

Networking Working Group
Internet Draft

Zafar Ali
Jean-Philippe Vasseur
Anca Zamfir
Cisco Systems, Inc.
Jonathan Newton
Cable and Wireless

Category: Informational
Expires: July 2008

January 2008

draft-ietf-ccamp-mpls-graceful-shutdown-05.txt

Graceful Shutdown in MPLS and Generalized MPLS
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>

This Internet-Draft will expire on July 2008.

Copyright Notice

Copyright (C) The IETF Trust (2008).

Expires July 2008

[Page 1]

Abstract

MPLS-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. MPLS-TE graceful shutdown mechanisms are tailored toward addressing planned outage in the network.

This document provides requirements and protocol mechanisms to reduce/eliminate traffic disruption in the event of a planned shutdown of a network resource. These operations are equally applicable to both MPLS and its Generalized MPLS (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](#) [[RFC2119](#)].

Table of Contents

1. Terminology.....	3
2. Introduction.....	3
3. Requirements for Graceful Shutdown.....	4
4. Mechanisms for Graceful Shutdown.....	5
4.1 OSPF/ ISIS Mechanisms for graceful shutdown.....	5
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
4.1.3 Graceful Shutdown of TE Node.....	6
4.1.4 Graceful Shutdown of Label Resource.....	6
4.2 RSVP-TE Signaling Mechanism 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 .	7
4.2.3 Graceful Shutdown of TE Node.....	8
4.2.2 Graceful Shutdown of a Label Resource.....	8
5. Security Considerations.....	8
6. IANA Considerations.....	9
7. Acknowledgments.....	9
8. Reference.....	9
8.1 Normative Reference.....	9
8.2 Informative Reference.....	9
9. Authors' Address:.....	10
10. Intellectual Property Considerations.....	10
11. Disclaimer of Validity.....	11
12. Copyright Statement.....	11

Expires July 2008

[Page 2]

1. Terminology

LSR (Label Switching Router): The terms node and LSR are used interchangeably in this document.

GMPLS: The term GMPLS is used in this document to refer to packet MPLS-TE, as well as GMPLS extensions to MPLS-TE.

LSP: An MPLS-TE/ GMPLS-TE Label Switched Path.

Head-end node: Ingress LSR that initiated signaling for the Path.

Border node: Ingress LSR of an LSP segment (S-LSP).

Path Computation Element (PCE): An entity that computes the routes on behalf of its clients (PCC).

TE Link: The term TE link refers to single or a bundle of physical link(s) or FA-LSP(s) on which traffic engineering is enabled [[RFC4206](#)], [[RFC4201](#)].

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) remove 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 gracefully 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 introduce a delay between disabling the resource in the control plane and removing the resource for forwarding. This is to allow the control plane to gracefully divert the traffic away from the resource being gracefully shutdown. The 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.

Expires July 2008

[Page 3]

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 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 nodes 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.
- Once an operator has initiated graceful shutdown of a network resource, no new TE LSPs may be set up that use the resource. Any signaling message for a new LSP that explicitly specifies the resource, or that would require the use of the resource due to local constraints, must be rejected as if the resource were unavailable.
- It is desirable for new LSP setup attempts that would be rejected because of graceful shutdown of a resource (as described in the previous requirement) to avoid any attempt to use the resource by selecting an alternate route or other resources.
- If the resource being shutdown is a last resort, it can be used. Time or decision for removal of the resource being shutdown is based on a local decision at the node initiating the graceful shutdown procedure.
- It is required to give the ingress node the opportunity to take actions in order to reduce/eliminate traffic disruption on the LSP(s) that are using the network resources which are about to be shutdown.
- Graceful shutdown mechanisms are equally applicable to intra-domain and TE LSPs spanning multiple domains. Here, a domain is defined as either an IGP area or an Autonomous System [[RFC4726](#)].
- 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 that when a node initiates the graceful shutdown of a resource, it

Expires July 2008

[Page 4]

identifies to all other network nodes the TE resource under graceful shutdown.

- Depending on switching technology, it may be possible to shutdown a label resource, e.g., shutting down a lambda in a Lambda Switch Capable (LSC) node.

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.

A node where a link or the whole node is being shutdown SHOULD first trigger the IGP updates as described in [Section 4.1](#), introduce a delay to allow network convergence and only then use the signaling mechanism described in [Section 4.2](#).

4.1 OSPF/ ISIS Mechanisms for graceful shutdown

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

4.1.1 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. A node MAY also specify a value for Minimum LSP bandwidth which is greater than the available 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.

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

Expires July 2008

[Page 5]

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.

[4.1.3](#) 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.

[4.1.4](#) Graceful Shutdown of Label Resource

If graceful shutdown procedure is performed on a label resource within a TE Link, the link attributes associated with the TE link are recomputed. If the removal of the label resource results in a significant change event, a new LSA is originated with the new traffic parameters.

[4.2](#) 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 PathErr and where available, Notify message, in order to achieve this requirement. These mechanisms apply to both existing and new LSPs.

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

The node where graceful shutdown of a link or a set of links is desired MUST trigger a PathErr message with the error code "Notify" and an error value of "Local link maintenance required" for all affected LSPs. The "Notify" error code is defined in [\[RFC3209\]](#) while the "local link maintenance required" error value is defined in [\[RFC4736\]](#). The PathErr message SHOULD include the ERROR_SPEC object containing IP address of the TE Link being gracefully shutdown. If TE link is unnumbered, the PathErr message SHOULD include the ERROR_SPEC object containing unnumbered ID and TE router ID for the TE Link being gracefully shutdown. If available, and where notify requests were included when the LSPs were initially setup, Notify message (as defined in [RFC 3471](#), [RFC 3473](#)) MAY also be used for delivery of this

information to the head-end node, border node or PCE.

