moy, J., "OSPF Version 2", RFC 2328, April 1998.

8. Acknowledgements

The authors would like to thank Joel Halpern for his comments.

Appendix A: OSPF fast convergence on link-down using FN

OSPF fast convergence is gated by how quickly the flooding algorithm can propagate the LSA throughout the area. This requires hop-by-hop processing and forwarding by control plane. If a FN service can transmit the link-down notification to all routers in the area then OSPF's fast convergence can be improved in the link-down scenario.

A.1. OSPF procedural changes OSPF's procedures must be modified to use the FN service as follows. OSPF transmits a copy of the updated Router LSA (on link-down) using a FN service in addition to the normal processing and flooding done by OSPF. The destination IP address of the link-state update (LSU) packet is set to the one dictated by the FN service. If Cryptographic authentication is required, a shared secret key must be configured for the area. The Cryptographic sequence number in the LSU must be set to zero. On receiving a LSU via FN, the router accepts it if authentication succeeds. There must be no acknowledgement for such an LSU. If the received LSA is older than the one in the LSDB, the received LSA is discarded. If the received LSA is newer, the LSA is stored alongside the older copy and a timer T-discard-FN-LSA is started. A flag FN-LSA-present is used in the LSDB entry to indicate that a newer version of the LSA (received via FN) is present. SPF is triggered. During SPF, if the FN-LSA-present flag is true then the LSA received via FN is used instead. When a LSA is received via the flooding procedure of [OSPF], and is determined to be newer, it is compared with the LSA copy received via FN (if one exists). If the two copies are the same, the LSA received via FN becomes the only entry in the LSDB. If the two copies are different, the LSA received through the flooding procedure of [OSPF] becomes the only copy in the LSDB and SPF is triggered. In both cases the flag FN-LSA-present is cleared and the timer T-discard-FN-LSA is canceled. When the timer T-discard-FN-LSA expires, the corresponding LSA copy received via FN is discarded (FN-LSA-present flag is cleared) and SPF is triggered.

A.2. FN service using spanning tree

One way to provide the FN service for this application is as follows. A multicast spanning tree (with a specially allocated multicast destination IP address) is used to send the link-down notification message. The tree must be consistently computed at all routers. It must be computed as a shortest path tree rooted at the highest router-id. During tree computation only routers that are capable of this FN service are picked. When multiple paths are available the neighboring node in the graph with highest LSID is picked. When multiple paths are available through multiple interfaces to a neighboring node, a numbered interface is preferred over an

unnumbered interface. A higher IP address is preferred among numbered interfaces and a higher ifIndex is preferred among unnumbered interfaces. Multicast forwarding state is installed using such a tree as a bi-directional tree. Each router on the tree can send packets to all other routers on that tree. Even when the topology changes such that the tree breaks, the link-down notification is delivered to all routers.

Authors' Addresses

Sriganesh Kini
Ericsson
300 Holger Way, San Jose, CA 95134
EMail: sriganesh.kini@ericsson.com

Wenhu Lu
Ericsson
300 Holger Way, San Jose, CA 95134
EMail: wenhu.lu@ericsson.com

Albert Tian
Ericsson
300 Holger Way, San Jose, CA 95134
EMail: albert.tian@ericsson.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 25, 2011

N. Sheth
L. Wang
J. Zhang
Juniper Networks
October 25, 2010

OSPF Hybrid Broadcast and P2MP Interface Type
draft-nsheth-ospf-hybrid-bcast-and-p2mp-01.txt

Abstract

This document describes a mechanism to model a broadcast network as a hybrid of broadcast and point-to-multipoint networks for purposes of OSPF operation. Neighbor discovery and maintenance as well as LSA database synchronization are performed using the broadcast model, but the network is represented using the point-to-multipoint model in the router LSAs of the routers connected to it. This allows an accurate representation of the cost of communication between different routers on the network, while maintaining the network efficiency of broadcast operation. This approach is relatively simple and requires minimal changes to OSPF.

Status of this Memo

This Internet-Draft is submitted to IETF 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 April 25, 2011.

Copyright Notice

Copyright (c) 2010 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 BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	3
2. Motivation	4
3. Operation	5
3.1. Interface Parameters	5
3.2. Neighbor Data Structure	5
3.3. Neighbor Discovery and Maintenance	5
3.4. Database Synchronization	5
3.5. Generating Network LSAs	5
3.6. Generating Router and Intra-Area-Prefix-LSAs	5
3.6.1. Stub Links in OSPFv2 Router LSA	6
3.6.2. OSPFv3 Intra-Area-Prefix-LSA	6
3.7. Next-Hop Calculation	6
3.8. Graceful Restart	6
4. Compatibility Considerations	8
5. Scalability and Deployment Considerations	9
6. Security Considerations	10
7. IANA Considerations	11
8. Normative References	12
Authors' Addresses	13

1. Introduction

OSPF [RFC2328] operation on broadcast interfaces takes advantage of the broadcast capabilities of the underlying medium for doing neighbor discovery and maintenance. Further, it uses a Designated Router and Backup Designated Router to keep the LSA databases of the routers on the network synchronized in an efficient manner. However, it has the limitation that a router cannot advertise different costs to each of the neighboring routers on the network in its router LSA.

Operation on point-to-multipoint interfaces could require explicit configuration of the identity of its neighboring routers. It also requires the router to send separate hellos to each neighbor on the network. Further, it mandates establishment of adjacencies to all all configured or discovered neighbors on the network. However, it gives the routers the flexibility to advertise different costs to each of the neighboring routers in their router LSAs.

