KARP Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 18, 2011

M. Bhatia
Alcatel-Lucent
S. Hartman
Painless Security
D. Zhang
Huawei Technologies co., LTD.
February 14, 2011

Security Extension for OSPFv2 when using Manual Key Management
draft-bhatia-karp-ospf-ip-layer-protection-03

Abstract

The current OSPFv2 cryptographic authentication mechanism as defined in the OSPF standards is vulnerable to both inter-session and intra-session replay attacks when its uses manual keying. Additionally, the existing cryptographic authentication schemes do not cover the IP header. This omission can be exploited to carry out various types of attacks.

This draft proposes an authentication scheme based on a challenge-response mechanism that will protect OSPFv2 from both inter and intra replay attacks when its using manual keys for securing its protocol packets. For comparison, an approach based on making sequence numbers unique is presented. Later we also describe some changes in the cryptographic hash computation so that we eliminate most attacks that result because of OSPFv2 not protecting the IP header.

Status of this Memo

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

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

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

This Internet-Draft will expire on August 18, 2011.

Copyright Notice

Copyright (c) 2011 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

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 cryptographic authentication mechanism as described in [[RFC2328]] uses per-packet sequence numbers to provide protection against replay attacks. The sequence numbers increase monotonically so that the attempts to replay the stale packets can be thwarted. The sequence number values are maintained as a part of adjacency states. Therefore, if an adjacency is broken down, the associated sequence numbers get re-initiated and the neighbors start all over again. Additionally, the cryptographic authentication mechanism does not specify how to deal with the rollover of a sequence number when it reaches its maximum limit. These omissions can be taken advantage of by attackers to implement various replay attacks ([RFC6039]). In order to address these issues, we propose a challenge/ response mechanism that introduces two additional random numbers to help routers generate distinguishable new states when the sequence numbers need to be re-initiated. Compared with the cryptographic authentication mechanism proposed in [RFC5709], the solution proposed does not impose any more security presumption.

The cryptographic authentication as described in [RFC2328] and later updated in [RFC5709] does not include the IP header. This also 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 recieved from. Changing the source address of a packet can thus, confuse the reciever which can be exploited to produce a number of denial of service attacks [RFC6039]. If the packet is interpreted as coming from a different neighbor, the 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]. This is referred to as the IP layer issue in [I-D.ietf-karp-threats-reqs].

[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 [RFC6094].

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

2. A Challenge and Response Solution

In OSPFv2, a non-decreasing sequence member is associated with each OSPF packet sent from a router in order to prevent replay attacks. However, as illustrated in [I-D.hartman-ospf-analysis] and [RFC6039], in the circumstances where automatic key management mechanisms are unavailable, any re-initiation of sequence numbers can potentially be taken advantage of to perform replay attacks. In this section, we introduce an extension of the OSPFv2 protocol, which uses challenge/ response to benefit the verification of the freshness of OSPF packets when the sequence numbers of routers are re-initiated. This solution eliminates the reliance on automatic key management mechanisms. However, it is assumed that a traffic key is shared between two communicating routers so that an attacker can play antique packets but lacks the capability to modify packets without being detected.

In this protocol, two random numbers (Session ID and Nonce) are introduced. The session ID is used to identify the session a packet is within and thus makes inter-session replay attacks difficult. The nonce is used to challenge the liveness of communicating routers so that states need not be maintained with routers that are not currently neighbors. In combination with the sequence number, the session ID can effectively resist intra-session replay attacks. When the sequence space is exhausted, a router simply chooses a new session ID.

Figure 1 illustrates how two routers A and B, challenge each other's liveness when they are initially connected to a link. First, A selects a new session ID (X1) and a new nonce (N1), and sends them out within a hello packet (see step 1). Particularly, X1and N1 are encapsulated in the OSPF header of the packet. Note that if A is on a multicast LAN, the packet is sent using multicast. Similarly, B sends a hello packet with its new session ID (X2) and Nonce (N2) (step 2). Upon receiving the hello packet from B, A sends a hello packet with X1 and N1. In the neighbor field of the packet, the router ID of B, X2, and N2 is encapsulated (Step 3). Upon receiving the packet sent in step 3, B can ensure the freshness of the packet if the attached session ID and nonce values of both routers are correct.

In the same way, after receiving the hello packet from A, B sends a hello packet with X2 and N2 in the OSPF header, and in the neighbor field of the packet, the router ID of A, X1, and N1 is listed to identify that A has been discovered. After receiving the packet, A can make sure the packet is fresh if the session IDs and the nonce of both routers contained in the packet are correct. After A and B discover each other, they start exchanging their database information (steps 5 and 6). During the exchange, every packet from Router A is

associated with X1 and N1, while every packet from Router B is associated with X2 and N2. Each of these packets also contains a sequence number as part of the cryptographic authentication option. The sequence number MUST increase for every packet sent.

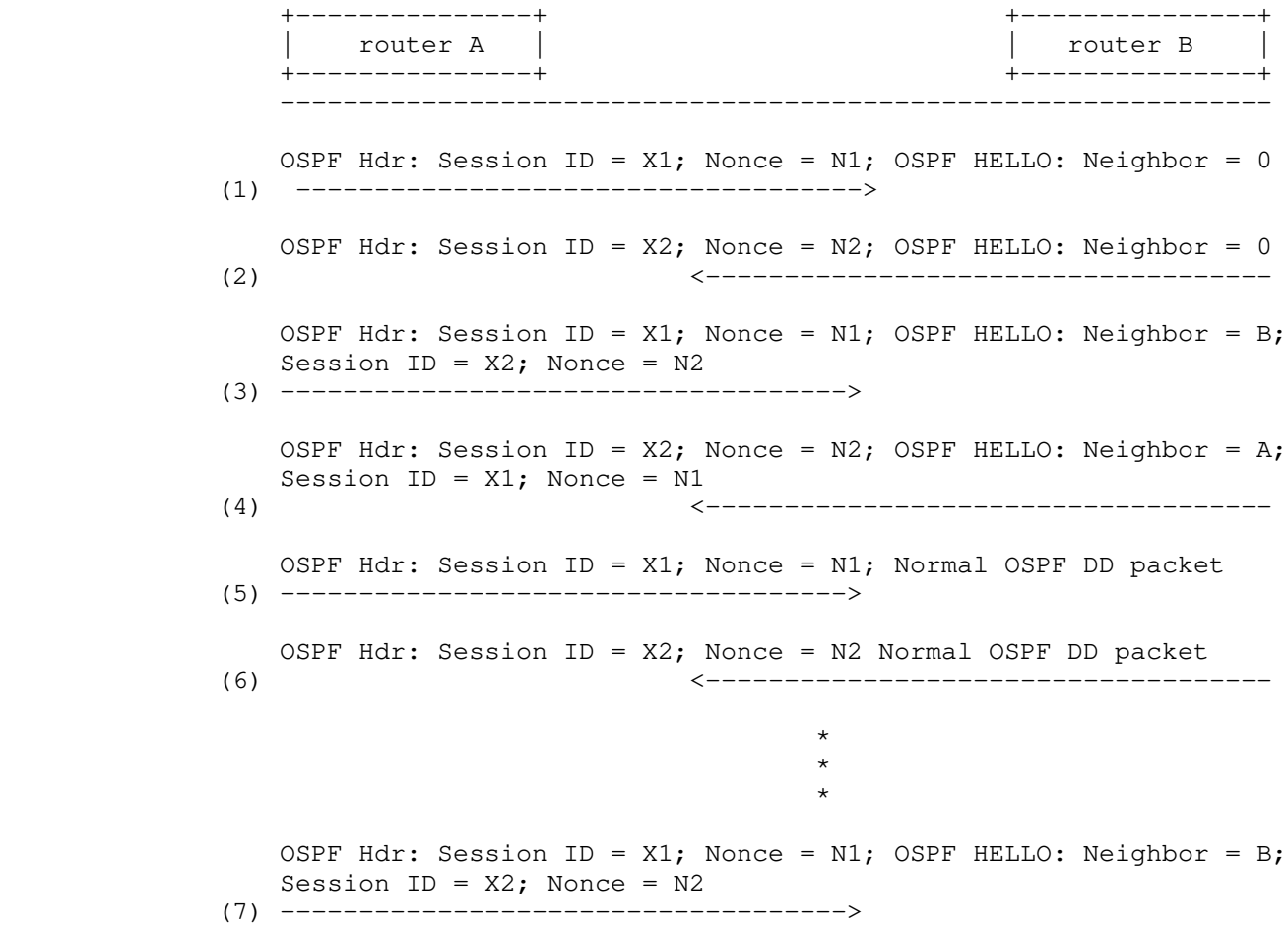


Figure 1 Scenario: two Routers coming up on a LAN

After A and B have generated a neighbor relationship, assume another router, C, is connected to the link. C finds the existence of A and intends to become a neighbor of A. The packets exchanged during this process are illustrated in Figure 2. Firstly, C selects a new session ID (X3) and a new nonce value (N3), and sends them out within a blank hello packet (see the second step of Figure 2). After receiving this packet, A sends out a hello packet with the information of C (router ID, X3, and N3) in the neighbor field.

Because A is challenging the liveness of a new neighbor, A selects a new nonce N1' and encapsulates it in the OSPF header of the hello packet to challenge whether the packet sent in step 2 is really from C. After receiving the packet from A, C can make sure the packet is valid since it consists of its current session ID and nonce (e.g., X3 and N3). Thus, C replies to A with a hello packet including the information of A (e.g., X1 and N1') in the neighbor field. After receiving this packet and checking the correctness of X1 and N1', A can ensure that the packet is fresh and C is currently online.

It worthwhile to note that during the challenge and response the hello packets sent immediately amnogst routers.

