                          OSPFv3 Graceful Restart

Status of This Memo

   This document specifies an Internet standards track protocol for the
   Internet community, and requests discussion and suggestions for
   improvements. Please refer to the current edition of the "Internet
   Official Protocol Standards" (STD 1) for the standardization state
   and status of this protocol. Distribution of this memo is unlimited.

Abstract

   This document describes the OSPFv3 graceful restart. The OSPFv3
   graceful restart is identical to that of OSPFv2 except for the
   differences described in this document. These differences include
   the format of the grace Link State Advertisements (LSAs) and other
   considerations.

Table of Contents

1. Introduction

   Graceful OSPF restart [GRACE] describes a mechanism to restart the
   control plane of an OSPFv2 [OSPFv2] router that 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 grace-LSA with a different
   format to signal that a router is initiating (or is about to
   initiate) 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

   An OSPFv3 router initiating a graceful restart of its OSPFv3 software
   originates grace-LSAs. A grace-LSA requests that the router's
   neighbors aid in its graceful restart by continuing to advertise the
   router as fully adjacent during the specified grace period. The
   grace-LSA contains the restarting router grace-period and the reason
   code indicating the reason for the graceful restart.

   In OSPFv3 (refer to section 2.11 of [OSPFv3]), neighboring routers on
   any link are always identified by their router IDs. This contrasts
   with the OSPFv2 behavior where neighbors on point-to-point networks
   and virtual links are identified by their Router IDs, while neighbors
   on broadcast, Non-Broadcast Multi-Access (NBMA), and point-to-
   multipoint links are identified by their IPv4 interface addresses.
   Consequently, there is no requirement for the router-address TLV
   [GRACE] for OSPFv3 graceful restart.

   The TLV formats of the grace-LSA described in [GRACE] remain
   unchanged.

2.1. Grace LSA - LS Type

   A grace-LSA is defined as an LSA with the LS type equal to 0x000b.
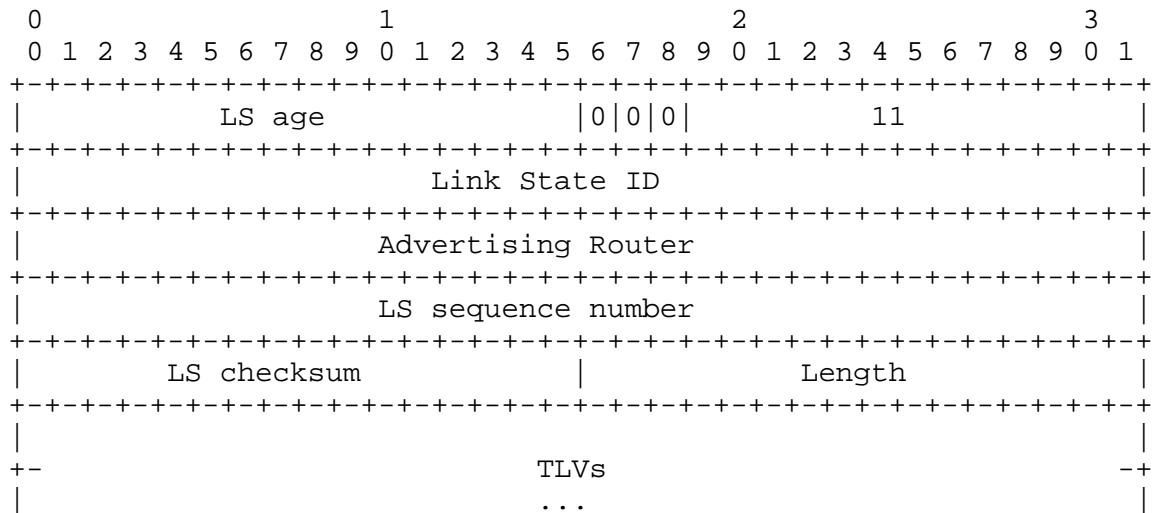
              LSA function code LS Type Description
              ----
              11   0x000b   GRACE-LSA

                    Grace-LSA Type and Function Code

   The S2-bit and S1-bit are set to 0 to indicate link-local flooding
   scope. The U-bit is set to 0 since it isn't applicable to LSAs with
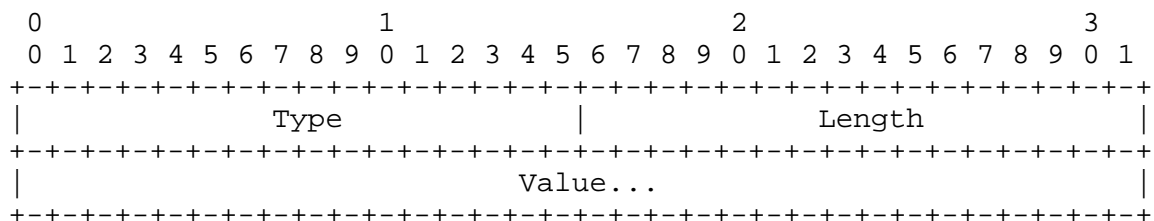   link-local flooding scope.

