CCAMP Working Group
Internet Draft
Category: Informational   Zafar Ali
Expires: September 08, 2009   Jean-Philippe Vasseur
                                                        Anca Zamfir
                                                  Cisco Systems, Inc.
                                                    Jonathan Newton
                                                  Cable and Wireless
                                                      March 09, 2009

           Graceful Shutdown in MPLS and Generalized MPLS
                    Traffic Engineering Networks

           draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

Status of this Memo

   This Internet-Draft is submitted to IETF in full conformance with
   the provisions of BCP 78 and BCP 79. This document may contain
   material from IETF Documents or IETF Contributions published or
   made publicly available before November 10, 2008. The person(s)
   controlling the copyright in some of this material may not have
   granted the IETF Trust the right to allow modifications of such
   material outside the IETF Standards Process. Without obtaining
   an adequate license from the person(s) controlling the copyright
   in such materials, this document may not be modified outside the
   IETF Standards Process, and derivative works of it may not be
   created outside the IETF Standards Process, except to format it
   for publication as an RFC or to translate it into languages other
   than English.

   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.

        draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

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.

Table of Contents

1. 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 indefinitely)
   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 traffic carried over TE LSPs in the
network by triggering notifications so as to gracefully reroute
such flows before the administrative procedures are started.

draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

These operations are equally applicable to both MPLS [RFC3209]
and its Generalized MPLS (GMPLS) extensions [RFC3471], [RFC3473].

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
is expected to 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.

This document describes the mechanisms that can be used to
gracefully shutdown MPLS/ GMPLS Traffic Engineering on a resource
such as a TE link, a component link within a bundled TE link, a
label resource or an entire TE node.

2. 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].

Last resort resource: If a path to a destination from a given
head-end node cannot be found upon removal of a resource (e.g.,
TE link, TE node), the resource is called last resort to reach
that destination from the given head-end node.

3. Requirements for Graceful Shutdown

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

draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

- Graceful shutdown is required to 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, label resource(s) 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, is required to 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
  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 solution based on [RFC3630], [RFC5305], [RFC4203] and
[RFC5307] are 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,
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 may 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.

OSPF and ISIS procedure for graceful shutdown of TE link(s) is
similar to graceful restart of OSPF and ISIS as described in
[RFC4203] and [RFC5307], respectively. Specifically, the node
where graceful-shutdown of a link is desired originates 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.

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. Similarly, 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 component link or label resource 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 related to TE link removal is
used.

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

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

        draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

4.2 RSVP-TE Signaling Mechanisms 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. For this purpose the
    Graceful Shutdown uses LSP rerouting mechanism as defined in
    [LSP-REROUTE].

Specifically, the node where graceful shutdown of an unbundled TE
link or an entire bundled TE link is desired triggers a PathErr
message with the error codes and error values of "Notify/Local
link maintenance required", for all affected LSPs. Similarly, the
node that is being gracefully shutdown triggers a PathErr message
with the error codes and error values of "Notify/ Local node
maintenance required", for all LSPs. For graceful shutdown of a
node, an unbundled TE link or an entire bundled TE link, the
PathErr message may contain either an [RFC2205] format ERROR_SPEC
object, or an IF_ID [RFC3473] format ERROR_SPEC object. In either
case, it is the address and TLVs carried by the ERROR_SPEC object
and not the error value that indicates the resource that is to be
gracefully shutdown.

MPLS TE Link Bundling [RFC4201] requires that an LSP is pinned
down to a component link. Consequently, 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 triggers a PathErr message with the
error codes and error values of "Notify/Local link maintenance
required". The rest of the ERROR_SPEC object is constructed using
Component Reroute Request procedure defined in [LSP-REROUTE].

If graceful shutdown of a label resource is desired, the node
initiating this action triggers a PathErr message with the error
codes and error values of "Notify/Local link maintenance
required". The rest of the ERROR_SPEC object is constructed using
Label Reroute Request procedure defined in [LSP-REROUTE].

When a head-end node, a transit node or a border node receive a
PathErr message with error codes and error values of "Notify/Local
link maintenance required" or "Notify/ Local node maintenance
required", it follows the procedures defined in [LSP-REROUTE] to
reroute the traffic around the resource being gracefully shutdown.
When performing path computation for the new LSP, the head-end node,
or border node avoids using the TE resources identified by the

draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

ERROR_SPEC object. If PCE is used for path computation, head-end

node or border node acts as PCC to request the PCE via PCEP for path
computation avoiding resource being gracefully shutdown. The
amount of time the head-end node, or border node avoid using the
TE resources identified by the IP address contained in the
PathErr is based on a local decision at head-end node or border
node.