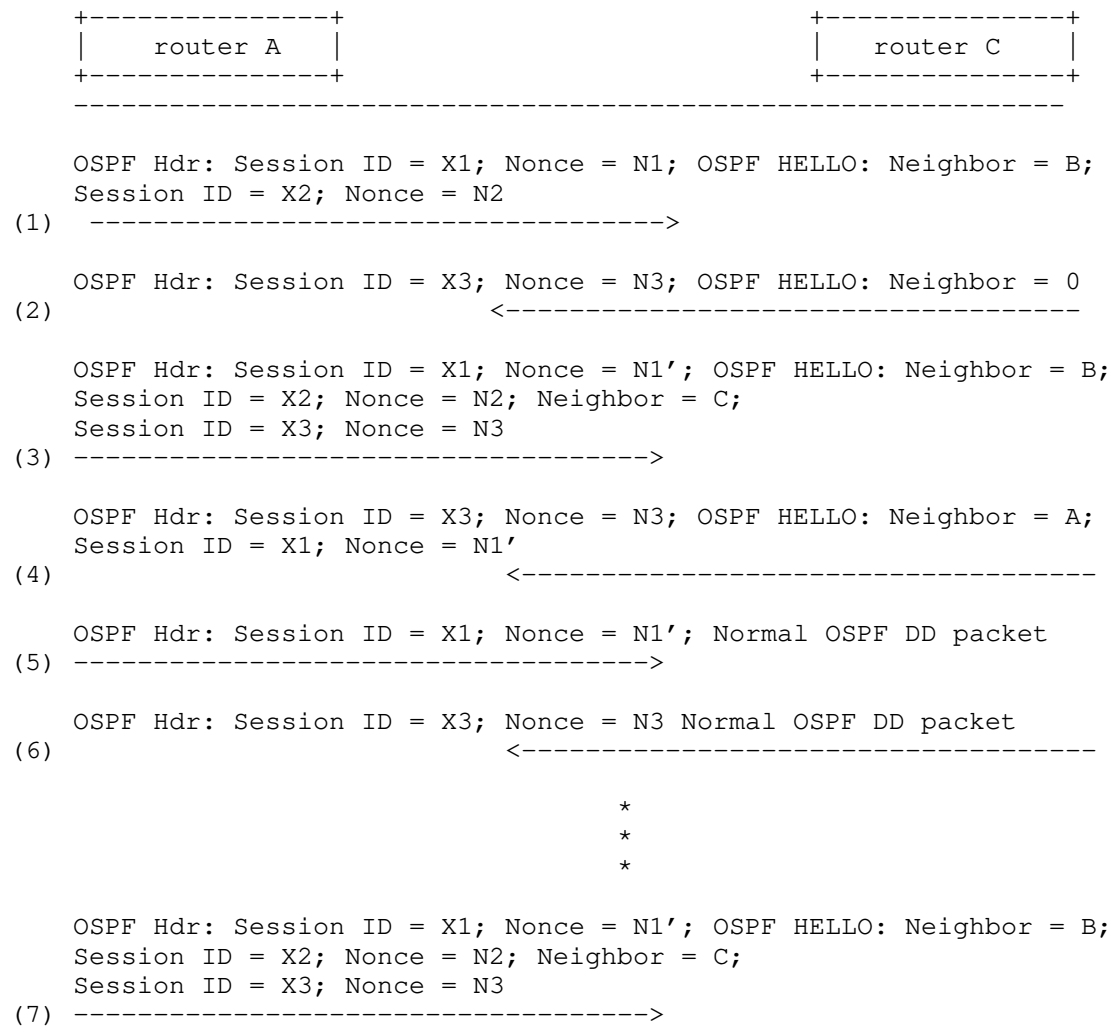


Figure 2. Scenario: another Router C comes up on that LAN

Figure 3 illustrates the scenario in which router A is rebooted. After the reboot, A lost its state and selects a new session ID (X4) and a new nonce value (N4). However, B still maintains the earlier session ID and nonce values of A (X1 and N1). In step 1, A sends a blank hello packet out with its new session ID and nonce value. After receiving the hello packet, B realizes that the session ID and the nonce value of A in the OSPF header are different from the ones maintained in its database. In order to distinguish a reboot from a replay of an old packet, B selects a new nonce value, N2', and

transports it as well as its session ID (X2) in a hello packet to check whether the packet is from A. In the neighbor field of this packet, B continues listing A with the earlier session ID and nonce values (i.e., X1 and N1). Therefore, if an attacker attempts to send an antique packet to masquerade as A, A would update its database with the new nonce of B and send a hello packet with its existing Session ID and nonce values (X1 and N1). In step 3, B receives a new hello packet consisting of B's new nonce value from A. Since this packet lists B with the new nonce value in the neighbors field of the hello and since the nonce is new, this packet cannot be a replay. Now, B can safely assume that A has indeed restarted and can start using the new session ID and the nonce values sent by A in the neighbor field of its hellos.

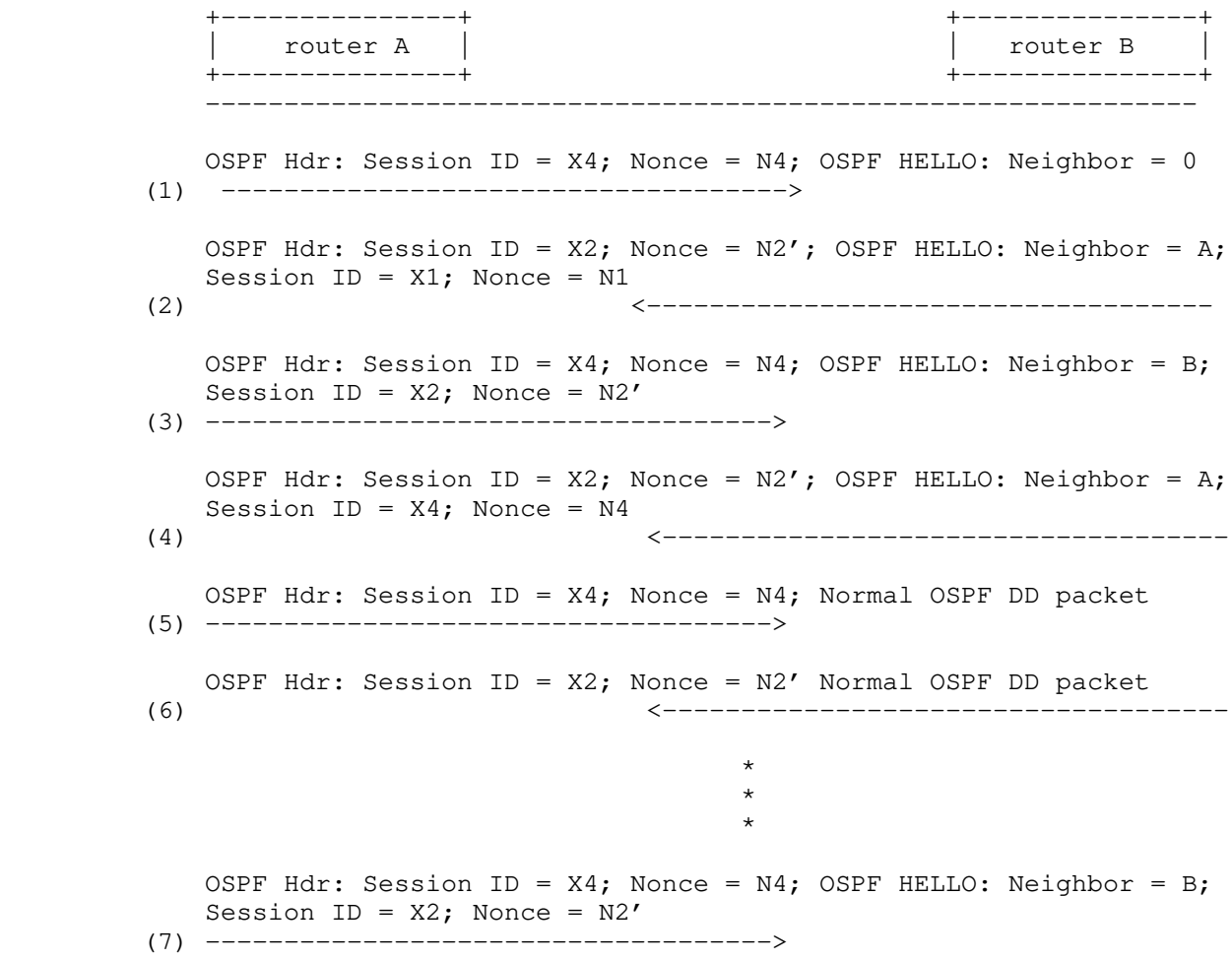


Figure 3. Scenario: Router A Reboot

2.1. Neighbor State Required

This authentication type requires the following additional fields be stored per neighbor:

- o The session ID most recently received from a neighbor
- o The nonce most recently received from a neighbor; this only needs to be kept up-to-date when the session ID changes or when establishing an adjacency

- o A set of sequence numbers for the neighbor; if packets are sometimes processed out of order, then a sequence number MAY be maintained for each type of packet

2.2. Receiver Behavior

This section describes how OSPF receivers will handle the reception of packets with the nonce and session ID.

If a packet is received for a neighbor in at least the 2-way state, then the session ID is compared to the one stored in the neighbor table. If the session ID does not match the session ID recorded with the neighbor, and the packet is not a hello, the packet is discarded. If the packet is a hello, then rules for hellos in following paragraphs apply. Otherwise, if the session ID matches, then if the sequence number in the cryptographic authentication option is not strictly greater than the sequence number associated with the neighbor for this type of packet, then the packet is discarded. If the cryptographic verification of the checksum fails, the packet is discarded. Otherwise, the packet is accepted by the cryptographic authentication and the sequence number associated with the neighbor for this packet type is updated to be the sequence number in the packet. The router MAY update the nonce associated with the neighbor to a nonce in a received hello packet. Updating the nonce is optional because the adjacency is already established. One case where a router implementationd would want to update nonces is where the router has recently changed session IDs without dropping all adjacencies. Such a session ID change is likely to be rare, either the result of a reboot that preserved adjacencies but might not preserve sequence numbers or running out of sequence number space.

If a hello is received for a neighbor that is not found or that has not reached 2-way state the following steps are performed. If a neighbor structure exists for the neighbor and the session ID match that stored in the neighbor structure, then the packet is processed as follows. The sequence number is checked and MUST be strictly greater than the sequence number in the neighbor structure. The cryptographic authentication is verified. If this router is listed in the set of neighbors in the hello packet, the nonce and session ID MUST match this router's current nonce and session ID. If any of these checks fail, the packet is discarded. Otherwise the packet is accepted past cryptographic processing.

By this point, the router has received a hello packet. Either no neighbor structure exists or the session ID has changed. Before permitting communication with this router, its liveness needs to be challenged. If a neighbor has been deleted (because of a timeout) since the last nonce trigger, then a nonce trigger (see Section 2.3is

performed and the packet is discarded. If this router is listed in the list of neighbors, it MUST be listed with its current session ID and nonce otherwise the packet is discarded. If verification of the cryptographic checksum fails, the packet is discarded. If the neighbor is already in 2-way state or greater and this router is not listed in the set of neighbors, the packet is discarded. Otherwise, the session ID, nonce and all sequence numbers associated with the neighbor are updated from the packet and the packet is accepted by cryptographic authentication processing.

2.3. Nonce Triggers

The router keeps track of whether a nonce trigger has happened since the last time a neighbor is deleted.

In order to test liveness, a router updates its current nonce to a new value. As a side effect, all routers on the link that do not already have an adjacency with this router will update the nonce associated with this router. More importantly, though, the router we are testing liveness with will update the nonce in its hello entry for this router. That will allow this router to confirm that the session ID is correct and corresponds to current replay state.

As part of a nonce trigger, the router updates its current nonce. If a hello has not been sent too recently, then a hello is sent with the new nonce. The nonce trigger state is updated to indicate that no new neighbors have been deleted since the last nonce trigger.