This document proposes a new interface type that can be used on layer 2 networks that have broadcast capability. In this mode, neighbor discovery and maintenance, as well as database synchronization are performed using existing procedures for broadcast mode. The network is modeled as a collection of point-to-point links in the router LSA, just as it would be in point-to-multipoint mode. This new interface type is referred to as hybrid-broadcast-and-p2mp in the rest of this document.

2. Motivation

There are some layer 2 networks that are broadcast capable but have a potentially different cost associated with communication between any given pair of nodes. The cost could be based on the underlying layer 2 topology as well as various link quality metrics such as bandwidth, delay and jitter among others.

It is not accurate to treat such networks as OSPF broadcast networks since that does not allow a router to advertise a different cost to each of the other routers. Using OSPF point-to-multipoint mode would satisfy the requirement to correctly describe the cost to reach each router. However, it would be inefficient in the sense that it would require forming $O(N^2)$ adjacencies when there are N routers on the network.

It is advantageous to use the hybrid-broadcast-and-p2mp type for such networks. This combines the flexibility of point-to-multipoint type with the advantages and efficiencies of broadcast interface type.

3. Operation

OSPF routers supporting the capabilities described herein should have support for an additional hybrid-broadcast-and-p2mp type for the Type data item described in section 9 of [RFC2328].

The following sub-sections describe salient aspects of OSPF operation on routers configured with a hybrid-broadcast-and-p2mp interface.

3.1. Interface Parameters

Routers MUST support configuration of the Router Priority for the interface.

The default value of the LinkLSASuppression is "disabled". It MAY be set to "enabled" via configuration.

3.2. Neighbor Data Structure

Routers MUST support an additional field called the Neighbor Output Cost. This is the cost of sending a data packet to the neighbor, expressed in the link state metric. The default value of this field is the Interface output cost. It MAY be set to a different value using mechanisms which are outside the scope of this document, like static per-neighbor configuration, or any dynamic discovery mechanism that is supported by the underlying network.

3.3. Neighbor Discovery and Maintenance

Routers send and receive Hellos so as to perform neighbor discovery and maintenance on the interface using the procedures specified for broadcast interfaces in [RFC2328] and [RFC5340].

3.4. Database Synchronization

Routers elect a DR and BDR for the interface and use them for initial and ongoing database synchronization using the procedures specified for broadcast interfaces in [RFC2328] and [RFC5340].

3.5. Generating Network LSAs

Since a hybrid-broadcast-and-p2mp interface is described in router LSAs using a collection of point-to-point links, the DR SHOULD NOT generate a network LSA for the interface.

3.6. Generating Router and Intra-Area-Prefix-LSAs

Routers describe the interface in their router LSA as specified for a point-to-multipoint interface in section 12.4.1.4 of [RFC2328] and section 4.4.3.2 of [RFC5340], with the following modifications for

Type 1 links:

- o If a router is not the DR, it MUST NOT add any Type 1 links if it does not have a full adjacency to the DR.
- o If a router is not the DR and has a full adjacency to the DR, it MUST add a Type 1 link corresponding to each neighbor that is in state 2-Way or higher.
- o The cost for a Type 1 link corresponding to a neighbor SHOULD be set to the value of the Neighbor Output Cost field as defined in Section 3.2

3.6.1. Stub Links in OSPFv2 Router LSA

Routers MUST add a Type 3 link for their own IP address to the router LSA as described in section 12.4.1.4 of [RFC2328]. Further, they MUST also add a Type 3 link with the Link ID set to the IP subnet address, Link Data set to the IP subnet mask, and cost equal to the configured output cost of the interface.

3.6.2. OSPFv3 Intra-Area-Prefix-LSA

Routers MUST add global scoped IPv6 addresses on the interface to the intra-area-prefix-LSA as described for point-to-multipoint interfaces in section 4.4.3.9 of [RFC5340]. In addition, they MUST also add all global scoped IPv6 prefixes on the interface to the LSA by specifying the PrefixLength, PrefixOptions, and Address Prefix fields. The Metric field for each of these prefixes is set to the configured output cost of the interface.

The DR SHOULD NOT generate an intra-area-prefix-LSA for the transit network for this interface since it does not generate a network LSA for the interface. Note that the global prefixes associated with the interface are advertised in the intra-area-prefix-LSA for the router as described above.

3.7. Next-Hop Calculation

Next-Hops to destinations that are directly connected to a router via the interface are calculated as specified for a point-to-multipoint interface in section 16.1.1 of [RFC2328].

3.8. Graceful Restart

The following modifications to the procedures defined in section 2.2, item 1 of [RFC3623] are required in order to ensure that the router correctly exits graceful restart.

- o If a router is the DR on the interface, it MUST NOT examine the pre-restart network LSA for the interface in order to determine the previous set of adjacencies.
- o If a router is in state DROther on the interface, it MUST consider an adjacency to non-DR and non-BDR neighbors as reestablished when the neighbor state reaches 2-Way.

4. Compatibility Considerations

All routers on the network must support the hybrid-broadcast-and-p2mp interface type for successful operation. Otherwise, the interface should be configured as a standard broadcast interface.

If some routers on the network treat the interface as broadcast and others as hybrid-broadcast-and-p2mp, neighbors and adjacencies will still get formed as for a broadcast interface. However, due to the differences in how router and network LSAs are built for these two interface types, there will be no traffic traversing certain pairs of routers. Note that this will not cause any persistent loops or black holing of traffic.

5. Scalability and Deployment Considerations

Treating a broadcast interface as hybrid-broadcast-and-p2mp results in $O(N^2)$ links to represent the network instead of $O(N)$, when there are N routers on the network. This will increase memory usage and have a negative impact on route calculation performance on all the routers in the area. Network designers should carefully weigh the benefits of using the new interface type against the disadvantages mentioned here.

