Network Working Group
Internet Draft

Anup Kumar T, Cisco System.
Sujay Gupta, Huawei Technologies.

Update to OSPF Graceful Restart procedure
draft-holla-ospf-update-graceful-restart-02.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.

Copyright Notice

Copyright (C) The Internet Society (2006).

Abstract

This document suggests improvements to OSPF v2 Graceful Restart. The basic protocol working remains as specified in [OSPFGR]. The draft describes a two pronged approach for improving the current graceful restart mechanism. It improves the restarting Router’s ability to
detect and react to topology changes. It also reduces the conservative behavior of the helper Router in reacting to topology changes. In addition, it adds certain clarifications for interpretation of the RFC 3623.

Conventions used in this document

In the text "GR" indicates Graceful Restart. The meaning of Â“Restarting RouterÂ” and Â“Helper RouterÂ” are as per [OSPFGR].

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.

Table of Contents

1. Introduction...................................................2
2. Terminology....................................................3
   2.1 Topology-Change LSAs....................................3
   2.2 Scope of the Topology-Change LSA.........................4
3. Enhancements to Graceful Restart mechanism.....................5
   3.1 Operations of the Helper Router..........................5
   3.2 Operations of the Restarting Router.......................7
   3.3 Explicit feedback mechanism...............................8
   3.4 Mechanism 1: Grace LSA TLV Signaling.....................8
   3.5 Mechanism 2: Link Local Signaling........................9
   3.6 Mechanism 3: One way hello Signaling....................9
   3.7 Recommendation........................................10
4. Security Considerations......................................10
5. IANA Considerations..........................................10
6. Backward Compatibility.......................................10
7. Special cases and considerations.............................10
8. Appendix................................................................11
Normative References............................................13
Informative References..........................................13
Acknowledgments................................................13
AuthorsÂ Addresses..............................................13
Intellectual Property Statement................................14

1. Introduction

This document suggests improvements to OSPF v2 Graceful Restart. The basic protocol working remains as specified in [OSPFGR].
Existing specification for Graceful restart in OSPFv2 has scope for optimization in terms of overall convergence time and avoidance of routing inconsistencies. This draft suggests methods to achieve the same. It also clarifies some helper behavior on reception of multiple grace LSA's.

This draft provides mechanisms for the restarting Router to detect and react to topology changes. It reduces the conservative behavior of the helper Router in reacting to topology changes. It also discusses introduction of an explicit mechanism through which a Router can communicate its non participation as a Helper Router. Further, certain clarification on interpreting multiple instances Grace LSA's from the restarting router have been added.

This document is organized as follows:

Section 2 introduces the concept of ÂToplogy-Change LSAÂ.

Section 3.1 describes a less conservative approach to be taken by the helper Router in reacting to topology changes.

Section 3.2 describes the additional checks a Router under restart should perform to ensure faster detection of a topology change to facilitate an early exit of Graceful Restart.

Section 3.3-7 introduces some explicit mechanisms that can be used by the helper, to communicate with the restarting router. These mechanisms signify the helper exiting the helper mode.

The proposed suggestions are backward compatible with [OSPFGR].

2. Terminology
2.1 Topology-Change LSAs

In [OSPFGR], helper routers detect topology changes by examining Âchanged LSAsÂ as prescribed in Section 3.3 [OSPFDC]. However, this notion of Âchanged LSAÂ is very broad and can be reduced in scope without affecting routing consistency. In this regard, the scope of a ÂTopology-Change LSAÂ is defined.

ÂTopology-Change LSAsÂ are Âbad newsÂ versions of their existing counterparts in the database. A detailed definition follows in section 2.2.

These LSAs definitely imply Âbad newsÂ it should cause helper routers to exit helper mode immediately. Other LSAs may as well trigger a router to exit helper mode. This is discussed in Section 3.1.
2.2 Scope of the Topology-Change LSA

For an LSA to qualify as a Topology-Change LSA, it must satisfy one of the following two requirements:

(a) A Max-Aged LSA received or originated for which there exists a previous database copy MUST be considered as a Topology-Change LSA

(b) A modified LSA received or originated for which there exists a previous database copy MUST be considered as a Topology-Change LSA

All other LSAs received or originated MUST NOT be considered as a Topology-Change LSA. Thus, this class excludes periodic refresh LSAs and also new LSAs without a previous database copy.

In a further attempt to reduce the scope of Topology-Change LSAs, an implementation MAY choose to interpret non max-aged modified LSAs (case (b) above) as follows:

If any existing information described in the previous database copy is missing or modified in the new LSA, then the new LSA MUST be considered as a Topology-Change LSA.