3. Packet Format

In the challenge/ response mechanism, every OSPFv2 packet MUST carry the current Session ID and the associated Nonce value. This section describes how this information is carried in the OSPFv2 packets.

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

When using this authentication scheme the 64 bit Authentication field in the OSPF packet header remains unchanged and is the same as defined in Section D.3 of [RFC2328]. NOTE to the WG: We can also increase the size of the Key ID. Currently it has been kept as, but nothing prevents us from changing this.

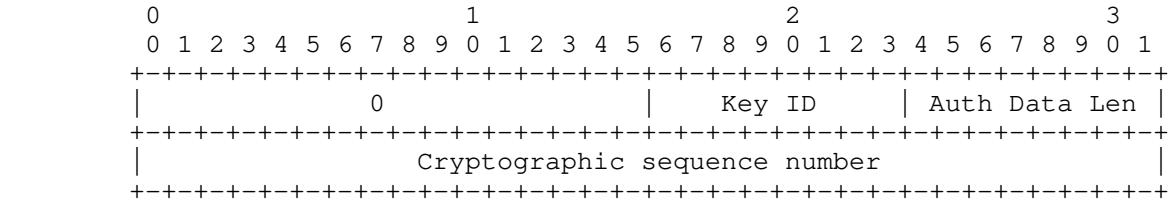


Figure 4.Usage of the Authentication field in the OSPF header when this mechanism is employed

The Session ID and the Nonce information is placed before the message digest that is appended to the OSPF packet. In this case too, the final Authentication data is not actually considered part of the OSPF protocol packet.

3.1. Extensions to OSPF packets

This section describes the new OSPFv2 packet format when this authentication scheme is being used.

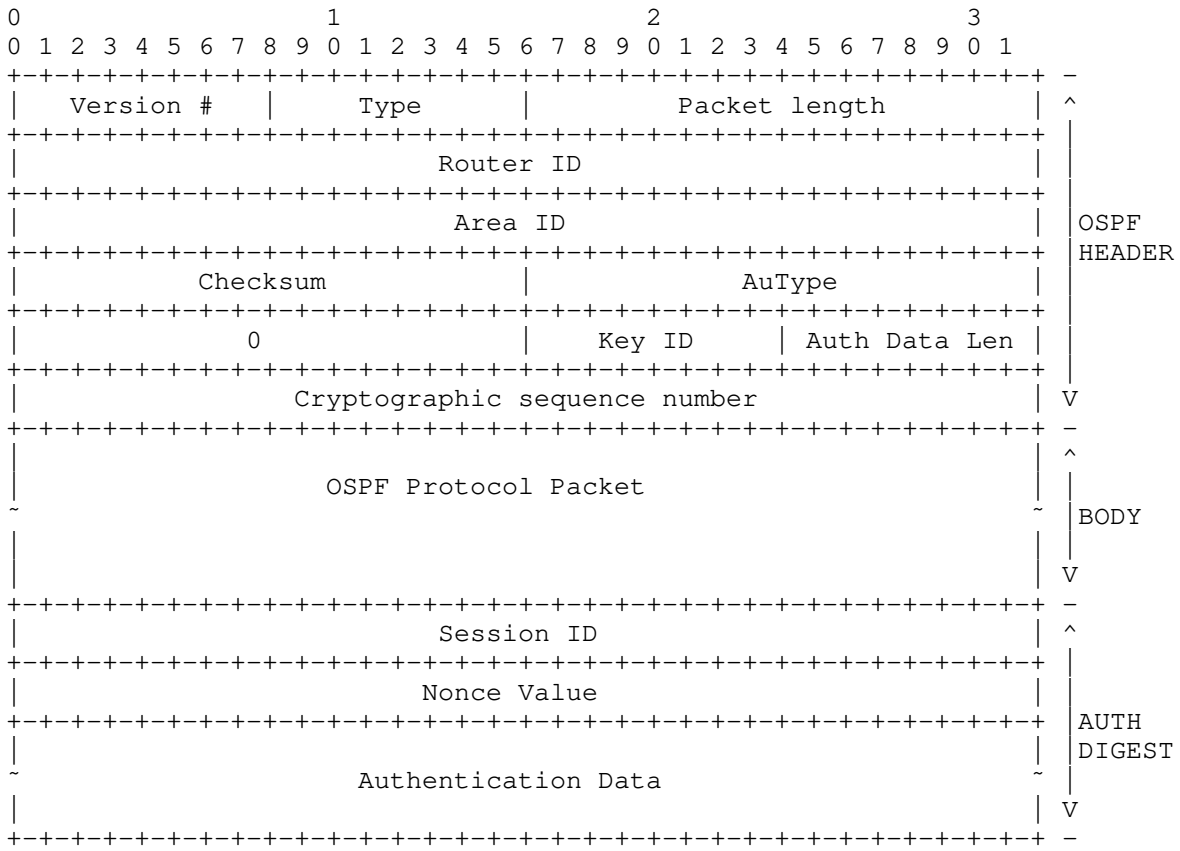
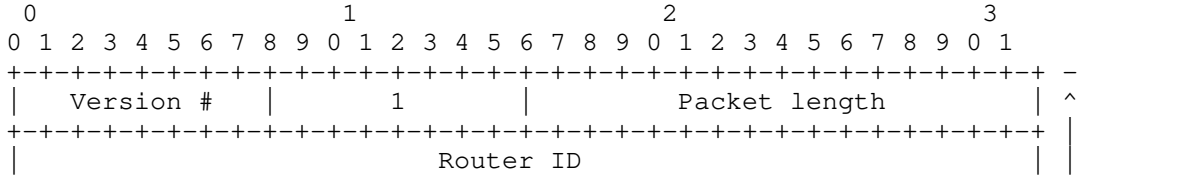


Figure 5.OSPFv2 Packet view

3.2. Extension of Hello Packet

The following figure shows an OSPF HELLO packet when this authentication scheme is being used. The HELLO payload has been modified to include each neighbor's Session ID and the Nonce value. The authentication data, as described above, carries the router's current Session ID and the Nonce value.



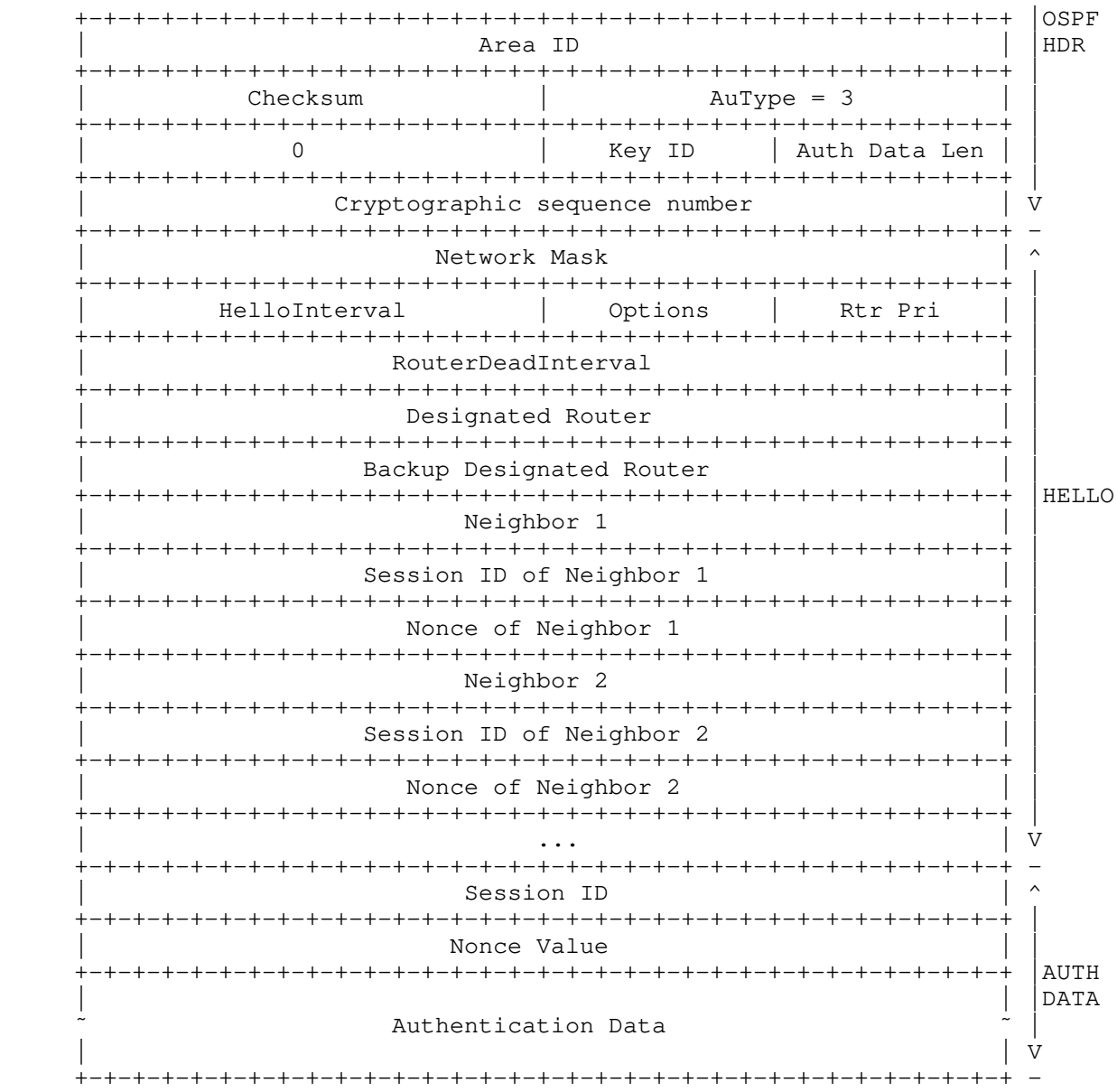


Figure 6.Extension of Protocol Packet

4. Key Selection in Processing OSPF Packets

This section introduces how the proposed security solution looks up long lived keys from key tables [I-D.ietf-karp-crypto-key-table]. Generally, a proper key selected to process an OSPFv2 packet should satisfy the requirements listed as follows:

the key is in its valid period; and

the key can be used for the desired security algorithm.

In the remainder of this section, other requirements that a selected key should particularly satisfy are depicted in different scenarios.

4.1. Key Selection in Sending Unicast OSPF Packets

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

the Peer field includes the router ID of R2;

the PeerKeyID field is not "unknown";

the Interfaces field includes I1; and

the Direction field is either "out" or "both".

When R1 and R2 are at the ends of a virtual link, the condition is a little more complex. Because the virtual link can be regarded as an unnumbered point-to-point network, the IP address of the interface actually used to send the packet (i.e., I1) is discovered during the routing table build process. Therefore, when the system operator deploys the keys to protect the virtual link, I1 has not been specified yet. Therefore, the key should be identified by the router IDs rather than by the interface originating the packet, and the third requirement introduced above should be changed to "the Interface field includes the router ID".