6. Security Considerations

This document raises no new security issues for OSPF. Security considerations for the base OSPF protocol are covered in [RFC2328] and [RFC5340].

7. IANA Considerations

This document has no IANA considerations.

This section should be removed by the RFC Editor to final publication.

8. 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.
- [RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.
- [RFC3623] Moy, J., Pillay-Esnault, P., and A. Lindem, "Graceful OSPF Restart", RFC 3623, November 2003.

Authors' Addresses

Nischal Sheth
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: nsheth@juniper.net

Lili Wang
Juniper Networks
10 Technology Park Dr.
Westford, MA 01886
US

Email: lililiw@juniper.net

Jeffrey Zhang
Juniper Networks
10 Technology Park Dr.
Westford, MA 01886
US

Email: zzhang@juniper.net

Network Working Group
Internet Draft
Intended status: Standard Track
Expires: April 2011

Pierre Peloso
Alcatel-Lucent
Julien Meuric
France Telecom
Giovanni Martinelli
Cisco

October 25, 2010

OSPF-TE Extensions for WSON-specific Network Element Constraints

draft-peloso-ccamp-wson-ospf-oeo-02.txt

Status of this Memo

This Internet-Draft is submitted to IETF 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 April 25, 2011.

Copyright Notice

Copyright (c) 2010 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 describes OSPF routing protocols extensions to support blocking nodes and O-E-O pools in all-optical networks under the control of Generalized MPLS (GMPLS).

Conventions used in this document

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].

Table of Contents

1. Introduction.....	2
2. Resource Block Attribute.....	3
2.1. Pool ID.....	5
2.2. Block Shared Access Wavelength Availability.....	5
2.3. Resource Element Information.....	5
2.4. Relation with Node.....	5
3. Security Considerations.....	6
4. IANA Considerations.....	6
4.1. Resource Block attributes.....	6
5. References.....	6
6. Author's Addresses.....	8
Intellectual Property Statement.....	8
Disclaimer of Validity.....	9

1. Introduction

The goal of all-optical meshed networks consists in the transport of optical circuit connections, with limited usage of Optical-Electrical-Optical conversion through photonic nodes. The gain brought by the use of fewer regenerators is balanced by the constraint of maintaining the optical signal continuity between the source and the destination nodes. In GMPLS controlled networks, the induced signal continuity brings the technological challenge of wavelength assignment using control plane protocols, which is discussed in [WSON-Frame].

The drawback of wavelength assignment computation in a single entity is the need to gather and convey all relevant and up-to-date information to this single entity. Whether the computing entity takes the form of a PCE or the form of a Constrained-Shortest-Path-First (C-SPF) engine in each node of the network, the IGP is supposed to do the job of gathering this information.

This document defines extensions to the OSPF routing protocol based on [WSON-Encode] to enhance the Traffic Engineering (TE) properties of GMPLS TE which are defined in [RFC3630], [RFC4202], and [RFC4203]. The enhancements to the Traffic Engineering (TE) properties of GMPLS TE links can be announced in OSPF TE LSAs. The TE LSA, which is an opaque LSA with area flooding scope [RFC3630], has only one top-level Type/Length/Value (TLV) triplet and has one or more nested sub-TLVs for extensibility. The top-level TLV can take one of three values (1) Router Address [RFC3630], (2) Link [RFC3630], (3) Node Attribute [RFC5786]. In this document, we introduce a new top-level TLV containing Resource Block Attribute (RBA).

[WSON-Encode] introduce the concept of RBA to include all information that are specific to WSON nodes. This information may introduce some additional constraints that needs to be considered to perform a correct RWA. This document does not define any additional encoding but maps information from [WSON-Info] and [WSON-Encode] on OSPF.

The detailed encoding of OSPF extensions are not defined in this document. [WSON-Encode] provides encoding detail.

2. Resource Block Attribute

This draft defines a new top-TLV named "Resource Block Attribute" TLV. It carries attributes related to a pool of Optical-Electric-Optical regeneration resource, thus allowing route computation to take into account available signal regenerators in the network.

Available OEO resource introduce different kind of constraints. One is the signal compatibility as defined in [WSON-Signal]. Another constraint comes from WSON node topologies (for technology reasons or cost of resources). This draft mainly refers to the latter.

Multiple O-E-O resources are logically gathered in a pool when they share a common transmission media before (and after) entering (exiting) the actual switching matrix of the node. A common transmission media is characterized by the sharing of at least a short section of fiber: hence an amplifier or a wavelength selective switch does also correspond to a common transmission media.

When several regenerators' pools are available on a node, several "Resource Block Attribute" will be used (one for each pool). As a matter of fact, the split into blocks of the O-E-O resources comes from the architectural structure of the node. This Node Attribute TLV contains two or more sub-TLVs.

The resource block attributes related to OEO pools in WSON nodes include Block ID, lists of available wavelengths on the ingress and egress side of the pool, and the features of the resources in the block. These pieces of information are described in this document and refer to . The Resource Block Attribute would also include some sub-TLVs identical to sub-TLVs of the TE-link top-TLV: TE-metric [rfc3630], Administrative Group [rfc3630], Link Local/Remote Identifiers [rfc4203], Shared-Risk Link Group [rfc4203].

The following new sub-TLVs are added to the "Resource Block Attribute" TLV. Detailed description for newly defined sub-TLVs is provided at the end of the section.

Sub-TLV Type	Length	Name
TBD	4 Bytes	Block ID
TBD	variable	Block Shared Access Wavelength Availability
TBD	fixed	Resource Element Information

In "Resource Block Attribute", the sub-TLV "Block Shared Access Wavelength Availability" and "Resource Block Information" are mandatory, the other sub-TLV listed above is optional.