2.2. Grace LSA Format

   The format of a grace LSA 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              |0|0|0|            11           |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                         Link State ID                         |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                       Advertising Router                      |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                       LS sequence number                      |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |        LS checksum            |              Length           |
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |                                                               |
      +-                             TLVs                            -+
      |                             ...                               |
```

                         Grace-LSA Format

   The Link State ID of a grace-LSA in OSPFv3 is the Interface ID of the
   interface originating the LSA.

   The format of each TLV is:

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

                           TLV Format

   Grace-LSA TLVs are formatted according to section 2.3.2 of [OSPF-TE].

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 summary and
external LSA IDs and the prefixes being advertised. However, in
OSPFv3, the LSA ID for inter-area prefix LSAs and external LSAs is
simply an unsigned 32-bit integer. Hence, to avoid network churn
during graceful restart, the restarting router MUST preserve the LSA
ID to prefix correspondence across graceful restarts.

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

In OSPFv3, the LSA ID for Link-LSAs and Network-LSAs and link
descriptions in Router-LSAs map to their corresponding Interface ID.
Changes in the Interface ID during graceful restart will result in a
mismatch between the restarting router's pre-restart LSAs and its
neighbor adjacency state. These disparities will cause the graceful
restart to terminate prematurely.

Synchronizing Interface ID changes between neighbors is possible.
However, placing the burden on the restarting router to preserve
Interface IDs across restarts provides for a more robust, more
deterministic, and simpler mechanism. Therefore, the OSPFv3
Interface ID, as described in section 3.1.2 of [OSPFv3], MUST be
preserved by the restarting router across restarts.

Many implementations currently use the interface's MIB-II IfIndex
[MIB-INTF] for Interface ID. The persistence of Interface ID across
reboots is described in section 3.1.5 of [MIB-PERS].

4. Security Considerations

   [OSPFv3-AUTH] relies on manual key distribution which precludes the
   use of replay protection that utilizes sequence numbers. The replay
   of an OSPF Link-Update containing a grace-LSA would allow an attacker
   to deceive neighboring routers into believing that a router that has
   been taken out of service (either intentionally or via a malicious
   action by the same attacker) is still active and is in the process of
   graceful restart. However, this attack is much more difficult than
   the obvious replay of standard OSPFv3 hello packets to accomplish the
   same thing by keeping the adjacency up. Since hello packets are sent
   more predictably and knowledge of the key is not required, the risk
   added by OSPFv3 graceful restart is insignificant. Hence, this
   document does not raise any new security concerns other than those
   covered in [OSPFv3], [OSPFv3-AUTH], and [GRACE].

5. IANA Considerations

   A new LSA function code has been assigned for the OSPFv3 grace-LSA.
   The assignment of 0x000b has been made in the "OSPFv3 LSA Function
   Codes" sub-registry of the "Open Shortest Path First v3 (OSPFv3)
   Parameters" registry. OSPFv3 grace-LSA TLVs and sub-TLVs use the
   "OSPFv2 Grace LSA Top Level TLV" IANA sub-registry of the "Open
   Shortest Path First v2 (OSPFv2) Parameters" registry.

6. Acknowledgments

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

   This document was produced using Marshall Rose's xml2rfc tool.

7. References

7.1. Normative References

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

   [OSPF-TE]   Katz, D., Yeung, D., and K. Kompella, "Traffic
                    Engineering Extensions to OSPF", RFC 3630,
                    September 2003.

    [OSPFv2]   Moy, J., "OSPF Version 2", STD 54, RFC 2328,
                        April 1998.

    [OSPFv3]   Moy, J., Ferguson, D., and R. Coltun, "OSPF for IPv6",
                        RFC 2740, March 1997.

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

7.2. Informative References

    [MIB-INTF]   McCloghrie, K. and M. Rose, "Management Information
                        Base for network management of TCP/IP-based internets:
                        MIB-II", STD 17, RFC 1213, March 1991.

    [MIB-PERS]   McCloghrie, K. and F. Kastenholz, "The Interfaces
                        Group MIB", RFC 2863, June 2000.

    [OSPFv3-AUTH] Gupta, M. and N. Melam, "Authentication/
                        Confidentiality for OSPFv3", RFC 4552, June 2006.

Authors' Addresses

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

    EMail: ppe@cisco.com

    Acee Lindem
    Redback Networks
    102 Carric Bend Court
    Cary, NC 27519
    USA

    EMail: acee@redback.com

Full Copyright Statement

Intellectual Property