4.2. Key Selection in Sending Multicast OSPF Packets

If a router R1 sends an OSPF packet from its interface I1 to a multicast address (e.g., AllSPFRouters, AllDRouters), it needs to

select a key according to the following requirements:

the Peer field includes the multicast address;

the PeerKeyID field is "group";

the Interfaces field includes I1; and

the Direction field is either "out" or "both".

4.3. Key Selection on Receiving OSPF Packets

When Cryptographic Authentication is employed, the ID of the adopted key is encapsulated within the authentication field of an OSPF packet header. Using this ID, it is relatively easy for a receiver to locate the key. The requirement is relatively simple:

the Peer field includes the router ID of the sender; and

the PeerKeyID field includes the key ID obtained from the authentication field

5. Existing Cryptographic Authentication Mechanism

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.

6. Mechanism to secure the IP header

This document updates the definition of Apad which is currently a constant defined in [RFC5709] to the source address thats 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 recieving 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.

7. Alternative Boot Count Approach

During discussion of the challenge/response authentication approach, a desire was expressed to have a simpler alternative to consider. This section presents an alternative that obtains most advantages of the challenge/response mechanism. Instead of adding nonces and session IDs, OSPF implementations are required to keep a count of the number of times they have booted in non-volatile storage. This requirement is also placed on agents by the SNMPv3 security architecture; the same boot count can be used both for SNMP and for this OSPF mechanism.

The OSPF sequence number is extended to be 64-bits rather than 32-bits. The most significant 32-bits are the boot count. The least significant 32-bits is a counter that increases for every packet sent.

A receiver verifies that the sequence number on a received packet is strictly greater than the sequence number of the previous packet received.

Requiring that each packet have a strictly greater sequence number is a change from the current OSPF security model. However this change is required for a number of the security guarantees.

This mechanism requires fewer changes to the OSPF packet than the challenge/response mechanism. Also, the implementation complexity is somewhat less.

However there are disadvantages. First, this mechanism requires that the boot count be maintained succesfully in nonvolatile storage. If the boot count ever goes backwards without changing the encryption key, then all the attacks against the current OSPF protocol become possible against this protocol until the time that the boot count reaches a value greater than the largest value ever used for this client. This can be particularly problematic if equipment is replaced, using a router ID that has been used previously on a link but with a fresh boot count.

Another disadvantage is that the boot count mechanism does not protect against a session replayed while a router is down. If a router crashes or is taken out of service, then an attacker can replay packets as soon as the adjacencies with the router time out. The vulnerabilities of this have not been fully analyzed. Potential vulnerabilities include attacks on the designated router election process and replays of complete sessions. So far it looks like it is not likely that an attacker could bring up a replayed session far enough to inject routes from a down router.

8. Security Considerations

This document attempts to fix the manual key management procedure that currently exists within OSPFv2, as part of the Phase 1 of the KARP Working Group. This therefore, only considers manual key management mechanism to be used for OSPFv2. Any solution that takes advantage of the automatic key management mechanism is beyond the scope of this document.

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

9. IANA Considerations

This document requests a new Auth Type to be defined for OSPFv2. It currently uses 3 to foster pre-standard deployments.

10. Acknowledgements

The authors would like to thank Acee Lindem for valuable contributions and helping to understand the tradeoffs surrounding various solutions.

11. References

11.1. 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.

11.2. 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-02 (work in progress), December 2010.

[I-D.ietf-karp-crypto-key-table] Housley, R. and T. Polk, "Database of Long-Lived Symmetric Cryptographic Keys", draft-ietf-karp-crypto-key-table-00 (work in progress), November 2010.

[I-D.ietf-karp-threats-reqs] Lebovitz, G., Bhatia, M., and R. White, "The Threat Analysis and Requirements for Cryptographic Authentication of Routing Protocols' Transports", draft-ietf-karp-threats-reqs-01 (work in progress), October 2010.

[RFC6039] Manral, V., Bhatia, M., Jaeggli, J., and R. White, "Issues with Existing Cryptographic Protection Methods for Routing Protocols", RFC 6039, October 2010.

[RFC6094] Bhatia, M. and V. Manral, "Summary of Cryptographic Authentication Algorithm Implementation Requirements for Routing Protocols", RFC 6094, February 2011.

Authors' Addresses

Manav Bhatia
Alcatel-Lucent
Bangalore,
India

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

Sam Hartman
Painless Security

Email: hartmans@painless-security.com

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

Phone:
Fax:
Email: zhangdacheng@huawei.com
URI:

Network Working Group
Internet Draft
Intended status: Proposed Standard
Expires: September 2011

S. Giacalone
Thomson Reuters

D. Ward
Juniper Networks

J. Drake
Juniper Networks

A. Atlas
Juniper Networks

March 4, 2011

OSPF Traffic Engineering (TE) Express Path
draft-giacalone-ospf-te-express-path-00.txt

Abstract

In certain networks, such as, but not limited to, financial information networks (e.g. stock market data providers), network performance criteria (e.g. latency) have become (or are becoming) as (or more) critical to data path selection than other metrics.

This document describes extensions to OSPF TE (RFC3630) such that network performance information can be distributed and collected in a scalable fashion. The information collected from OSPF TE Express Path can then be used to make path selection decisions. Additionally, the information passed in these extensions will permit granular network performance monitoring.

Note that this document only covers the mechanisms with which network performance information is distributed. The mechanisms for measuring network performance or acting on that information, once distributed, are outside the scope of this document.

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), 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/1id-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 September 4, 2011.

Copyright Notice

Copyright (c) 2011 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

In certain networks, such as, but not limited to, financial information networks (e.g. stock market data providers), network performance information (e.g. latency) have (or are becoming) as (or more) critical to data path selection than other metrics. In many of these networks, bandwidth is relatively rich and homogeneous (e.g. a core network of all 10 or 20 Gigabit Ethernet links, or greater), however path length (and therefore latency) can vary in between end-points (e.g. PE nodes), and segment length or latency can change based on the path protection scheme used. In these networks, extremely large amounts of money rest on the ability to predictably make trades faster than the competition and the ability to access real time market data.

In certain financial services networks, hop count, cost, and bandwidth are only tangentially important. Rather, it would be

beneficial to be able to granularly monitor network performance and/or make path selection decisions based on performance data (such as latency) in a cost-effective and scalable way. In addition, since these networks may be built as overlays on top of multiple service provider networks, strict link-by-link service level agreement monitoring and enforcement mechanisms are needed.

This document describes extensions to OSPF TE (hereafter called "OSPF TE Express Path"), that can be used to distribute various pieces of network performance information (such as link latency). The mechanisms described in this document only disseminate performance information. The methods for initially gathering that performance information, or acting on it once it is distributed are outside the scope of this document. OSPF Express Path provides a number of benefits:

The data distributed by OSPF TE Express Path can be used to make path selection decisions. Using the link-by-link performance information data distributed by OSPF TE Express Path, end-to-end path selection can be performed based on performance metrics, as part of the normal operation of various routing protocols (e.g. by replacing cost with latency) or by using "second order" control plane protocols such as CSPF, RSVP-TE [RFC3209], etc.

OSPF TE Express Path enables a scalable, open mechanism for link-by-link SLA compliance monitoring, which is an important issue in large, diverse networks that use transport services from various providers. In networks like this, end-to-end latency is not always useful for enforcement of "underlying" SLAs (since various links from different providers may make up a path). This link-by-link performance monitoring data could easily be gathered by looking at a routing protocol's state database (on any router in an area, depending on what is being monitoring and disseminated by the routing protocol), using SNMP [RFC1441] on a per device basis, or in other ways.

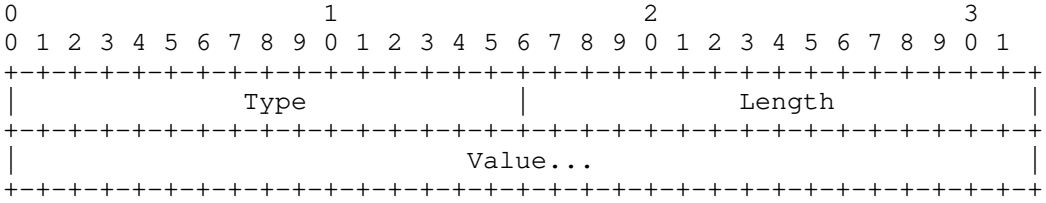
2. Conventions used in this document

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 RFC-2119 [RFC2119].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

3. Express Path Extensions to OSPF TE

The extensions in this document build on the ones provided in OSPF TE (RFC3630) and GMPLS (RFC4203) to permit path selection and network monitoring based on various network performance items. As such, this document proposes new OSPF TE sub-TLVs that can be announced in OSPF TE LSAs. OSPF TE LSAs (RFC3630) are opaque LSAs (RFC5250) with area flooding scope. Each TLV has one or more nested sub-TLVs which permit the TE LSA to be readily extended. There are two main types of OSPF TE LSA; the Router Address or Link TE LSA. Like the GMPLS extensions (RFC4203), this document proposes additional sub-TLVs for the Link TE LSA. As background, all OSPF TE TLVs and sub-TLVs use the same general format (RFC3630):



As per (RFC3630) the Length field defines the length of the value portion of the sub-TLV in octets (thus a TLV with no value portion would have a length of zero). TLVs are padded to four-octet alignment; padding is not included in the length field (so a three octet value would have a length of three, but the total size of the TLV would be eight octets). Unrecognized types are ignored.

OSPF TE Express Path defines several new sub-TLVs. These sub-TLVs fall into 2 distinct categories; "Routine" or "Significant". Routine and Significant sub-TLVs are intended to be used for different purposes (i.e. monitoring or control plane manipulation, respectively). The technical differences between Routine and Significant sub-TLVs are related to the averaging periodicity and announcement frequency of each category of sub-TLV. More information on this subject can be found in section 5.