The following sub-TLVs to the "Resource Block Attribute" TLV are identical to the ones defined respectively in [RFC3630] and [RFC4203], and being defined for the TE-link top-TLV. Detailed description for newly defined sub-TLV is provided at the end of the section.

Sub-TLV Type	Length	Name
TBD	4 Bytes	TE-metric [alike RFC3630]
TBD	4 Bytes	Administrative Group [alike RFC3630]
TBD	8 Bytes	Link Local/Remote Identifiers [alike RFC4203]
TBD	variable	Shared Risk Link Group [alike RFC4203]

In "Resource Block Attribute", the sub-TLV "Link Local/Remote Identifiers" is mandatory as it is needed to ensure the consistency with the Node Information described in [Gen-OSPF] and [Gen-Encode]. The other sub-TLVs listed above are optional.

2.1. Pool ID

This optional sub-TLV can be used to provide an identifier to the regenerator pool.

2.2. Block Shared Access Wavelength Availability

This block includes information from [WSON-Encode] section 4.4 "Block Shared Access Wavelength Availability". It is used to describe the wavelengths available on the shared fibers (ingress and egress sides) of the pool.

At every RWA process the OEO pool may or may-not be used. The status of the wavelength availability will change. The information is fairly dynamic.

2.3. Resource Element Information

This sub-TLV advertises information that describes the features of the resource elements inside the resource block itself. The features are the accepted bit-rates, modulation format, FEC formats, etc...

Actually this sub-TLV is replicated in a list of such sub-TLVs in order to depict all the resource elements available in the pool. The description of the encoding of this sub-TLV is available in [WSON-encode] section 5 (Hence needs a slight adaptation of TLV described in 5.1: Resource Block Information).

The features of a given element are fairly static as they refer to the characteristics of the device, which mean that the content of a given sub-TLV is static. On the other hand, the elements composing the list are subject to change, when a device is used, its corresponding sub-TLV will disappear from the list.

2.4. Relation with Node

Accessing resource block is also subject to switching constraints. These switching constraints can be both spatial and spectral.

In order to convey this information, the Connectivity Matrix sub-TLV shall depict the ports of the O-E-O pool, and referring their Link Local/Remote Identifiers sub-TLV as described in section 2.

Hence the number of ports described by the connectivity matrix is:

Ingress ports (CM): # incoming links (Node) + # O-E-O pools

Egress ports (CM): # outgoing links (Node) + # O-E-O pools

3. Security Considerations

This document does not introduce any further security issues other than those discussed in [RFC 3630], [RFC 4203].

4. IANA Considerations

[RFC3630] says that the top level Types in a TE LSA and Types for sub-TLVs for each top level Types must be assigned by Expert Review, and must be registered with IANA.

IANA is requested to allocate new Types for the sub-TLVs as defined in Sections 2, 3, 3.1, 3.2 and 3.3 as follows:

4.1. Resource Block attributes

This document introduces the "O-E-O Pool Attribute" top-TLV, value TBD with the following sub-TLVs:

Type	Name
TBD	Pool ID
TBD	Block Shared Access Wavelength Availability
TBD	Resource Element Information
TBD	TE-metric [alike RFC3630]
TBD	Administrative Group [alike RFC3630]
TBD	Link Local/Remote Identifiers [alike RFC4203]
TBD	Shared Risk Link Group [alike RFC4203]

5. References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC3630] Katz, D., Kompella, K., and Yeung, D., "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC4202] Kompella, K., Ed., and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4202, October 2005

[RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, October 2005.

[RFC3945] E. Mannie, Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, October 2004.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC5786] R. Aggarwal and K. Kompella, "Advertising a Router's Local Addresses in OSPF TE Extensions", RFC 5786, March 2010.

[WSON-Frame] G. Bernstein, Y. Lee, W. Imajuku, "Framework for GMPLS and PCE Control of Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-wson-framework-07.txt, October 2010.

[RWA-Info] Y. Lee, G. Bernstein, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Model for Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-info-06.txt, October 2010.

[Gen-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "General Network Element Constraint Encoding for GMPLS Controlled Networks", work in progress: draft-ietf-ccamp-general-ext-encode-00.txt.

[WSON-Encode] G. Bernstein, Y. Lee, D. Li, W. Imajuku, "Routing and Wavelength Assignment Information Encoding for Wavelength Switched Optical Networks", work in progress: draft-ietf-ccamp-rwa-wson-encode-06.txt, October 2010.

[Gen-OSPF] F. Zhang, Y. Lee, J. Han, G. Bernstein, "OSPF-TE Extensions for General Network Element Constraints", work in progress: draft-zhang-ccamp-general-constraints-ospf-ext-00.txt, September 2010.

6. Author's Addresses

Pierre Peloso
Alcatel-Lucent
Rte de Villejust
91620 Nozay, France

Phone: +33 130 702 662
Email: pierre.peloso@alcatel-lucent.com

Julien Meuric
France Telecom
2, av Pierre Marzin
22307 Lannion Cedex, France

Phone: +33 296 052 828
Email: julien.meuric@orange-ftgroup.com

Giovanni Martinelli
Cisco
Via Philips 12
20052 Monza, Italy

Phone: +39 039 2092044
Email: giomarti@cisco.com

Intellectual Property Statement

The IETF Trust takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in any IETF Document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights.

