

Networking Working Group

Zafar Ali
Jean-Philippe Vasseur
Anca Zamfir
Cisco Systems, Inc.

IETF Internet Draft
Category: Standard track
Expires: December 2006

June 2006

draft-ali-ccamp-mpis-graceful-shutdown-04.txt

Graceful Shutdown in GMPLS Traffic Engineering Networks

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>.

Abstract

GMPLS-TE Graceful shutdown is a method for explicitly notifying the nodes in a Traffic Engineering (TE) enabled network that the TE capability on a link or on an entire Label Switching Router (LSR) is going to be disabled. GMPLS-TE graceful shutdown mechanisms are tailored towards addressing the planned outage in the network.

This document provides requirements and protocol mechanisms so as to reduce/eliminate traffic disruption in the event of a planned shutdown of a network resource. These operations are equally applicable for both MPLS and its GMPLS extensions.

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](#) [i].

Table of Contents

1. Terminology.....	2
2. Introduction.....	3
3. Requirements for Graceful Shutdown.....	3
4. Mechanisms for Graceful Shutdown.....	4
4.1 RSVP-TE Signaling Mechanism for graceful shutdown.....	4
4.1.1 Graceful Shutdown of TE link(s).....	5
4.1.2 Graceful Shutdown of Component Link(s) in a Bundled TE Link...5	5
4.1.3 Graceful Shutdown of TE Node.....	5
4.2 OSPF/ ISIS Mechanisms for graceful shutdown.....	6
4.2.1 Graceful Shutdown of TE link(s).....	6
4.2.2 Graceful Shutdown of Component Link(s) in a Bundled TE Link...6	6
4.2.3 Graceful Shutdown of TE Node.....	6
5. Security Considerations.....	6
6. IANA Considerations.....	6
7. Intellectual Property Considerations.....	7
8. Full Copyright Statement.....	7
9. Acknowledgments.....	7
10. Reference.....	7
10.1 Normative Reference.....	7
10.2 Informative Reference.....	8

[1. Terminology](#)

LSR - Label Switching Device.

LSP - An MPLS Label Switched Path

Head-end or Ingress node: In this document the terms head-end node equally applies to the Ingress node that initiated signaling for the

Path, or an intermediate node (in the case of loose hops path computation) or a Path Computation Element (PCE) that computes the routes on behalf of its clients (PCC).

GMPLS - The term GMPLS is used in this document to refer to both classic MPLS, as well as the GMPLS extensions to MPLS.

TE Link - The term TE link refers to a physical link or an FA-LSP, on which traffic engineering is enabled. A TE link can be bundled or unbundled.

The terms node and LSR will be used interchangeably in this document.

2. Introduction

When outages in a network are planned (e.g. for maintenance purpose), some mechanisms can be used to avoid traffic disruption. This is in contrast with unplanned network element failure, where traffic disruption can be minimized thanks to recovery mechanisms but may not be avoided. Hence, a Service Provider may desire to gracefully (temporarily or definitely) disable Traffic Engineering on a TE Link, a group of TE Links or an entire node for administrative reasons such as link maintenance, software/hardware upgrade at a node or significant TE configuration changes. In all these cases, the goal is to minimize the impact on the GMPLS traffic engineered flows carried over TE LSPs in the network by triggering notifications so as to graceful reroute such flows before the administrative procedures are started.

Graceful shutdown of a resource may require several steps. These steps can be broadly divided into two sets: disabling the resource in the control plane and removing the resource for forwarding. The node initiating the graceful shutdown condition SHOULD delay the removal of the resources for forwarding, for some period determined by local policy. This is to allow control plane to gracefully divert the traffic away from the resource being gracefully shutdown. Similarly, trigger for the graceful shutdown event is a local matter at the node initiating the graceful shutdown. Typically, graceful shutdown is triggered for administrative reasons, such as link maintenance or software/hardware upgrade at a node.

This document describes the mechanisms that can be used to gracefully shutdown GMPLS Traffic Engineering on a resource. As mentioned earlier, the graceful shutdown of the Traffic Engineering capability on a resource could be incorporated in the traditional shutdown operation of an interface, but it is a separate step that is taken before the IGP on the link is brought down and before the interface is brought down at different layers. This document only addresses TE node and TE resources.