Expires July 2008

[Page 6]

When a graceful shutdown operation is performed along the path of a protected LSP, based on a local decision, the PLR or branch node MAY redirect the traffic onto the local detour or protecting segment. In all cases, the PLR or branch node MUST forward the PathErr to the head-end node, border node, or PCE.

When a head-end node, border node, or PCE receives a PathErr (or Notify) message with error value of "Local link maintenance required", it MAY trigger a make-before-break procedure. When performing path computation for the new LSP, the head-end node, border node, or PCE SHOULD avoid using the TE resources identified by the IP address contained in the PathErr (or Notify message)

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

MPLS TE Link Bundling [[RFC4201](#)] requires that an LSP is pinned down to component link(s). Hence, when a component link is shutdown, the TE LSPs affected by this action need to be resigned.

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 bundled TE link may still be able to admit new LSPs.

The node where graceful shutdown of a component link is desired MUST trigger a PathErr message with the error code "Notify" and the new error value of "Local component link maintenance required" for all affected LSPs. The "Notify" error code is defined in [[RFC3209](#)] while the "local component link maintenance required" error value is introduced by this proposal:

12 (TBA) Local component link maintenance required

Error value for "Local component link maintenance required" is to be assigned by IANA.

The PathErr message should include in the ERROR_SPEC the TE Link ID address.

If the last component link is being shutdown, the procedure outlined in [Section 4.2.1](#) is used.

When a head-end node, border node, or PCE receives an RSVP PathErr or Notify message with error value "local component link maintenance required" Flag set, it MAY immediately perform a make-before-break to avoid traffic loss. The head-end node, border node, or PCE MAY still use the IP address contained in the

PathErr or Notify message in performing path computation for rerouting the LSP. This is because, this address is an IP address of the TE link and the flag is an implicit indication that the TE

Expires July 2008

[Page 7]

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 previously selected MAY be avoided.

4.2.3 Graceful Shutdown of TE Node

The node that is being gracefully shutdown MUST trigger a PathErr message with the error code "Notify" and an error value of "Local node maintenance required" for all LSPs. The "Notify" error code is defined in [[RFC3209](#)] while the "local node maintenance required" error value is defined in [[RFC4736](#)].

The PathErr message should include in the ERROR_SPEC object the MPLS-TE Node ID address

4.2.2 Graceful Shutdown of a Label Resource

The node where graceful shutdown of a label resource is desired MUST trigger a PathErr message with the error code "Notify" and the new error value of "Local component link maintenance required" for the affected LSP. The "Notify" error code is defined in [[RFC3209](#)] while the "local component link maintenance required" error value is introduced by this proposal:

13 (TBA) Local label resource maintenance required

Error value for "Local label resource maintenance required" is to be assigned by IANA.

The PathErr message should include in the ERROR_SPEC the TE Link ID address.

If the last component link is being shutdown, the procedure outlined in [Section 4.2.1](#) is used.

When a head-end node, border node, or PCE receives an RSVP PathErr or Notify message with error value "local label resource maintenance required" Flag set, it MAY immediately perform a make-before-break to avoid traffic loss. The head-end node, border node, or PCE MAY still use the IP address contained in the PathErr or Notify message in performing path computation for rerouting the LSP. This is because, this address is an IP address of the TE link and the flag is an implicit indication that the TE link may still have capacity to admit new LSPs.

5. Security Considerations

This document introduces no new security considerations beyond those already addressed for existing RSVP PathErr or Notify messages, or advertisement of TE LSA/LSP containing Link TLV. In

this regard, the security considerations specified in [[RFC2205](#)], [[RFC3209](#)] and [[MPLS-GMPLS-SECURITY](#)] remain relevant.

Expires July 2008

[Page 8]

6. IANA Considerations

The following assignment is required in the "Notify" subsection of "Error Codes and Values" section of the "RSVP PARAMETERS" registry (located at <http://www.iana.org/assignments/rsvp-parameters>):

- 12 (TBA) - "Local component link maintenance required" flag.
- 13 (TBA) Local label resource maintenance required.

7. Acknowledgments

The authors would like to thank Adrian Farrel for his detailed comments and suggestions. The authors would also like to acknowledge useful comments from David Ward, Sami Boutros, and Dimitri Papadimitriou.

8. Reference

8.1 Normative Reference

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

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

[RFC4736] Jean-Philippe Vasseur, et al "Reoptimization of MPLS Traffic Engineering loosely routed LSP paths", [RFC 4736](#).

8.2 Informative Reference

[RFC2205] Braden, et al, "Resource ReSerVation Protocol (RSVP) Version 1, Functional Specification", [RFC 2205](#), December 1997.

[RFC4726] Adrian Farrel, Jean-Philippe Vasseur, Arthi Ayyangar, "A Framework for Inter-Domain MPLS Traffic Engineering", [RFC 4726](#).

[RFC4201] Kompella, K., Rekhter, Y., Berger, L., "Link Bundling in MPLS Traffic Engineering", [RFC 4201](#).

[RFC4206] Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE), [RFC 4206](#).

[MPLS-GMPLS-SECURITY] Fang, et al, "Security Framework for MPLS and GMPLS Networks", [draft-fang-mpls-gmpls-security-framework-](#)

[00.txt](#), work in progress.

Expires July 2008

[Page 9]

9. 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

Jonathan Newton
Cable and Wireless
jonathan.newton@cw.com

10. 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.

Expires July 2008

[Page 10]

11. 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, THE IETF TRUST 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.

12. Copyright Statement

Copyright (C) The IETF Trust (2007).

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.

Expires July 2008

[Page 11]