In this regard, the criteria for an LSA to be considered as a Topology-Change LSA are:

1. For Router LSA:
   a. Change in the options field or any of the V, E or B bits in the LSA.
   b. Absence of any link piece (link ID, link Data tuple) previously present in the database copy.
   c. Change of the contents of an existing link piece (Ex: metric)

2. For Network LSA:
   a. Change in the options field of the LSA.
   b. Change in Network mask.
   c. Absence of any previously attached router.

3. Any modification in the summary LSAs should result in considering it as a Topology-Change LSA.

4. For External and NSSA LSAs:
   a. Change in the options field in the LSA.
   b. Change of the network mask.
   c. Change in the E bit setting (indicating the type of External metric) or the metric.
A non max-aged changed LSA which does not exhibit any of the
modifications listed above, MAY NOT be considered as a Topology-
Change LSA. This implies that if the change in the LSA is the
presence of any new information not found in the previous copy (Ex: a
new link piece in a Router LSA, a new attached router in a network
LSA etc), the LSA MAY NOT be considered as a Topology-Change LSA.

3. Enhancements to Graceful Restart mechanism.

3.1 Operations of the Helper Router

This section describes some modifications to the helper router
behavior. The basic protocol behavior remains as specified in Section
3 of [OSPFGR].

Whenever a helper router encounters a non-refresh LSA (section 3.2
[OSPFGR]) which has to be flooded to a restarting router, it exits
helper mode. This results in the restarting router quitting GR. This
constraint is overkill, and can be optimized.

3.1.1 Entering Helper mode

One of the criteria for a router to enter helper mode (section 3.1
[OSPFGR]) requires that all LSAs on the retransmit list on the helper
router for the restarting router, should be periodic refresh LSAs.
This criterion is conservative and can be relaxed.
In this draft, this has been achieved through an explicit definition
of a Topology-Change LSA (Section 2.1). This definition reduces the
scope of topology changes that necessitate a router to terminate
helper mode.

In light of this, the checks that a Router (say Router Y) should
perform (Section 3.1 [OSPFGR]) while entering helper mode for a
restarting router (say Router X), are modified as follows:
On receiving Grace LSA from X, Y should examine the retransmit list
for X. If Y finds any Topology-Change LSA on the retransmit list
and, Y MUST NOT enter helper mode.

Rest of the behavior should be as per section 3.1 [OSPFGR].

3.1.2 Refusal to enter Helper mode

In certain scenarios, a router may refuse to enter helper mode after
receiving Grace LSAs from a restarting router. This could be due to a
local policy configured on the router which prevents it from entering
helper mode for the particular restarting router, or the type of
graceful restart (planned/unplanned) or the duration of graceful
restart. A router could refuse being a helper also due to the presence of a Topology-Change LSA on the retransmit list for the restarting router.

In such cases, the router SHOULD immediately explicitly signal to the restarting router indicating its reluctance to enter helper mode. The possible methods of achieving explicit signaling are explained in Section 3.3.

This helps the restarting router in making an informed decision on whether to enter or abandon a Graceful Restart. In the absence of this mechanism, the restarting router will enter a Graceful Restart and later during adjacency formation, on examining the neighbor’s router/network LSA, it would exit graceful restart. With this mechanism, the restarting router can immediately abandon graceful restart. Therefore minimizing any possible routing inconsistencies and improving the overall convergence time.

3.1.3 Exiting Helper mode

The existing criteria for exiting helper mode are very conservative. They can be relaxed to ensure that the restarting router is made to exit GR only when it is necessary to prevent routing inconsistencies. This has been achieved through an explicit definition of a Topology-Change LSA in Section 2.1.

1. Action on receiving or originating a Topology-Change LSA:

   If a helper router receives or originates a Topology-Change LSA and the LSA has to be flooded to a restarting router, the helper router MUST immediately terminate helper mode for the restarting router.

   This ensures that the restarting router’s routing table does not have any invalid routes.

   Other LSAs are to be flooded without the helper router exiting helper mode. However, sometimes it may lead to routing holes, to avoid it the helper router must additionally perform the below mentioned check when performing route computation.

2. Action on Route Calculation:

   When a helper runs SPF [OSPF], if a route is added or modified making a restarting router as the next hop, then the helper MUST immediately exit helper mode for that restarting router.
It may be argued that the action on route calculation is sufficient and there is no need for a explicit definition and handling of topology-change LSAs. However, such an approach is suboptimal since routing holes are introduced for short periods. Also, routing holes will be created on the GR router since the SPF is not run for inter-area and external destinations. As such inconsistencies could last until the completion of GR.

Using both the mechanisms together is provably complete and optimal.