If node initiating the graceful shutdown procedure received path
setup request for a new tunnel using resource being gracefully
shutdown, it sends a Path Error message with "Notify" error code
in the ERROR SPEC object and an error value consistent with the
type of resource being gracefully shutdown. However, based on a
local decision, if an existing tunnel continues to use the
resource being gracefully shutdown, the node initiating the
graceful shutdown procedure may allow resource being gracefully
shutdown to be used as a "last resort". The node initiating the
graceful shutdown procedure can distinguish between new and
existing tunnels based on the tunnel ID in the SESSION object.

Time or decision for removal of the resource being shutdown from
forwarding is based on a local decision at the node initiating
the graceful shutdown procedure. For this purpose, the node
initiating graceful shutdown procedure follows the Reroute
Request Timeout procedure defined in [LSP-REROUTE].

5. Security Considerations

This document introduces no new security considerations as this
document describes usage of existing formats and mechanisms. This
document relies on existing procedures for advertisement of TE
LSA/LSP containing Link TLV. Tampering with TE LSAs may have an
effect on traffic engineering computations, and it is suggested
that any mechanisms used for securing the transmission of normal
OSPF LSAs/ ISIS LSPs be applied equally to all Opaque LSAs/ LSPs
this document uses. Existing security considerations specified
in [RFC3630], [RFC5305], [RFC4203], [RFC5307] and [MPLS-GMPLS-
SECURITY] remain relevant and suffice. Furthermore, security
considerations section in [LSP-REROUTE] and the Section 9 of
 [RFC4736] should be used for understanding the security
considerations related to the formats and mechanisms used in this
document.

6. IANA Considerations

This document has no IANA actions.

7. Acknowledgments

The authors would like to thank Adrian Farrel for his detailed
comments and suggestions. The authors would also like to

draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

acknowledge useful comments from David Ward, Sami Boutros, and
Dimitri Papadimitriou.

8. Reference

8.1 Normative Reference

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

[LSP-REROUTE] Berger, L., Papadimitriou, D., and J. Vasseur,
"PathErr Message Triggered MPLS and GMPLS LSP Reroute", draft-
ietf-mpls-gmpls-lsp-reroute (work in progress).

8.2 Informative Reference

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

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

[RFC3630] Katz D., Kompella K., Yeung D., "Traffic Engineering
(TE) Extensions to OSPF Version 2", RFC 3630.

[RFC5305] Smit, H. and T. Li, "Intermediate System to
Intermediate System (IS-IS) Extensions for Traffic Engineering
(TE)", RFC 5305.

[RFC4203] Kompella, K., Ed., and Y. Rekhter, Ed., "OSPF
Extensions in Support of Generalized Multi-Protocol Label
Switching (GMPLS)", RFC 4203.

[RFC5307] Kompella, K., Ed., and Y. Rekhter, Ed., "Intermediate
System to Intermediate System (IS-IS) Extensions in Support of
Generalized Multi-Protocol Label Switching (GMPLS)", RFC 5307.

[RFC3471] Berger, L., "Generalized Multi-Protocol Label
Switching (GMPLS) Signaling Functional Description", RFC 3471.

[RFC3473] Berger, L., "Generalized Multi-Protocol Label

Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic
Engineering (RSVP-TE) Extensions", RFC 3473.

[RFC4726] Farrel A, Vasseur, J.-P., Ayyangar A., "A Framework for
Inter-Domain MPLS Traffic Engineering", RFC 4726, November 2006.

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

_____

        draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

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

[MPLS-GMPLS-SECURITY] Luyuan Fang, Ed. "Security Framework for
MPLS and GMPLS Networks", draft-ietf-mpls-mpls-and-gmpls-
security-framework, work in progress.

9. Authors' Address:

Zafar Ali
Cisco systems, Inc.,
Email: zali@cisco.com

Jean Philippe Vasseur
Cisco Systems, Inc.
Email: jpv@cisco.com

Anca Zamfir
Cisco Systems, Inc.
Email: ancaz@cisco.com

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

draft-ietf-ccamp-mpls-graceful-shutdown-10.txt

10. Copyright Notice

   Copyright (c) 2009 IETF Trust and the persons identified as the
   document authors. All rights reserved.

   This document is subject to BCP 78 and the IETF Trust's Legal
   Provisions Relating to IETF Documents in effect on the date of
   publication of this document (http://trustee.ietf.org/license-info).
   Please review these documents carefully, as they describe your
   rights and restrictions with respect to this document.

11. Legal

   This documents and the information contained therein 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 THEREIN WILL NOT
   INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY
   OR FITNESS FOR A PARTICULAR PURPOSE.

   This document may contain material from IETF Documents or IETF
   Contributions published or made publicly available before November
   10, 2008. The person(s) controlling the copyright in some of this
   material may not have granted the IETF Trust the right to allow
   modifications of such material outside the IETF Standards Process.
   Without obtaining an adequate license from the person(s)
   controlling the copyright in such materials, this document may not
   be modified outside the IETF Standards Process, and derivative
   works of it may not be created outside the IETF Standards Process,

except to format it for publication as an RFC or to translate it
into languages other than English.