Copies of Intellectual Property disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement

any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: March 1, 2011

S. Tsuchiya, Ed.
G. Van de Velde
Cisco Systems
T. Yamagata
KDDI Corporation
August 28, 2010

OSPFv3 Stub Router Advertisement
draft-shishio-ospf-ospfv3-stub-01

Abstract

OSPFv3 accommodates for the possibility to indicate through the R-bit if a router is an active router and should be taken into consideration as a transit device. Another method available is the v6-bit indicating if a router or link should be excluded from IPv6 routing calculations.

A direct result is that OSPFv3 has "no transit capability" potentially based upon the setting of R-bit and V6-bit, unlike the stub OSPFv2 router functionality. This feature proposal has as purpose to re-introduce existing OSPFv2 stub router behavior into OSPFv3 to keep the operational service provider experience used to deploy, troubleshoot and be familiar with OSPFv2 stub routing.

OSPFv3 has similar metric field information field of all of LSAs, with exception of the Link-LSA, so RFC3137 method can be re-utilized in OSPFv3.

To drive consistency between OSPFv2 and OSPFv3, there should be next to supporting both R-bit and v6-bit be support for "max-metric".

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 March 1, 2011.

Copyright Notice

Copyright (c) 2010 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. Motivation	4
2. Requirements Language	4
3. OSPFv2 operation	4
3.1. Wait for BGP during booting	5
3.2. LDP synchronization	5
3.3. Configuration change	6
4. OSPFv3 operation	6
4.1. R-bit and v6-bit	6
4.2. OSPFv2 compatibility mode	6
5. Acknowledgements	7
6. IANA Considerations	7
7. Security Considerations	7
8. Normative References	7
Appendix A. Additional Stuff	7
Authors' Addresses	8

1. Motivation

OSPF Stub Router Advertisement [RFC3137] describes a set of situations when the Service provider has a desire to utilize this functionality.

- o The router is in a critical condition resulting in either a very high CPU load, or not enough memory to store all LSAs, or doesn't succeed to build the routing table
- o Graceful introduction or removal of a router to or from the network
- o Other (administrative or traffic engineering) reasons

Even if the network will be moved or migrated towards from IPv4 in combination with OSPFv2 towards IPv6 using OSPFv3 [OSPFv3] technology, it remains important that the operational behaviour remains similar between routing protocols.

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 RFC 2119 [RFC2119].

3. OSPFv2 operation

RFC3137 [RFC3137] describes in section 2 in detail the behavior of link cost metrics. i.e. Router X announces its Router LSA to the neighbor with costs of all non-stub links which are set to LSInfinity(0xFFFF), while stub links are announced with interface cost.

Often the operator will check interface metric of ospf database assuming he would like to confirm whether the router is announcing LSInfinity.

Many service provider operators are using OSPF stub router advertisement [RFC3137] for OSPFv2 [OSPFv2]. This feature is supported by majority of OSPFv2 implementations. Use cases are below;

3.1. Wait for BGP during booting

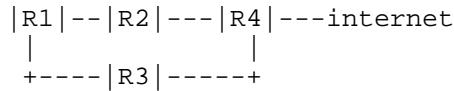


Figure 1

In this example R2 can be assumed it is the best path towards the Internet from R1. When R2 reloads it would result that R3 would be best path for going towards Internet. From the moment R2 is reloaded and OSPF has converged, there may be a situation when BGP is still not converged to the full. If in this situation R1 should not send traffic towards R2 just yet. R2 should send LSInfinity(0xFFFF) to indicate that R1 should wait for R2's BGP converge. Once BGP is fully converged, R2 LSA's out with correct interface metric value in OSPFv2 area which will result in R2 being reintroduced as the primary path.

3.2. LDP synchronization

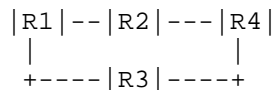


Figure 2

Assume that from R1 the best path to R4 is via R2 in this MPLS network. When R2 is reloaded, then R3 is the only and hence also the best path. At some point in time R2 is fully successfully reloaded resulting that OSPF has converged also. This does not necessary mean LDP has fully converged either. In this situation R1 should not send traffic to R2 immediatly. In that case R2 could send LSInfinity(0xFFFF) resulting in a situation where R1 must wait for R2 to be fully be available and transit states have been passed. From the moment LDP converged on R2, it can distribute the traditional Interface OSPF metric value. This operation will result that OSPF and LDP have a converged behaviour. Thimportance and the description of this behaviour can be found in [LDP-Sync].

3.3. Configuration change

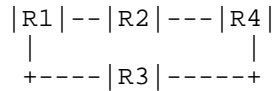


Figure 3

When operator needs R2 configuration change, R2 sends LSInfinity(0xFFFF) for traffic engineering. R2 configuration completed, R2 sends correct metric value in OSPFv2 area.

4. OSPFv3 operation

4.1. R-bit and v6-bit

[OSPFv3] explains at section 2.7 the following: If the "R-bit" is clear, an OSPF speaker can participate in OSPF topology distribution without being used to forward transit traffic. The V6-bit specializes the R-bit; if the V6-bit is clear, an OSPF speaker can participate in OSPF topology distribution without being used to forward IPv6 datagrams. If the R-bit is set and the V6-bit is clear, IPv6 datagrams are not forwarded but datagrams belonging to another protocol family may be forwarded.

This protocol implementation is useful in multi address family environment such as [OSPFv3-AF]. But Service Provider operators have to check both the "R-bit" and "v6-bit" during their operation and introduce both training and operational changes to make this a true usable technology. Operators have believe that a usefull approach would be to rely upon successfull IPv4 OSPFv2 behaviour and to add a "OSPFv2 compatibility mode" in IPv6 only environment to mimic OSPFv2 behaviour in this environment.

The functionality of the R-bit and v6-bit operations is described in [OSPFv3]'s Errata 2078 more detail.