3.1.4
Operation of Router when acting as Helper

The Helper router (say Y) when supporting a graceful restart for a specified router (say X) should additionally maintain the following checks

1. If the Helper router receives a new instance of a grace LSA from the restarting router. It should serve as helper for the new grace period as specified in the new grace LSA, provided the cumulative sum of the grace periods since running as helper for the restarting router(X) is less than LSRefreshTime (1800 second) [OSPF]. This is to prevent possibility DOS attack.

2. If the Helper router receives a new grace LSA from the restarting router with changed graceful restart reason TLV or IP interface address [OSPFGR]. The helper should exit helping mode and follow operations as specified in Section 3.1.4.

3.1.5
Operation of the Router on Exiting Helper Mode

The helper router on exiting helper mode before the grace LSAs are flushed by the restarting router SHOULD, explicitly signal this event to the restarting router (See Section 3.3), in addition to the operations specified in Section 3.1[OSPFGR].

3.2
Operations of the Restarting Router

This section describes some modifications to the restarting Router behavior. Some responsibility of detecting topology changes has now been placed on the restarting router. The basic protocol behavior remains as specified in Section 3 [OSPFGR].

3.2.1
Exit Graceful Restart
A restarting router SHOULD perform the following actions along with those mentioned in the “Exit Graceful Restart” in Section 2.2 [OSPFGR].

1. If on the restarting router, the inactivity timer fires for a neighbor, or there is a downward transition from 2-way to 1-way with a neighbor, the router SHOULD exit graceful restart.

2. The restarting router, on receiving the explicit ‘exit-helper’ signal (See description in Section 3.3) from the helper must immediately quit GR.

3.3
Explicit feedback mechanism.

The restarting router learns about a router not supporting helper mode by examining router and network LSAs (section 2.2 [OSPFGR]). This mechanism is sub-optimal, and in some scenarios, ineffectual (Section 8: Case A).

This draft proposes some methods for the helper routers to signal events to the restarting router, which will improve the convergence time as well as reduce the possibility of routing holes. The next section discusses three mechanisms which may be used to this effect.

3.4
Mechanism 1: Grace LSA TLV Signaling

It introduces an additional TLV for the Grace LSA, using which the helper can explicitly signal its non-participation as a helper router.

3.4.1
Sending Grace LSA with the Exit-Grace TLV.

The Grace LSA with the Exit-Grace TLV SHOULD be sent as unicast packets to the restarting router on Broadcast, NBMA and P2MP networks. However, on P2P interfaces, the packet should be sent to AllSPFRouters. [OSPF]
When this TLV is included, the mandatory Grace TLV (described in [OSPFGR]) MUST NOT be included in the LSA.
The sending of this Exit-Grace TLV MAY be made reliable.

3.4.2
Receiving Grace LSA with Exit-Grace TLV.

A router, on receiving grace LSA containing the Exit-grace TLV from a neighbor, MUST acknowledge the LSA. The LSA MUST NOT be stored in the database or flooded. If the receiving router is under graceful-restart, it must immediately exit graceful restart. If the receiving router is not under Graceful Restart, the TLV will have no impact on the operations of the router.
3.4.3 Format of Exit-Grace LSA TLV

(See Appendix for Format)

3.5 Mechanism 2: Link Local Signaling
This mechanism uses the ideas presented in [OSPFLLS]. It introduces an additional LLS Exit-Grace TLV to indicate the helper’s non-participation (See Appendix for Format). This LLS TLV should be sent in Hello packets [OSPFLLS].

3.5.1 Sending LLS Exit-Grace TLV in the Hello packet.

The helper router if LLS capable should append the Exit-Grace TLV to all Hello packets sent out of its interface to the restarting router after exiting helper mode. As signaling using Hello packets is unreliable, an implementation MAY continue signaling until the grace-period expires or the grace-LSA is flushed.

3.5.2 Receiving LLS Exit-Grace TLV

A router if LLS capable and undergoing graceful restart, on receiving an LLS Exit-Grace TLV should examine if the Router ID in the TLV matches with its own. If it matches it should immediately exit graceful restart.

3.5.3 Format of Exit-Grace LLS TLV

(See Appendix for Format)

3.6 Mechanism 3: One way hello Signaling
The helper router sends a one-way hello to the restarting router indicating its non-participation as a helper.

3.6.1 Operation of the Helper on sending one-way hello
The helper router should send a one-way hello to the restarting router but in doing so should not change its adjacency relationship with the restarting router. In particular the one way hello should not trigger a one way event. Other operations remain as specified in [OSPF] and [OSPFGR]. The restarting router on receiving a one-way hello should exit GR. This is covered is Section 3.2.1 [1].
3.7 Recommendation

Among the above specified mechanisms of explicit signaling it is recommended that the Helper use Mechanism 2 as specified above in Section 3.5.

4. Security Considerations

This draft raises no new security considerations other than that covered in [OSPFGR] and [OSPF].