The following sub-TLVs are defined in OSPF TE Express Path:

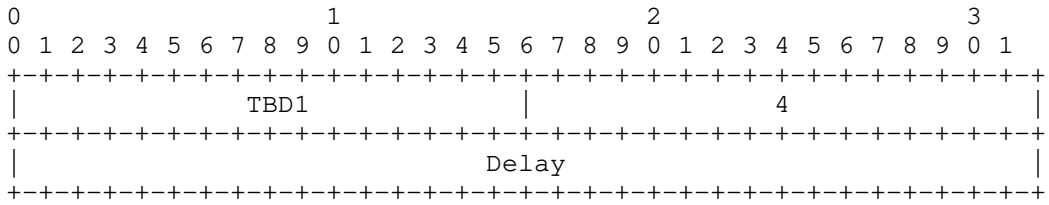
Value	Length	Name

TBD1	4	Routine Unidirectional Link Delay
TBD2	4	Routine Unidirectional Delay Variation
TBD3	4	Routine Unidirectional Link Loss
TBD4	4	Significant Unidirectional Link Delay
TBD5	4	Significant Unidirectional Link Loss

4. Sub TLV Details

4.1. Routine Unidirectional Link Delay Sub-TLV

This TLV advertises the average link delay between two directly connected OSPF neighbors. The delay advertised by this sub TLV MUST be the delay from the local neighbor to the remote one (i.e. the forward path latency). The format of this sub-TLV is shown in the following diagram:



4.1.1. Type

This sub-TLV has a type of TBD1

4.1.2. Length

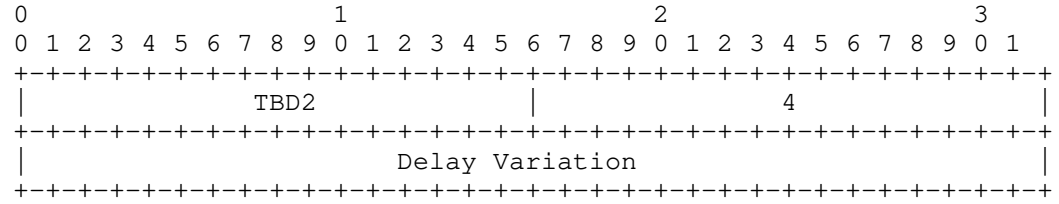
The length is 4

4.1.3. Delay Value

This field carries the average link delay over a configurable interval in micro-seconds, encoded as an IEEE floating point single precision value.

4.2. Routine Unidirectional Delay Variation Sub-TLV

This TLV advertises the average link delay variation between two directly connected OSPF neighbors. The delay variation advertised by this sub-TLV MUST be the delay from the local neighbor to the remote one (i.e. the forward path latency). The format of this sub-TLV is shown in the following diagram:



4.2.1. Type

This sub-TLV has a type of TBD2

4.2.2. Length

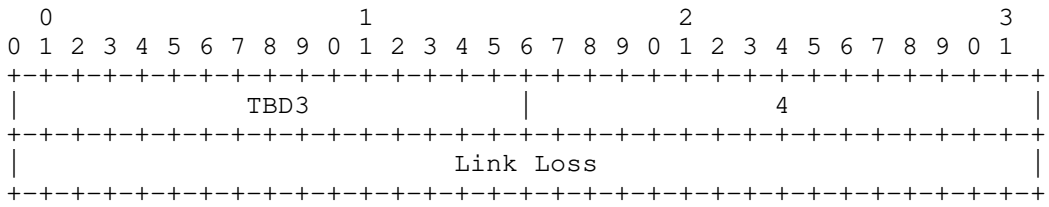
The length is 4

4.2.3. Delay Variation

This field carries the average link delay variation over a configurable interval in micro-seconds, encoded as an IEEE floating point single precision value.

4.3. Routine Unidirectional Link Loss Sub TLV

This TLV advertises the loss (as a packet percentage) between two directly connected OSPF neighbors. The link loss advertised by this sub-TLV MUST be the packet loss from the local neighbor to the remote one (i.e. the forward path loss). The format of this sub-TLV is shown in the following diagram:



4.3.1. Type

This sub-TLV has a type of TBD3

4.3.2. Length

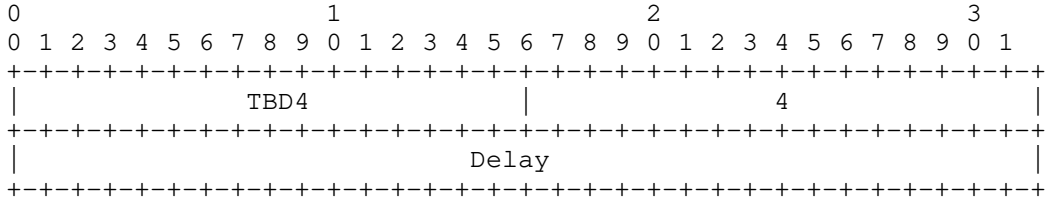
The length is 4

4.3.3. Link Loss

This field carries the link packet loss as a percentage of the total traffic sent over a configurable interval, encoded as an IEEE floating point single precision value.

4.4. Significant Unidirectional Link Delay Sub-TLV

This TLV advertises the average link delay between two directly connected OSPF neighbors. This TLV is announced when either a configurable maximum average delay or a configurable reuse delay threshold is passed. The delay advertised by this sub TLV MUST be the delay from the local neighbor to the remote one (i.e. the forward path latency). The format of this sub-TLV is shown in the following diagram:



4.4.1. Type

This sub-TLV has a type of TBD4

4.4.2. Length

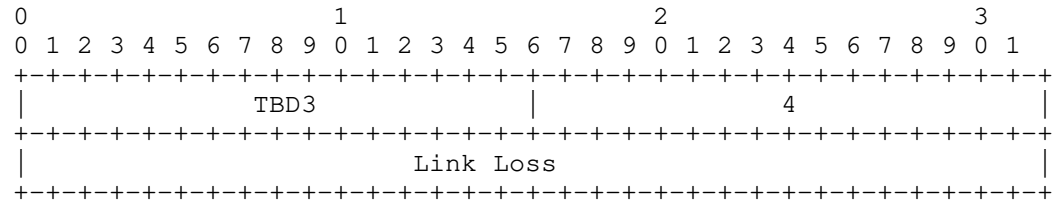
The length is 4

4.4.3. Delay Value

This field carries the average link delay over a configurable interval in micro-seconds, encoded as an IEEE floating point single precision value.

4.5. Significant Unidirectional Link Loss Sub TLV

This TLV advertises the loss (as a packet percentage) between two directly connected OSPF neighbors. This TLV is announced when either a configurable loss threshold or a configurable loss reuse threshold is passed. The link loss advertised by this sub-TLV MUST be the packet loss from the local neighbor to the remote one (i.e. the forward path loss). The format of this sub-TLV is shown in the following diagram:



4.5.1. Type

This sub-TLV has a type of TBD5

4.5.2. Length

The length is 4

4.5.3. Link Loss

This field carries the link packet loss as a percentage of the total traffic sent over a configurable interval, encoded as an IEEE floating point single precision value.

5. Announcement Periodicity

Routine announcements are intended to announce data for trending applications (e.g. advertising small variations in performance occurring over a longer period of time). Significant sub-TLVs are intended to announce the occurrence of more dramatic events that affect network performance (e.g. protection switching). A primary function of Significant sub-TLVs are to manipulate the control plane.

Since Routine and Significant sub-TLVs have generally different goals, implementations SHOULD permit them to be announced using different thresholds and filtering (i.e. rolling average) parameters.

6. Announcement Suppression

Implementations MAY suppress Routine announcements when performance metrics averages do not change by more than a certain amount. These suppression thresholds SHOULD be configurable.

Significant announcements MUST only be sent when configurable thresholds are surpassed.

7. Compatibility

As per (RFC3630), unrecognized TLVs should be silently ignored

8. Security Considerations

This document does not introduce security issues beyond those discussed in [RFC3630] and [RFC5329].

9. IANA Considerations

IANA maintains the registry for the sub-TLVs. OSPF TE Express Path will require one new type code per sub-TLV defined in this document.

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.

10.2. Informative References

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

[RFC3031] Rosen, E., Viswanathan, A., Callon, R., "Multiprotocol Label Switching Architecture", January 2001

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

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

[RFC5250] Berger, L., Bryskin I., Zinin, A., Coltun, R., "The OSPF Opaque LSA Option", RFC 5250, July 2008.

11. Acknowledgments

The authors would like to recognize Ayman Soliman for his contributions.

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

12. Author's Addresses

Spencer Giacalone
Thomson Reuters
195 Broadway
New York NY 10007, USA

Email: Spencer.giacalone@thomsonreuters.com

Dave Ward
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089, USA

Email: dward@juniper.net

John Drake
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089, USA

Email: jdrake@juniper.net

Alia Atlas
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089, USA

Email: akatlas@juniper.net

OSPF Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 23, 2011

M. Bhatia
Alcatel-Lucent
V. Manral
IP Infusion
A. Lindem
Ericsson

February 19, 2011

Supporting Authentication Trailer for OSPFv3
draft-ietf-ospf-auth-trailer-ospfv3-03

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 August 23, 2011.

Copyright Notice

Copyright (c) 2011 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

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.

[RFC4552] 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.

[RFC4552] 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 [RFC4552] 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 [RFC6039].

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].

1.1. Requirements Section

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

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

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.

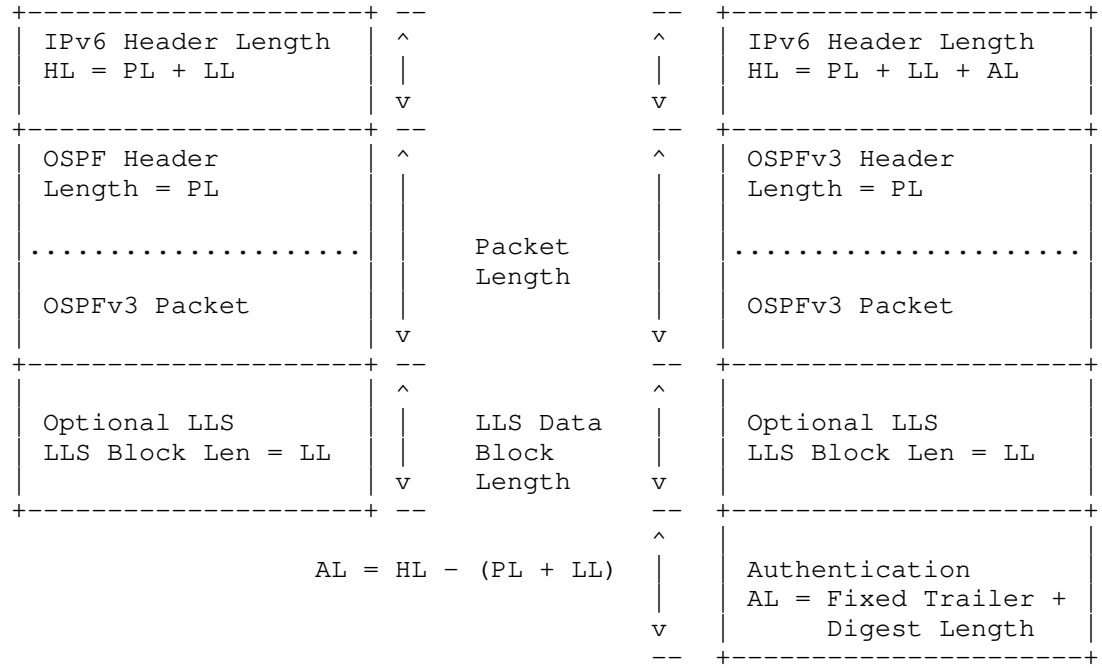


Figure 1: Authentication Trailer in OSPFv3

The presence of the Link Local Signaling (LLS) [RFC5613] block, is determined by the L-bit setting in OSPFv3 options field in OSPFv3 Hello and Database Description packets. If present, the LLS block is included along with the OSPFv3 packet in the cryptographic authentication computation.