A Router should support a "R-bit" know with a clear wait for BGP or waiting-before-becoming-active time on start-up same as [RFC3137] indicates.

4.2. OSPFv2 compatibility mode

An OSPFv3 routing device has through the area scope LSAs metric information of all of devices. As result the router can announce the interface metric LSInfinity(0xFFFF). This is simple implementation

model not requiring operational service provider changes.

5. Acknowledgements

Tsuyohi Momose

6. IANA Considerations

This document has no actions for IANA.

7. Security Considerations

The technique described in this document does not introduce any new security issues into OSPFv3 protocol.

8. Normative References

[LDP-Sync]

M. Jork, A. Atlas, L. Fang, "LDP IGP Synchronization", March 2009, <<http://tools.ietf.org/html/rfc5443>>.

[OSPFv2]

J. Moy, "OSPF Version 2", 1998, <<http://tools.ietf.org/html/rfc2328>>.

[OSPFv3]

R. Coltun, D. Ferguson, J. Moy, A. Lindem, "OSPF for IPv6", July 2008, <<http://tools.ietf.org/html/rfc5340>>.

[OSPFv3-AF]

A. Lindem, S. Mirtorabi, A. Roy, M. Barnes, R. Aggarwal, "Support of Address Families in OSPFv3", April 2010, <<http://tools.ietf.org/html/rfc5838>>.

[RFC2119]

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3137]

A. Retana, L. Nguyen, R. White, A. Zinin, D. McPherson, "OSPF Stub Router Advertisement", June 2001, <<http://tools.ietf.org/html/rfc3137>>.

Appendix A. Additional Stuff

This becomes an Appendix.

Authors' Addresses

Shishio Tsuchiya (editor)
Cisco Systems
Shinjuku Mitsui Building, 2-1-1, Nishi-Shinjuku
Shinjuku-Ku, Tokyo 163-0409
Japan

Phone: +81 3 6434 6543
Email: shtsuchi@cisco.com

Gunter Van de Velde
Cisco Systems
Pegasus Parc
De kleetlaan 6a, DIEGEM, BRABANT 1831
BELGIUM

Phone: +32 2 704 5473
Email: gunter@cisco.com

Tomohiro Yamagata
KDDI Corporation
Garden Air Tower, 3-10-10, Iidabashi
Chiyoda-Ku, Tokyo 102-8460
Japan

Phone: +81 3 6678 3089
Email: to-yamagata@kddi.com

OSPF Working Group
Internet Draft
Intended status: Standards Track
Expires: April 14, 2011

Y. Yang
A. Retana
A. Roy
Cisco Systems
October 11, 2010

Hiding Transit-only Networks in OSPF
<draft-yang-ospf-hiding-00.txt>

Abstract

A transit-only network is defined as a network connecting routers only. In OSPF, transit-only networks are usually configured with routable IP addresses, which are advertised in LSAs but not needed for data traffic. In addition, remote attacks can be launched against routers by sending packets to these transit-only networks. This document presents a mechanism to hide transit-only networks to speed up network convergence and minimize remote attack vulnerability.

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 14, 2011.

Copyright Notice

Copyright (c) 2010 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	3
2. Hiding IPv4 Transit-only Networks in OSPFv2	4
2.1. Point-to-Point Networks	4
2.1.1. Advertising Point-to-Point Networks	4
2.1.2. Hiding Point-to-Point Networks	5
2.2. Broadcast Networks	5
2.2.1. Advertising Broadcast Networks	5
2.2.2. Hiding Broadcast Networks	6
2.2.2.1. Sending Network LSA	6
2.2.2.2. Receiving Network LSA	6
2.2.2.2.1. Backward Compatibility	6
2.3. Non-Broadcast Networks	7
2.3.1. NBMA	7
2.3.2. Point-to-MultiPoint	7
2.3.2.1. Advertising Point-to-MultiPoint Networks	8
2.3.2.2. Hiding Point-to-MultiPoint Networks	8
3. Hiding IPv6 Transit-only Networks in OSPFv3	9
4. Hiding AF Enabled Transit-only Networks in OSPFv3	9
5. Operational Considerations	10
6. Security Considerations	10
7. IANA Considerations	10
8. References	10
8.1. Normative References	10
8.2. Informative References	10
Appendix A. Acknowledgments	11
Authors' Addresses	11

1. Introduction

A transit-only network is defined as a network connecting routers only. In OSPF, transit-only networks are usually configured with routable IP addresses, which are advertised in LSAs but not needed for data traffic. In addition, remote attacks can be launched against routers by sending packets to these transit-only networks. This document presents a mechanism to hide transit-only networks to speed up network convergence and minimize remote attack vulnerability.

Hiding transit-only networks will not impact reachability to the end hosts.

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 [KEYWORD].

2. Hiding IPv4 Transit-only Networks in OSPFv2

In [OSPFv2], networks are classified as point-to-point, broadcast, or non-broadcast. In the following sections, we will review how these OSPF networks are being advertised and discuss how to hide them consequently.

2.1. Point-to-Point Networks

A point-to-point network joins a single pair of routers. Figure 1 shows a point-to-point network connecting routers RT1 and RT2.

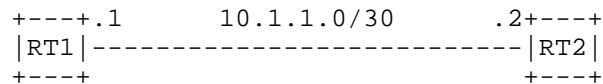


Figure 1 Physical point-to-point network

2.1.1. Advertising Point-to-Point Networks

For each numbered point-to-point network, a router has 2 link descriptions in its router LSA, one Type 1 link (point-to-point) regarding the neighboring router, and one Type 3 link (stub) regarding the assigned IPv4 address.