5. IANA Considerations

This document describes a new TLV for Grace LSAs. The TLV type is to be assigned by IANA. The suggested value is 4. (Refer IANA). This document suggests using an Exit Grace LLS TLV. The TLV type for which is to be assigned by IANA.

6. Backward Compatibility

The compliance of the feature mentioned here, allows for faster and optimal detection of topology changes and thus faster convergence. If one or more neighbors do not support Exit-Grace LSA TLV/Exit Grace LLS TLV and/or other features mentioned above and support only RFC 3623, the protocol behavior specified in [OSPFGR] remains unimpaired. However, one anomaly for backward compatibility introduced by this draft is documented in Section 7.

7. Special cases and considerations

Area 0
RT1-----------------X-----------------RT2---------RT3

In the above topology RT1, X, RT2 and RT3 are all in area 0 X and RT2 are implementations which are based on this draft. However, RT1 is based on [OSPFGR] only.

Scenario:
X having initiated GR has completed adjacency formation with helper RT1, but, is still in the process of forming adjacency with RT2. At this point, RT3 originates a NEW type-5 LSA. Due to the nature of the modifications proposed in this draft, RT2 will not exit helper mode on receipt of this LSA from RT3 (refer sections 2.1 and 3.1.3). However, since X will flood this LSA to RT1, RT1 will exit helper mode.
mode. Since, RT1 and X are full with each other, X cannot detect RT1 quitting helper mode. Therefore, X continues with GR WITHOUT adding a route to the destination described by the new LSA, as X does not update its forwarding table during GR. This results in a routing hole. This condition will persist until X exits graceful restart.

8. Appendix

Format of Exit-Grace LSA TLV:

The Exit-Grace LSA TLV is enclosed in the Opaque link-local scoped Grace LSA. The format of the Grace LSA is specified in [OSPFGR].

The format of the Exit-Grace TLV is described below.

```
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 type field of the Exit-Grace TLV is assigned an experimental value of 4. The length field, defining the length of the value field in octets, MUST be 4. For non P2P networks, the value MUST be the IP address of the interface through which the helper sends the Grace LSA. For P2P networks, the value MUST be the Router-ID of the helper router.

If the above TLV is present in a Grace LSA, it is mandatory for the Grace-Period TLV to be absent. The absence of the Grace-Period TLV ensures the Receiving router dropping the LSA, as the Grace-Period TLV is specified as mandatory in [OSPFGR]. In particular this ensures that implementations not supporting the Exit-Grace TLV will not process the LSA.

The format of the Grace LSA other than the addition of the above TLV remains same as that mentioned in [OSPFGR].

Format of the Exit-Grace LLS TLV:

The Exit-Grace LSA TLV is sent in the Hello packet by the helper router.
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 type field of the Exit-LLS TLV is to be assigned a value. The length field, defining the length of the value field in octets, MUST be 4. The value MUST contain Router-ID of the restarting router and not that of the Helper router.

Case A

Since the protocol [OSPFGR] does not provide a method for a helper router to notify the restarting router, its exit of helper mode after the adjacency formation is complete, there are possibilities of routing holes existing for a bounded period of time, grace LSA being flushed or grace period expiry.

Ex:

Area 0
RT1-----------------X-----------------RT2---------RT3

Area 1
| RT4

In the above topology RT1, X, RT2 and RT3 are all in area 0; X and RT4 are in Area 1 which is a stub area.

Scenario:
X having initiated GR has completed adjacency formation with helpers RT1 and RT2 and still in the process of forming adjacency with RT4. At this point, RT3 originates a new type-5 LSA. Due to the nature of the protocol, RT2 and RT1 will exit helper mode on receipt of this LSA from RT3 and X respectively. Since, X will not send this LSA to RT4, RT4 will not exit helper mode. RT1 and RT2 have no mechanism to inform X about their quitting of helper mode (since they are already full with X). Therefore, X continues with GR WITHOUT adding a route to the destination described by the new LSA, as X will not update the forwarding table during GR. This results in a routing hole being created at X. This condition will persist until X exits GR.
Normative References


Informative References


Acknowledgments

The authors wish to thank Acce Lindem, Vishwas Manral, Don Goodspeed, Pradeep Shastry, K.L Srini, Saravanak, Abhay, Abhishek and Sandeep for their insightful comments and suggestions.

Authors’ Addresses

Ashok Chandrashekhar Holla
Dartmouth College
NH, U.S.A
ashok.chandrashekar@dartmouth.edu

Anup Kumar T
Cisco Systems
Bangalore, India.
anut@cisco.com

Sujay Gupta
Huawei Technologies
Bangalore, India
sujayg@huawei.com

Holla, Kumar & Gupta
Expires - March 2007
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.