3. Requirements for Graceful Shutdown

This section lists the requirements for graceful shutdown in the context of GMPLS Traffic Engineering.

- Graceful shutdown must address graceful removal of one TE link, one component link within a bundled TE link, a set of TE links, a set of component links or an entire node.
- It is required to prevent other network nodes to use the network resources that are about to be shutdown, should new TE LSP be set up. Similarly it is required to reduce/eliminate traffic disruption on the LSP(s) using the network resources which are about to be shutdown.
- Graceful shutdown mechanisms are required to address TE LSPs spanning multiple domains, as well as intra domain TE LSPs. Here, a domain is defined as either an IGP area or an Autonomous System [[INTER-AREA-AS](#)].
- Graceful shutdown is equally applicable to GMPLS-TE, as well as packet-based (MPLS) TE LSPs.
- In order to make rerouting effective, it is required to communicate information about the TE resource under graceful shutdown.

4. Mechanisms for Graceful Shutdown

An IGP only based solution is not applicable when dealing with Inter-area and Inter-AS traffic engineering, as IGP LSA/LSP flooding is restricted to IGP areas/levels. Consequently, RSVP based mechanisms are required to cope with TE LSPs spanning multiple domains. At the same time, RSVP mechanisms only convey the information for the transiting LSPs to the router along the upstream Path and not to all nodes in the network. Furthermore, it must be noted that graceful shutdown notification via IGP flooding is required to discourage a node from establishing new LSPs through the resources being shutdown. In the following sections the complementary mechanisms for RSVP-TE and IGP for Graceful Shutdown are described.

4.1 RSVP-TE Signaling Mechanism for graceful shutdown

As discussed in [Section 3](#), one of the requirements for the signaling mechanism for graceful shutdown is to carry information about the resource under graceful shutdown. The Graceful Shutdown mechanism outlined in the following section, uses Path Error and where available, Notify message, in order to achieve this requirement. Such mechanisms relying on signaling are only applicable to the existing LSPs.

Setup request for new LSPs over the TE resource being gracefully shutdown SHOULD be rejected using the existing mechanisms that are applied when the TE resource is not available.

4.1.1 Graceful Shutdown of TE link(s)

The node where graceful shutdown of a link or a set of links is desired MUST trigger a Path Error message with local link maintenance required sub-code for all affected LSPs. The local TE link maintenance required error code is defined in [\[PATH-REOPT\]](#). If available, and where notify requests were included when the LSPs were initially setup, Notify message (as defined in [\[\]](#)) MAY also be used for delivery of this information to the head-end nodes.

When a head-end LSR receives a Path Error (or Notify) message with sub-code "Local Maintenance on TE Link required Flag", it SHOULD immediately trigger a make-before-break procedure. A head-end node SHOULD avoid the IP address contained in the PathErr (or Notify message) when performing path computation for the new LSP.

[4.1.2](#) Graceful Shutdown of Component Link(s) in a Bundled TE Link

MPLS TE Link Bundling [\[BUNDLE\]](#) requires that an LSP is pinned down to component link(s). Hence, when a component link is shutdown, the TE LSPs affected by such maintenance action needs to be resingaled.

Graceful shutdown of a component link in a bundled TE link differs from graceful shutdown of unbundled TE link or entire bundled TE link. Specifically, in the former case, when only a subset of component links and not the entire TE bundled link is being shutdown, the remaining component links of the TE links may still be able to admit new LSPs. Consequently a new error sub-code for PathError and Notify message is needed:

9 (TBA)	Local component link maintenance required
---------	---

Error Sub-code for Local component link maintenance required is to be assigned by IANA.

If the last component link is being shutdown, the procedure outlined in [Section 5.1](#) is used.

When a head-end LSR receives an RSVP Path Error or Notify message with sub-code "local component link maintenance required Flag set, it SHOULD immediately perform a make-before-break to avoid traffic loss. The head-end LSR MAY still use the IP address contained in the Path Error or Notify message in performing path computation for rerouting the LSP. This is because, this address is an IP address of the component link and the flag is an implicit indication that the TE link may still have capacity to admit new LSPs. However, if the ERO is computed such that it also provides details of the component link selection(s) along the Path, the component link selection with IP address contained in the Path Error or Notify message SHOULD be avoided.

