OSPFv3 Graceful Restart
draft-ietf-ospf-ospfv3-graceful-restart-04.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of 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 November 9, 2006.

Copyright Notice

Copyright (C) The Internet Society (2006).

Abstract

This memo describes the OSPFv3 graceful restart. For OSPFv3, graceful restart is identical to OSPFv2 except for the differences described in this memo. These differences include the format of the grace Link State Advertisements (LSA) and other considerations.
Table of Contents

1. Introduction ............................................. 3
2. Grace Link State Advertisement .......................... 4
   2.1 Grace LSA - LS Type ................................. 4
   2.2 Grace LSA Format .................................... 5
3. Additional Considerations for OSPFv3 Graceful Restart ..... 7
   3.1 Preservation of LSA ID to Prefix Correspondence .... 7
   3.2 Preservation of Interface IDs for Link-LSAs, Network
       LSAs, and Router-LSAs ............................... 7
4. Security Considerations ................................... 8
5. IANA Considerations ...................................... 9
6. Acknowledgments ......................................... 10
7. Normative References ................................. 10
   Authors’ Addresses ................................... 10
   Intellectual Property and Copyright Statements ....... 12
1. Introduction

Graceful OSPF restart [GRACE] describes a mechanism to restart the control plane of an OSPFv2 [OSPFv2] router which still has its forwarding plane intact with a minimum of disruption to the network.

In general, the methods described in [GRACE] work for OSPFv3 [OSPFv3] as well. However, OSPFv3 will use a different grace LSA to signal that a router is (or is about) to attempt a graceful restart. This document describes other OSPFv3 differences as well.

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

Grace-LSAs are originated by an OSPFv3 router that wishes to execute a graceful restart of its OSPFv3 software. A grace-LSA requests that the router’s neighbors aid in its graceful restart by continuing to advertise the router as fully adjacent during a specified grace period. The grace-LSA contains the restarting router grace-period and the reason code indicate the reason for the graceful restart.

In OSPFv3 (refer 2.11 of [OSPFv3]), neighboring routers on a given link are always identified by router ID. This contrasts with the IPv4 behavior where neighbors on point-to-point networks and virtual links are identified by their Router IDs, and neighbors on broadcast, NBMA and point-to-multipoint links are identified by their IPv4 interface addresses. Consequently, there is no requirement for the router-address TLV used for OSPFv3 graceful restart [GRACE].

The grace-LSA body format will remain the same as described in [GRACE].

2.1 Grace LSA - LS Type

A grace-LSA is defined as link-local scope LSA with the LS type equal to 0x000b.

<table>
<thead>
<tr>
<th>LSA function code</th>
<th>LS Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>0x000b</td>
<td>Grace LSA</td>
</tr>
</tbody>
</table>

The U-bit is set to 0 to since this is a link local scoped LSA and the flooding scope is not impacted by whether or not the LSA is known. The S2-bit and S1-bit are also set to 0 to indicate link-local flooding scope.
2.2 Grace LSA Format

The format of a grace LSA format is:

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| LS age              |          11             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Link State ID                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Advertising Router                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    LS sequence number                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                LS checksum            |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             TLVs                              |
|                                                                 |
|                             ...                               |
|                                                               |
```

The Link State ID of a grace-LSA in OSPFv3 is the interface ID of the interface the LSA is originated on.

The length field defines the length of the value portion in octets (thus a TLV with no value portion would have a length of zero). The TLV is 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). Nested TLVs are also 32-bit aligned. For example, a one byte value would have the length field set to 1, and three bytes of padding would be added to the end of the value portion of the TLV. Unrecognized types are ignored.

The format of each TLV is:

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

The format of the TLVs within the body of a grace-LSA is the same as the TLV format used by the Traffic Engineering Extensions to OSPF [OSPF-TE]. The TLV header consists of a 16-bit Type field and a 16-bit length field. The header is followed by zero or more bytes of
value. The length field indicates the length of the value portion in bytes. The value portion is padded to four-octet alignment, but the padding is not included in the length field.

The following is the list of TLVs that can appear in the body of a grace-LSA.

Grace Period (Type=1, length=4). The number of seconds that the router’s neighbors should continue to advertise the router as fully adjacent, regardless of the state of database synchronization between the router and its neighbors. This TLV MUST always appear in a grace-LSA.

Graceful restart reason (Type=2, length=1). Encodes the reason for the router restart, as one of the following: 0 (unknown), 1 (software restart), 2 (software reload/upgrade) or 3 (switch to redundant control processor). This TLV MUST always appear in a grace-LSA.
3. Additional Considerations for OSPFv3 Graceful Restart

This section describes OSPFv3 unique considerations in addition to those described in [GRACE].

3.1 Preservation of LSA ID to Prefix Correspondence

In OSPFv2 there is a direct correspondence between type 3 and type 5 LSA IDs and the prefixes being advertised. For OSPFv3, the LSA ID for inter-area prefix LSAs and external LSAs is simply an unsigned 32 bit integer. To avoid network churn during graceful restart, the restarting router SHOULD preserve the LSA ID to prefix correspondence across graceful restarts.

3.2 Preservation of Interface IDs for Link-LSAs, Network LSAs, and Router-LSAs

The OSPFv3 interface ID, as described in section 3.1.2 [OSPFv3], MUST be preserved by the restarting router across restarts. It is used as the LSA ID for link-LSAs and network-LSAs and is included in the link descriptions in router-LSAs. Failure to preserve interface IDs would result in a mismatch between the restarting router’s pre-restart LSAs and its neighbor adjacency state. This, in turn, would make synchronizing an interface ID change between the restarting router and its helping routers much more difficult (if not impossible) and would most likely result in unreachability or premature graceful restart termination. Placing the burden on the restarting router to preserve interface IDs across restarts provides for a more robust, more deterministic, and simpler mechanism.

Many implementations are using the interface’s MIB-II IfIndex ([INTFMIB]) for Interface ID and are already preserving it across restarts.
4. Security Considerations

This document doesn’t raise any new security concerns other than those covered in [OSPFv3], [OSPFv3-AUTH], and [GRACE]. This is based on the fact that [OSPFv3-AUTH] relies on manually key distribution which precludes the use of replay protection utilizing sequence numbers.
5. IANA Considerations

A new LSA function code will be required for the OSPFv3 grace LSA. Assignment of 0x000b has been suggested herein. Grace LSA TLVs and sub-TLVs will share the same IANA registry as the TLVs and sub-TLVs used by the OSPFv2 grace opaque LSA.
6. Acknowledgments

Many thanks to Kireeti Kompella with whom much of this was discussed. The authors also wish to thank Kunihiro Ishiguro and Vivek Dubey for their comments.

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

7. Normative References


Authors’ Addresses

Padma Pillay-Esnault
Cisco Systems, Inc
3750 Cisco Way
San Jose, CA  95134
USA

Email: ppe@cisco.com
Acee Lindem
Cisco Systems, Inc
7025 Kit Creek Road
Research Triangle Park, NC  27709
USA

Email: acee@cisco.com
Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

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

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

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

Copyright Statement

Copyright (C) The Internet Society (2006). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

Acknowledgment

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