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. For OSPFv3 Hello and Database Description packets, the AT-bit indicates the AT is present. For other OSPFv3

packet types, the OSPFv3 AT bit setting is preserved from the OSPFv3 Hello/Database Description setting.

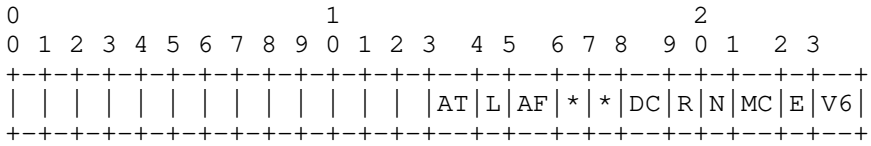


Figure 2: OSPFv3 Options Field

The AT-bit must be set in all OSPFv3 Hello and Database Description packets that contain an authentication trailer.

2.2. Basic Operation

The procedure followed for computing the Authentication Trailer is much same as described in [RFC5709] and [RFC2328]. One difference is that the LLS block, if present, is included in the cryptographic authentication computation.

The way the authentication data is carried in the Authentication Trailer is very similar to how it is done in case of [RFC2328]. The only difference between the OSPFv2 authentication trailer and the OSPFv3 authentication trailer is that information in addition to the message digest is included.

Consistent with OSPFv2 cryptographic authentication [RFC2328], both OSPFv3 header checksum calculation and verification are omitted when the OSPFv3 authentication mechanisms described in this specification are used.

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:

- Security Association Identifier (SA 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 SA 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 SA IDs makes changing keys while maintaining protocol operation convenient. Each SA 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 SA ID can indicate a key with a different authentication protocol. This allows the introduction of new authentication mechanisms without disrupting existing OSPFv3 adjacencies.

- Authentication Algorithm

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

  Currently, the following algorithms are supported:

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

- 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:

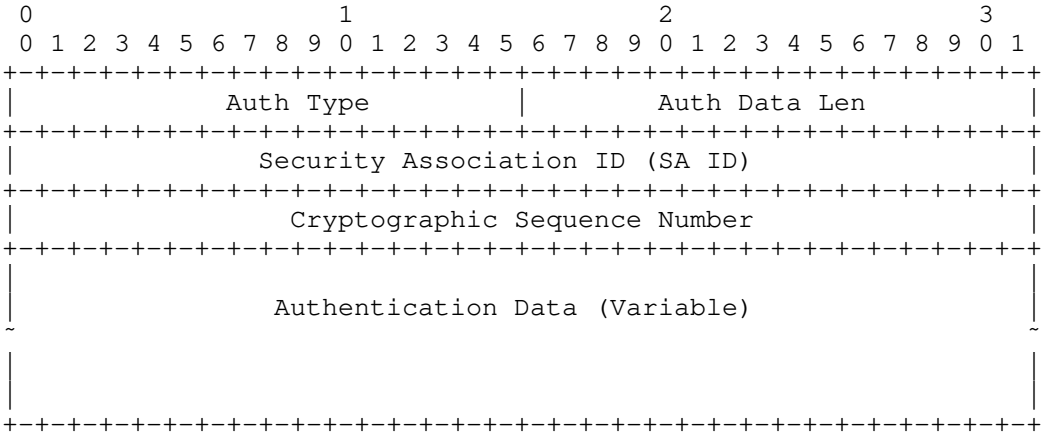


Figure 3: Authentication Trailer Format

The various fields in the Authentication trailer are:

- Auth Type

  16-bit field identifying the type of authentication. The following values are defined in this specification:

  0 - Reserved.
  1 - Cryptographic Authentication as described herein.

- Auth Data Len

  The length in bytes of the message digest appended to the OSPF packet.

- Security Association Identifier (SA ID)

  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.

- Cryptographic Sequence Number

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

- Authentication Data

  Variable data that is carrying the digest for the protocol packet and optional LLS block.

4.2. OSPFv3 Header Checksum

Both OSPFv3 header checksum calculation and verification are omitted when the OSPFv3 authentication mechanisms described in this specification are used. This implies:

- For OSPFv3 packets to be transmitted, the OSPFv3 header checksum computation is omitted and the OSPFv3 header checksum SHOULD be set to 0 prior to computation of the OSPFv3 Authentication Trailer message digest.

- For received OSPFv3 packets including an OSPFv3 Authentication Trailer, OSPFv3 header checksum verification MUST be omitted.

4.3. Cryptographic Authentication Procedure

As noted earlier, the SA ID implicitly specifies the algorithms used to generate and verify the message digest. 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.

4.4. 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:

First-Hash = H(Ko XOR Ipad || (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 and, if present, the LLS block[RFC5613].

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 both the optional LLS block and 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.

4.5. 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.

The Authentication Trailer (AT) is accessed using the OSPFv3 packet header length to access the data after the OSPFv3 packet and, if an LLS Data Block [RFC5613] is present, using the LLS Data Block Length to access the data after the LLS Data Block. The L-bit in the OSPFv3 options in Hello and Database Description packets is examined to determine if an LLS Data Block is present. If an LLS block is present (as specified by the L-bit), it is included along with the OSPFv3 Hello or Database Description packet in the cryptographic authentication computation.

Due to the placement of the AT following the LLS block and the fact that the LLS block is included in the cryptographic authentication computation, OSPFv3 routers supporting this specification MUST minimally support examining the L-bit in the OSPFv3 options and using the length in the LLS block to access the AT. It is RECOMMENDED that OSPFv3 routers supporting this specification fully support OSPFv3 Link Local Signaling, [RFC5613].

If usage of the Authentication Trailer (AT), as specified herein, is configured for an OSPFv3 link, OSPFv3 Hello and Database Description packets with the AT-bit clear in the options will be dropped. All OSPFv3 packet types will be dropped if AT is configured for the link and the IPv6 header length is less than the amount necessary to include an 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.

5. Migration and Backward Compatibility

In general, all OSPFv3 routers participating on a link should be migrated to OSPFv3 Authentication at the same time. As with OSPFv2 authentication, a mismatch in the SA ID, Authentication Type, or message digest will result in failure to form an adjacency. For multi-access links, communities of OSPFv3 routers could be migrated using different interface instance IDs. However, at least one router would need to form adjacencies between both the OSPFv3 routers including and not including the authentication trailer. This would result in sub-optimal routing, as well as, added complexity and is only recommended in cases where authentication is desired on the link and it isn't feasible to migrate all the routers on the link at the same time.

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.

6. Security Considerations

The document proposes extensions to OSPFv3 which would make it more secure than [RFC5340]. 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.

7. IANA Considerations

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

IANA is also requested to create new OSPFv3 "Authentication Trailer Types Registry"

Value/Range	Designation	Assignment Policy
0	Reserved	Reserved
1	Cryptographic Authentication	Already assigned
2-65535	Unassigned	Standards Action

OSPFv3 Authentication Types

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.

[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.

8.2. 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.

[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.

[RFC4552] Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, 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.

[RFC6039] Manral, V., Bhatia, M., Jaeggli, J., and R. White, "Issues with Existing Cryptographic Protection Methods for Routing Protocols", RFC 6039, October 2010.

Appendix A. Acknowledgments

First and foremost, thanks to the authors of RFC5709[RFC5709] from which this work was derived.

Thanks to Michael Barnes for numerous comments and strong input on the coverage of LLS by the Authentication Trailer (AT).

Thanks to Rajesh Shetty for numerous comments including the suggestion to include an Authentication Type field in the Authentication Trailer for extendibility.

Thanks to Srinivasan K L, Shraddha H, Alan Davey, and Glen Kent for their review comments.

Thanks to Alan Davey, Russ White, Stan Ratliff, and others for their support of the draft.

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

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 S. Kini
Internet-Draft W. Lu
Intended Status: Standards Track A. Tian
Expires: September 2011 Ericsson
March 14, 2011

OSPF Fast Notification
draft-kini-ospf-fast-notification-01.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/1id-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 September 15, 2011.

Copyright Notice

Copyright (c) 2011 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

The current OSPF link-state database (LSDB) flooding mechanism involves control-plane processing and forwarding at each hop. The delay due to the involvement of the control-plane adversely affects OSPF convergence. This document describes a mechanism to transmit link-state advertisements (LSA) multiple hops away without involving control-plane processing at the intermediate routers. This helps to achieve faster convergence. It complements the current LSDB flooding mechanism.

Table of Contents

1. Introduction

The LSDB flooding mechanism of OSPF is described in [OSPF]. On receiving a LSA from an adjacent neighbor, the router performs several consistency checks and also compares it with the LSA instance in its LSDB to determine the more recent version. The next step in the flooding procedure involves sending the LSA to its adjacent neighbors and that includes acknowledgements and retransmission to ensure reliability. These procedures involve the control-plane and are therefore gated by the processing and forwarding speed of the control-plane at each hop.

The solution described in this document does not involve control-plane processing at the intermediate nodes. Details are provided in section 3.

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. Solution

The document describes a mechanism to quickly notify a LSA to all the routers within an OSPF area. This does not require control-plane processing at an intermediate hop and such a mechanism is also referred to as a fast notification (FN) as described in [FN-FRWK]. The solution in this document uses such a fast-notification mechanism and makes it work with the OSPF procedures (including its flooding mechanism) and is henceforth referred to as OSPF-FN. OSPF-FN does not aim to replace the current OSPF flooding mechanism. The details of this solution are described below.

A variety of mechanisms to encode and transport FN messages are described in [FN-TRNS]. For message encoding, this document uses the minimal extra encapsulation i.e, it uses the multicast FN address 'MC-FN' as the destination address of a OSPF-FN link-state update (LSU) packet. Note that when the redundant-tree mode as described in [FN-TRNS] is used, there will be two such multicast addresses. Such an LSU packet is henceforth referred to as the OSPF-FN-LSU packet. This LSU packet contains LSAs that need to be quickly notified to all routers within the OSPF area. At this time the redundant-tree mode of transporting OSPF-FN-LSU packets is the preferred method.