[4.1.3](#) Graceful Shutdown of TE Node

When graceful shutdown at node level is desired, the node in question follows the procedure specified in the previous section for all TE Links.

[4.2](#) OSPF/ ISIS Mechanisms for graceful shutdown

The procedures provided in this section are equally applicable to OSPF and ISIS.

[4.2.1](#) Graceful Shutdown of TE link(s)

The node where graceful-shutdown of a link is desired MUST originate the TE LSA/LSP containing Link TLV for the link under graceful shutdown with Traffic Engineering metric set to 0xffffffff, 0 as unreserved bandwidth, and if the link has LSC or FSC as its Switching Capability then also with 0 as Max LSP Bandwidth. This would discourage new LSP establishment through the link under graceful shutdown.

Neighbors of the node where graceful shutdown procedure is in progress SHOULD continue to advertise the actual unreserved bandwidth of the TE links from the neighbors to that node, without any routing adjacency change.

[4.2.2](#) Graceful Shutdown of Component Link(s) in a Bundled TE Link

If graceful shutdown procedure is performed for a component link within a TE Link bundle and it is not the last component link available within the TE link, the link attributes associated with the TE link are recomputed. If the removal of the component link results in a significant bandwidth change event, a new LSA is originated with the new traffic parameters. If the last component link is being shutdown, the routing procedure outlined in [Section 4.2.1](#) is used.

[4.2.3](#) Graceful Shutdown of TE Node

When graceful shutdown at node level is desired, the node in question follows the procedure specified in the previous section for all TE Links.

[5](#). Security Considerations

This document does not introduce new security issues. The security considerations pertaining to the original RSVP protocol [[RSVP](#)] remain relevant.

[6](#). IANA Considerations

A new error sub-code for Path Error and Notify message is needed for Local component link maintenance required flag.

7. Intellectual Property Considerations

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.

8. Full 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.

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.

9. Acknowledgments

The authors would like to acknowledge useful comments from David Ward, Sami Boutros, Adrian Farrel and Dimitri Papadimitriou.

10. Reference

10.1 Normative Reference

[RSVP] Braden, et al, "Resource ReSerVation Protocol (RSVP) - Version 1, Functional Specification", [RFC 2205](#), September 1997.

[RSVP-TE] Awduche, et al, "Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.

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

[RFC3473] L. Berger, et al, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#).

[RFC4203] K. Kompella, Y. Rekhter, et al, OSPF Extensions in Support of Generalized MPLS , [draft-ietf-ccamp-ospf-gmpls-extensions-12.txt](#).

[RFC4205] K. Kompella, Y. Rekhter, et al, IS-IS Extensions in Support of Generalized MPLS , [draft-ietf-isis-gmpls-extensions-19.txt](#).

[PATH-REOPT] Jean-Philippe Vasseur, and Y. Ikejiri, Reoptimization of MPLS Traffic Engineering loosely routed LSP paths , [draft-ietf-ccamp-loose-path-reopt-02.txt](#).

10.2 Informative Reference

[INTER-AREA-AS] Adrian Farrel, Jean-Philippe Vasseur, Arthi Ayyangar, A Framework for Inter-Domain MPLS Traffic Engineering , [draft-ietf-ccamp-inter-domain-framework-04.txt](#).

[BUNDLE] Kompella, K., Rekhter, Y., Berger, L., "Link Bundling in MPLS Traffic Engineering", [draft-ietf-mpls-bundle-04.txt](#) (work in progress)

Authors' Address:

Zafar Ali
Cisco systems, Inc.,
2000 Innovation Drive
Kanata, Ontario, K2K 3E8
Canada.
Email: zali@cisco.com

Jean Philippe Vasseur
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough , MA - 01719
USA
Email: jpv@cisco.com

Anca Zamfir
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, Ontario, K2K 3E8
Canada
Email: ancaz@cisco.com