An example of router LSA originated by RT1 would look like

```

LS age = 0                      ;newly (re)originated
LS type = 1                      ;router-LSA
Link State ID = 1.1.1.1          ;RT1's Router ID
Advertising Router = 1.1.1.1     ;RT1's Router ID
#links = 2
  Link ID = 2.2.2.2              ;RT2's Router ID
  Link Data = 10.1.1.1           ;Interface IP address
  Type = 1                      ;connects to RT2
  Metric = 10

  Link ID= 10.1.1.0              ;Interface IP address
  Link Data = 255.255.255.252    ;Subnet's mask
  Type = 3                      ;Connects to stub network
  Metric = 10

```

The Type 1 link will be used for SPF calculation while the Type 3 link will be used for RIB installation.

2.1.2. Hiding Point-to-Point Networks

To hide a transit-only point-to-point network, the Type 3 link MUST be removed from the router LSA.

An example of router LSA originated by RT1, hiding the point-to-point network depicted in Figure 1, would look like

```

LS age = 0                      ;newly (re)originated
LS type = 1                      ;router-LSA
Link State ID = 1.1.1.1          ;RT1's Router ID
Advertising Router = 1.1.1.1     ;RT1's Router ID
#links = 1
  Link ID = 2.2.2.2              ;RT2's Router ID
  Link Data = 10.1.1.1           ;Interface IP address
  Type = 1                       ;connects to RT2
  Metric = 10

```

2.2. Broadcast Networks

A broadcast networks joins many (more than two) routers, and supports the capability to address a single physical message to all of the attached routers. Figure 2 shows a broadcast network connecting router RT3, RT4, and RT5.

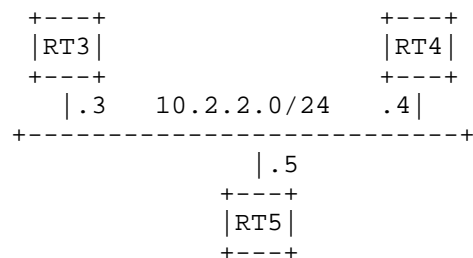


Figure 2 Broadcast network

2.2.1. Advertising Broadcast Networks

For each broadcast network, a designated router (DR) describes it in its network LSA. Assuming RT3 is elected as the DR in Figure 2, an example of the network LSA originated by RT3 would look like

```
LS age = 0                               ;newly (re)originated
LS type = 2                               ;network-LSA
Link State ID = 10.2.2.3                 ;IP address of the DR (RT3)
Advertising Router = 3.3.3.3             ;RT3's Router ID
Network Mask = 255.255.255.0
  Attached Router = 3.3.3.3               ;Router ID
  Attached Router = 4.4.4.4               ;Router ID
  Attached Router = 5.5.5.5               ;Router ID
```

OSPF obtains the IP network number from the combination of the Link State ID and the Network Mask. In addition, Link State ID is also being used for 2-way connectivity check.

2.2.2. Hiding Broadcast Networks

2.2.2.1. Sending Network LSA

To hide a transit-only broadcast network, a special network mask value 255.255.255.255 MUST be used in the network LSA. While a broadcast network connects more than routers, using 255.255.255.255 will not hide an access broadcast network accidentally.

As there is no change of the Link State ID, the 2-way connectivity check would proceed normally.

An example of network LSA originated by RT3, hiding the broadcast network depicted in Figure 2, would look like

```
LS age = 0                               ;newly (re)originated
LS type = 2                               ;network-LSA
Link State ID = 10.2.2.3                 ;IP address of the DR (RT3)
Advertising Router = 3.3.3.3             ;RT3's Router ID
Network Mask = 255.255.255.255           ;special subnet mask
  Attached Router = 3.3.3.3               ;Router ID
  Attached Router = 4.4.4.4               ;Router ID
  Attached Router = 5.5.5.5               ;Router ID
```

2.2.2.2. Receiving Network LSA

It's RECOMMENDED that all routers in an area be upgraded as a whole to process the modified network LSA correctly and consistently.

When a router receives a network LSA, it MUST check the 2-way connectivity as normal. However, if the network mask in the network LSA is 255.255.255.255, the router MUST NOT install the route in the RIB.

2.2.2.2.1. Backward Compatibility

When a not-yet-upgraded router receives a modified network LSA, as specified in section 2.2.2.1, a host route to the originating DR will be installed. This is not ideal but better than the current result, which exposes the whole subnet.

In a partial deployment scenario, upgraded routers and not-yet-upgraded routers may mix up. The former do not have the host routes aforementioned, while the latter do have. Such inconsistency creates routing black holes, which should be avoided normally. In this case, however, as packets destined for the transit-only networks are dropped somewhere in the network, the black holes actually help DRs defend from the remote attacks.

In summary, the modification of the network LSA, as specified in section 2.2.2.1, is backward compatible with the current specification of [OSPFv2], even in a partial deployment case.

2.3. Non-Broadcast Networks

A non-broadcast networks joins many (more than two) routers, but does NOT support the capability to address a single physical message to all of the attached routers. As mentioned in [OSPFv2], OSPF runs in one of two modes over non-broadcast networks: NBMA or Point-to-MultiPoint.

2.3.1. NBMA

In NBMA mode, OSPF emulates operation over a broadcast network: a Designated Router is elected for the NBMA network, and the Designated Router originates an LSA for the network.

To hide a NBMA transit-only network, OSPF adopts the same modification over the broadcast transit-only network, as defined in section 2.2.2.

2.3.2. Point-to-MultiPoint

In point-to-MultiPoint mode, OSPF treats the non-broadcast network as a collection of point-to-point links.

Figure 3 shows a non-broadcast network connecting router RT6, RT7, RT8, and RT9. In this network, all routers can communicate directly, except for routers RT7 and RT8.