OSPF-FN works in conjunction with the OSPF flooding mechanism as follows. An OSPF router that originates a LSA that is determined to require FN, creates a OSPF-FN-LSU packet containing the LSA and

transports it using a method chosen from [FN-TRNS]. Note that this is in addition to any OSPF procedures and specifically does not change the current flooding mechanism in OSPF. If the redundant tree mode is used then an OSPF-FN-LSU packet is sent on each of the redundant trees. When a LSA is received through a OSPF-FN-LSU packet, the normal OSPF procedures on receiving a LSU packet should be executed with the exception that there is no acknowledgement for the LSU packet. Also, the LSA is noted as received by OSPF-FN and the older instance of the LSA MUST be retained. The older LSA instance is discarded only if the current OSPF flooding mechanism confirms that the LSA from the OSFP-FN-LSU is the correct instance. This confirmation should occur within a short time of receiving the OSPF-FN-LSU packet. This time period is henceforth called T-discard-FN-LSA time and has a recommended default of 5 seconds. If this timer expires then the LSA received via the OSPF-FN-LSU packet is discarded and the older instance is used as the correct instance.

After updating the LSDB with the LSA received from the OSPF-FN-LSA as described above, all the route processing is performed except that the new and changed routes are not activated in the forwarding plane. The activation is done on receiving the the LSA via the current OSPF flooding procedures.

It is desirable for OSPF-FN-LSU packets to use area-wide authentication parameters so that OSPF-FN-LSU messages can be forwarded by intermediate routers similar to any normal data packet. The LSA sequence number can provide protection against replay attacks. A separate per-link (or network) authentication parameter introduces the complexity of rewriting the packet at each hop as described in [FN-TRNS]. This is an area for further study.

The determination of which LSAs require a fast notification is dependent on the event that caused the LSA update. The benefits of using FN for a link or adjacency going down is straightforward. Other cases require further study.

4. Security Considerations

Area wide authentication (if used) parameters could bring associated security concerns. This is an area for further study.

5. IANA Considerations

A TLV id to identify this capability and multicast IP addresses to transport the OSPF-FN-LSU messages are required.

6. References

6.1. 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.

[FN-FRWK] Lu, W., et al, "Fast Notification Framework", draft-lu-fast-notification-framework-01 (Work in progress), March 2011.

[FN-TRNS] Lu, W., et al, "Transport of Fast Notfication Messages", draft-lu-fn-transport-00 (Work in progress), March 2011.

7. Acknowledgements

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

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

OSPF
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2011

W. Lu
Ericsson

October 18, 2010

OSPF TE Extension for Area IDs
draft-lu-ospf-area-tlv-00

Abstract

For multi-area path computation, it is desirable to have the knowledge of the boarder areas and the corresponding boarder routers. This memo defines a TLV to the OSPF TE extensions to meet such need.

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 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. 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

The Traffic Engineering Database (TED) based on OSPF is sufficient for the intra-area path computation. However because TED is of area scope, the path computation cannot be used for inter-area scenarios without the help of area boarder routers.

Although the Router LSA offers a B bit to signify an ABR router, the identity of the attached area is unkown. This will force a router to contact every ABRs if it wants TED info from each area.

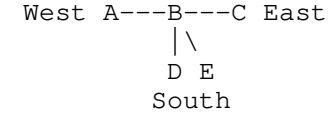


Figure 1: Sample Topology

For example, as shown in Figure 1, the router B has three neighbor areas West, East, and South. It can reach these areas through ABRs A, C, D and E. Both D and E connect to the area South. Ideally B only needs to contact A, c, D (or E) to obtain TED info of the three areas. However, since it does not know that D and E share the boarder between this area and the area South, it has to blindly send the request to both D and E.

If instead each ABR provides its exit area's information, such as A(West), C(East), D(South), E(South), the router B will be able to make a sound decision to utilize only three ABRs. For this purpose we define an area ID TLV, detailed in Section 2.

Moreover since the TLV is for TE purpose, it is added under the OSPF TE LSA as defined in OSPF TE Extensions [RFC3630].

1.1. Requirements Language

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

1.2. Acronyms

ABR - Area Boarder Router

TE - Traffice Engineering

TED - Traffice Engineering Database

PCE - Path Computation Element

LSP - Label Switched Path

OSPF - Open Shortest Path First

LSA - Link State Advertisement

2. Area ID TLV

[RFC3630] section 2.4 defined two TLVs. This memo adds a third TLV, the Area ID TLV.

The area ID TLV is type TBD (suggest 3), has a length of 4, and a value that is the four octet integer. It may have zero or more occurances in one Traffic Engineering LSA originated by a router.

The value is the area ID of an exit area for which the ABR has TE enabled. An ABR may join multiple areas. Therefore it may generate m-1 area ID TLVs, where m is the total number of areas the router joins. For a non-ABR router, it does not have any exit area, hence its TE TLV has zero occurance of the area ID TLV.

3. Applications

With the area ID TLVs in TED, when performing inter-area path computation, a PCE [RFC4655] will gain additional knowledge of the surrounding areas and the boarder routers to reach each area. The PCE may elect one boarder router for each area and request TE info from it.

Alternatively, the PCE may relay the path computation job to the PCE which is also an ABR.

4. Acknowledgements

TBD

5. IANA Considerations

This document defines the following TLV to the OSPF TE Extensions under TE LSA:

Type	Name	Source
TBD (recommend 3)	Area ID TLV	This document

6. Security Considerations

There are no specific security considerations within the scope of this document.

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.

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

7.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

Author's Address

Wenhu Lu
Ericsson
300 Holger Way
San Jose, California 95134
USA

Phone: 408 750-5436
Email: wenhu.lu@ericsson.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: June 3, 2011

A. Retana
Cisco Systems, Inc.

November 30, 2010

OSPF Incremental Link State Database Synchronization
draft-retana-ospf-ils-00

Abstract

The ability of OSPF to transport non-routing information to be used by other applications was defined by the Opaque LSA Option. In order to not impact the convergence of routing information, this document describes a simple process to incrementally synchronize the routing and non-routing information residing in an OSPF link-state database.

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 June 3, 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

Opaque LSAs [RFC5250] provide the ability for OSPFv2 [RFC2328] to transport non-routing information to be used by other applications. A similar capability exists in OSPFv3 [RFC5340] through the use of the U-bit and an appropriate LSA Function Code. Throughout this document Opaque LSAs and ones with unrecognized link-state types will be referred to simply as "opaque".

The presence of opaque information in the OSPF Link-State Database (LSDB) may result in longer database exchange times, especially in cases where the amount of data is significantly larger than the routing-specific information. In order to not impact the convergence of routing information, this document describes a simple process to incrementally synchronize the information residing in an OSPF LSDB. The process uses existing mechanisms.

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. Incremental LSDB Synchronization Process

The Incremental LSBD Synchronization (ILS) process consists of the following steps:

LSA Prioritization
: The contents of the local LSDB are classified to determine which LSAs require prioritized synchronization.

  In general, LSAs containing routing-specific information SHOULD be classified as requiring prioritized synchronization, while other LSAs MAY be classified as not requiring it.

Prioritized LSDB Synchronization
: This step corresponds to the adjacency establishment process as described in RFC 2328 [RFC2328].

  LSAs classified as not requiring prioritized synchronization MUST NOT be included in Database Description (DBD) Packets during the Database Exchange Process. The OSPF routing table structure SHOULD be calculated before moving on to the next step.

Final LSDB Synchronization
: In this step, any remaining LSAs in the LSDB SHOULD be synchronized. The routers MUST use the Out-of-Band LSDB Resynchronization [RFC4811] (OOB Resync) mechanism, which provides a way to resynchronize the LSDB without affecting the advertised neighbor state.

The process is described in terms of LSAs containing (or not) routing-specific information, but it may be generalized to include any other criteria considered significant in the local network and protocol instance.

The last step in the process MAY be used recursively to achieve an incremental LSDB synchronization through different types of data, making it also applicable to environments where only non-routing information exists.

3.1. Graceful Restart

The restart of the OSPF software in a router also presents an opportunity for LSDB synchronization. Because the restarting router is still in the forwarding path, it is important for the routing information in the LSDB to be synchronized as fast as possible. ILS can be used, with minor modifications, to reduce the synchronization time and the probability of network topology changes.

Graceful OSPF Restart
: Graceful OSPF Restart for OSPFv2 [RFC3623] and OSPFv3 [RFC5187] don't specify any changes to the adjacency establishment process.

  ILS can be used by the Helper Neighbor during the Grace Period; if so, then the Helper Node MUST include any Grace-LSAs in the DBD Packets during the Prioritized LSDB Synchronization step.

OSPF Restart Signaling
: OSPF Restart Signaling [RFC4812] defines a mechanism to inform neighbors about a local restart, in which the LSDB synchronization is achieved using OOB Resync. In other words, the Prioritized LSDB Synchronization step would use OOB Resync if the non-restarting router uses ILS. No other changes to the process are needed.

4. Backward Compatibility

The operation of ILS depends on the support of OOB Resync during synchronization; no backwards compatibility issues exist there [RFC4811]. If OOB Resync is not supported by one of the routers, then the LSDB synchronization would fall back to the adjacency establishment process as described in RFC 2328 [RFC2328].

If OOB Resync is supported, but ILS has not been implemented by all the routers involved, the operation is still backwards compatible. Note that the process (Section 3) depends on the database description by the local router. In other words, a router may decide to not fully describe the contents of its LSDB to its neighbor during the adjacency establishment process, and later use OOB Resync to incrementally describe the difference; the receiver doesn't need to be aware of ILS. The benefits of ILS may only be partially realized if not supported by all the routers involved in synchronization.

5. IANA Considerations

This memo includes no request to IANA.

6. Security Considerations

The process described in this document does not introduce any new security issues into the OSPF protocol.

7. Acknowledgements

The author would like to thank Abhay Roy and Liem Nguyen for their comments, and Dimitri Papadimitriou for his comments and for providing the motivation for this document.

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.

[RFC4811] Nguyen, L., Roy, A., and A. Zinin, "OSPF Out-of-Band Link State Database (LSDB) Resynchronization", RFC 4811, March 2007.

8.2. Informative References

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

[RFC3623] Moy, J., Pillay-Esnault, P., and A. Lindem, "Graceful OSPF Restart", RFC 3623, November 2003.

[RFC4812] Nguyen, L., Roy, A., and A. Zinin, "OSPF Restart Signaling", RFC 4812, March 2007.

[RFC5187] Pillay-Esnault, P. and A. Lindem, "OSPFv3 Graceful Restart", RFC 5187, June 2008.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, July 2008.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

Author's Address

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

Phone: +1 919 392 2061
Email: aretana@cisco.com

Network Working Group
Internet-Draft
Updates: RFC5820
(if approved)
Intended status: Experimental
Expires: September 8, 2011

A. Retana
S. Ratliff
Cisco Systems, Inc.

March 7, 2011

Use of the OSPF-MANET Interface in Single-Hop Broadcast Networks
draft-retana-ospf-manet-single-hop-00

Abstract

This document describes the use of the OSPF-MANET interface in single-hop broadcast networks. It includes a mechanism to deterministically reduce the number of adjacencies using Smart Peering and other considerations due to the nature of the network.

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 September 8, 2011.

Copyright Notice

Copyright (c) 2011 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

The OSPF-MANET interface [RFC5820] uses the point-to-multipoint adjacency model over a broadcast media to allow the following:

- all router-to-router connections are treated as if they were point-to-point links.
- Link metric can be set on a per-neighbor basis.
- Broadcast and multicast can be accomplished through the Layer 2 broadcast capabilities of the media.

It is clear that the characteristics of the MANET interface can also be beneficial in fixed network deployments; specifically in single-hop broadcast capable networks which may have a different cost associated with any pair of nodes.

This document describes the use of the MANET interface in single-hop broadcast networks.

1.1. Single-Hop Broadcast Networks

The OSPF extensions for MANET networks assume the ad-hoc formation of a network over bandwidth-constrained wireless links, where packets may traverse several intermediate nodes before reaching their destination (multi-hop paths on the interface). By contrast, a single-hop broadcast network (as considered in this document) is one that is structured in such a way that all the nodes in it are directly connected to each other. An Ethernet interface is a good example of the connectivity model.

Furthermore, the single-hop networks considered may have different link metrics associated to the connectivity between a specific pair of neighbors. The OSPF broadcast model [RFC2328] can't accurately describe these differences. A point-to-multipoint description is more appropriate given that each node can reach every other node directly.

In summary, the single-hop broadcast interfaces considered in this document have the following characteristics:

- direct connectivity between all the nodes
- different link metrics may exist per-neighbor
- it has broadcast/multicast capabilities

1.2. MANET Interface Considerations

The operation of the MANET interface doesn't change when implemented on a single-hop broadcast interface. However, some of the proposed enhancements are not needed; explicitly, Incremental Hellos and Overlapping Relays are not required due to the connectivity model. If Overlapping Relays are used, then the A-bit SHOULD NOT be set by any of the nodes: the result is an empty set of Active Overlapping Relays.

Smart Peering can be used to reduce the burden of requiring a full mesh of adjacencies. In short, a new adjacency is not required if reachability to the node is already available through the existing STP. In general, the reachability is verified on a first-come-first-served basis; i.e. in a typical network, the neighbors with which a FULL adjacency is set up depend on the order of discovery. Section 3 explains the use of Router Priority to create a deterministic mechanism to select which nodes to form FULL adjacencies with.

Section 4 explains the operation with unsynchronized adjacencies.

The operation described in this document uses already defined mechanisms and requires no additional on-the-wire changes.

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].

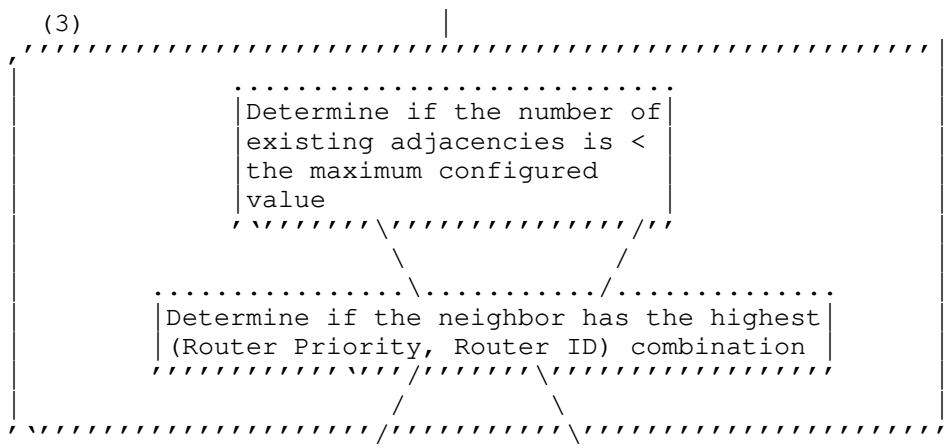
3. Use of Router Priority

The Smart Peering state machine [RFC5820] allows for the definition of heuristics, beyond the SPT reachability, to decide whether or not it considers a new adjacency to be of value. This section describes one such heuristic to be used in Step (3) of the state machine.

The Router Priority (as defined in OSPFv2 [RFC2328] and OSPFv3 [RFC5340]) is used in the election of the (Backup) Designated Router, and can be configured only in broadcast and NBMA interfaces. The MANET interface is a broadcast interface using the point-to-multipoint adjacency model, which means that no (Backup) Designated Router is elected. For its use with the MANET interface, the Router Priority is defined as:

Router Priority
: An 8-bit unsigned integer. Used to determine the precedence of which router(s) to establish a FULL adjacency with during the Smart Peering selection process. When more than one router attached to a network is present, the one with the highest Router Priority takes precedence. If there is still a tie, the router with the highest Router ID takes precedence.

The heuristic for the smart peering state machine is described as:



Smart Peering Algorithm

In order to avoid churn in the selection and establishment of the adjacencies, every router SHOULD wait Wait Time [RFC2328] before running the Smart Peering state machine. Note that this wait should cause the selection process to consider all the nodes on the link, instead of being triggered based on receiving a Hello message from a potential neighbor. The nodes selected using this process are referred to simply as Smart Peers.

It is RECOMMENDED that the maximum number of adjacencies be configured to at least 2.

4. Unsynchronized Adjacencies

An unsynchronized adjacency [RFC5820] is one for which the database synchronization is postponed, but that is announced as FULL because SPT reachability can be proven. A single-hop broadcast network has a connectivity model in which all the nodes are directly connected to each other. This connectivity results in a simplified reachability check through the SPT: the adjacency to a specific peer MUST be

advertized as FULL by at least one Smart Peer.

The single-hop nature of the interface allows then the advertisement of the reachable adjacencies as FULL without additional signaling. Flooding SHOULD be enabled for all the unsynchronized adjacencies to take advantage of the broadcast nature of the media. As a result, all the nodes in the interface will be able to use all the LSAs received.

5. IANA Considerations

This document includes no request to IANA.

6. Security Considerations

No new security concerns beyond the ones expressed in RFC 5820 [RFC5820] are introduced in this document. In fact, due to the application in fixed networks, some of the concerns may actually be reduced.

7. Acknowledgements

The authors would like to thank Anton Smirnov for his comments.

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.

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

8.2. Informative References

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

[RFC5820] Roy, A. and M. Chandra, "Extensions to OSPF to Support Mobile Ad Hoc Networking", RFC 5820, March 2010.

Authors' Addresses

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

Email: aretana@cisco.com

Stan Ratliff
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
USA

Email: sratliff@cisco.com

Network Working Group
Internet-Draft
Obsoletes: RFC3137
(if approved)
Intended status: Informational
Expires: August 29, 2011

A. Retana
L. Nguyen
R. White
A. Zinin
Cisco Systems, Inc.
D. McPherson
Verisign, Inc.

February 25, 2011

OSPF Stub Router Advertisement
draft-retana-ospf-rfc3137bis-00

Abstract

This memo describes a backward-compatible technique that may be used by OSPF (Open Shortest Path First) implementations to advertise unavailability to forward transit traffic or to lower the preference level for the paths through such a router. In some cases, it is desirable not to route transit traffic via a specific OSPF router. However, OSPF does not specify a standard way to accomplish this.

Status of this Memo

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

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

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

This Internet-Draft will expire on August 29, 2011.

Copyright Notice

Copyright (c) 2011 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

In some situations, it may be advantageous to inform routers in a network not to use a specific router as a transit point, but still route to it. Possible situations include the following:

- o The router is in a critical condition (for example, has very high CPU load or does not have enough memory to store all LSAs or build the routing table).
- o Graceful introduction and removal of the router to/from the network.
- o Other (administrative or traffic engineering) reasons.

Note that the proposed solution does not remove the router from the topology view of the network (as could be done by just flushing that router's router-LSA), but prevents other routers from using it for transit routing, while still routing packets to the router's own IP addresses, i.e., the router is announced as a stub.

It must be emphasized that the proposed solution provides real benefits in networks designed with at least some level of redundancy so that traffic can be routed around the stub router. Otherwise, traffic destined for the networks reachable through such a stub router will be still routed through it.

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. Proposed Solution

The solution described in this document solves two challenges associated with the outlined problem. In the description below, router X is the router announcing itself as a stub.

1) Making other routers prefer routes around router X while performing the Dijkstra calculation.

2) Allowing other routers to reach IP prefixes directly connected to router X.

Note that it would be easy to address issue 1) alone by just flushing

router X's router-LSA from the domain. However, it does not solve problem 2), since other routers will not be able to use links to router X in Dijkstra (no back link), and because router X will not have links to its neighbors.

To address both problems, router X announces its router-LSA to the neighbors with the costs of all non-stub links (links of the types other than 3) set to LSInfinity (16-bit value 0xFFFF, rather than 24-bit value 0xFFFFFF used in summary and AS-external LSAs).

The solution above applies to both OSPFv2 [RFC2328] and OSPFv3 [RFC5340].

4. Compatibility issues

Some inconsistency may be seen when the network is constructed of the routers that perform intra-area Dijkstra calculation as specified in RFC 1247 [RFC1247] (discarding link records in router-LSAs that have LSInfinity cost value) and routers that perform it as specified in RFC 1583 [RFC1583] and higher (do not treat links with LSInfinity cost as unreachable). Note that this inconsistency will not lead to routing loops, because if there are some alternate paths in the network, both types of routers will agree on using them rather than the path through the stub router. If the path through the stub router is the only one, the routers of the first type will not use the stub router for transit (which is the desired behavior), while the routers of the second type will still use this path.

5. Security Considerations

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

6. Acknowledgements

The authors of this document do not make any claims on the originality of the ideas described. Among other people, we would like to acknowledge Henk Smit for being part of one of the initial discussions around this topic.

We would also like to thank Shishio Tsuchiya, Gunter Van de Velde and Tomohiro Yamagata for reminding us of the need to document the OSPFv3 behavior.

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.

7.2. Informative References

[RFC1247] Moy, J., "OSPF Version 2", RFC 1247, July 1991.

[RFC1583] Moy, J., "OSPF Version 2", RFC 1583, March 1994.

[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.

Appendix A. Changes from RFC 3137

- Edited in support for OSPFv3.
- Updated references and author information.
- Miscellaneous edits.

Authors' Addresses

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

Email: aretana@cisco.com

Liem Nguyen
Cisco Systems, Inc.
3750 Cisco Way
San Jose, CA 95134
USA

Phone: +1 408 527 0670
Email: lhnguyen@cisco.com

Russ White
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
USA

Email: russwh@cisco.com

Alex Zinin
Cisco Systems, Inc.
Capital Tower, 168 Robinson Rd.
Singapore, Singapore 068912
Singapore

Email: azinin@cisco.com

Danny McPherson
Verisign, Inc.
21345 Ridgetop Circle
Dulles, VA 20166
USA

Email: dmcpherson@verisign.com