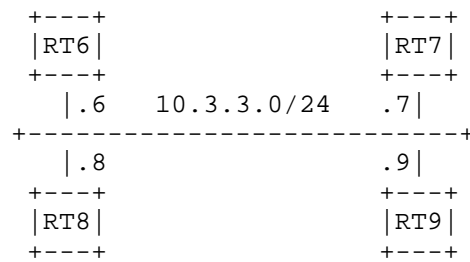


Figure 3 Non-Broadcast network

2.3.2.1. Advertising Point-to-MultiPoint Networks

For a point-to-multipoint network, a router has multiple link descriptions in its router LSA, one Type 1 link (point-to-point) for EACH directly communicable router, and one Type 3 link (stub) regarding the assigned IPv4 address.

An example of router LSA originated by RT7 would look like

```

LS age = 0                               ;newly (re)originated
LS type = 1                               ;router-LSA
Link State ID = 7.7.7.7                   ;RT7's Router ID
Advertising Router = 7.7.7.7               ;RT7's Router ID
#links = 3
  Link ID = 6.6.6.6                       ;RT6's Router ID
  Link Data = 10.3.3.7                     ;Interface IP address
  Type = 1                                 ;connects to RT6
  Metric = 10

  Link ID = 9.9.9.9                       ;RT9's Router ID
  Link Data = 10.3.3.7                     ;Interface IP address
  Type = 1                                 ;connects to RT9
  Metric = 10

  Link ID= 10.3.3.7                        ;Interface IP address
  Link Data = 255.255.255.255              ;Subnet's mask
  Type = 3                                 ;Connects to stub network
  Metric = 0

```

2.3.2.2. Hiding Point-to-MultiPoint Networks

To hide a transit-only point-to-multipoint network, the Type 3 link MUST be removed from the router LSA.

An example of router LSA originated by RT7, hiding the point-to-

point network depicted in Figure 3, would look like

```
LS age = 0 ;newly (re)originated
LS type = 1 ;router-LSA
Link State ID = 7.7.7.7 ;RT7's Router ID
Advertising Router = 7.7.7.7 ;RT7's Router ID
#links = 2
    Link ID = 6.6.6.6 ;RT6's Router ID
    Link Data = 10.3.3.7 ;Interface IP address
    Type = 1 ;connects to RT6
    Metric = 10

    Link ID = 9.9.9.9 ;RT9's Router ID
    Link Data = 10.3.3.7 ;Interface IP address
    Type = 1 ;connects to RT9
    Metric = 10
```

3. Hiding IPv6 Transit-only Networks in OSPFv3

In [OSPFv3], addressing semantics have been removed from the OSPF protocol packets and the main LSA types, leaving a network-protocol-independent core.

More specifically, Router-LSAs and network-LSAs no longer contain network addresses, but simply express topology information. A new LSA called the intra-area-prefix-LSA has been introduced. This LSA carries all IPv6 prefix information that in [OSPFv2] is included in router-LSAs and network-LSAs.

Such changes simplify the process to hide the IPv6 addresses of the transit-only networks in [OSPFv3] -- simply removing the correspondent IPv6 unicast prefixes from the intra-area-prefix-LSA will do the trick.

4. Hiding AF Enabled Transit-only Networks in OSPFv3

[OSPF-AF] supports multiple address families (AFs) by mapping each AF to a separate Instance ID and OSPFv3 instance.

In the meantime, each prefix advertised in OSPFv3 has a prefix Length field [OSPFV3], which facilitates advertising prefixes of different lengths in different AFs. The existing LSAs defined in OSPFv3 are used for prefix advertising and there is no need to define new LSAs.

In other words, intra-area-prefix-LSAs are still being used to advertise the attached networks, and same method explained in section

3 can also be used to hide those AF enabled transit-only networks.

5. Operational Considerations

By eliminating the ability to reach transit-only networks, the ability to manage these interfaces may be reduced. In order to not reduce the functionality and capability of the overall network, it is recommended that extensions such as RFC5837 be also implemented.

6. Security Considerations

One motivation for this document is to reduce remote attack vulnerability by hiding transit-only networks. The result should then be that fewer OSPF core networks will be exposed to unauthorized access.

While the steps described in this document are meant to be applied to transit-only networks ONLY, they could be used to hide other networks as well. It is expected that the same care that users put on the configuration of other routing protocol parameters is used in the configuration of this extension.

7. IANA Considerations

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

8. References

8.1. Normative References

- [KEYWORD] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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.
- [OSPF-AF] Lindem, A., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC5838, April 2010.

8.2. Informative References

[RFC5837] Atlas, A., Bonica, R., Pignataro, C., Shen, N., and JR. Rivers, "Extending ICMP for Interface and Next-Hop Identification", RFC5837, April 2010.

Appendix A. Acknowledgments

The draft text was produced using Stefan Santesson's NroffEdit application.

The idea of using a special subnet mask to hide broadcast networks in OSPF was originally introduced in US pending patent application of "Apparatus and Method to Hide Transit Only Multi-Access Networks in OSPF" (Publication Number: US 2008/0080494 A1), authored by Yi Yang, Alvaro Retana, James Ng, Abhay Roy, Alfred Lindem, Sina Mirtorabi, Timothy Gage, and Khalid Raza.

Authors' Addresses

Yi Yang
Cisco Systems
7025 Kit Creek Road
RTP, NC 27709
USA

EMail: yiya@cisco.com

Alvaro Retana
Cisco Systems
7025 Kit Creek Road
RTP, NC 27709
USA

EMail: aretana@cisco.com

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

Email: akr@cisco.com