

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 10, 2011

K. Hu
J. Luo
B. Wu
ZTE Corporation
July 9, 2010

**LDP Graceful Restart for Pseudowire
draft-jiang-pwe3-ldp-gr-01.txt**

Abstract

LDP graceful restart mechanism defined in [[RFC 3478](#)] can be applied to pseudowire graceful restart. This document introduces the LDP graceful restart capability to allow LSR to negotiate FEC element graceful restart capability with its neighbor. And then it describes an optimized pseudowire graceful restart (GR) mechanism that helps to minimize the negative effects on pseudowire traffic during the process of pseudowire graceful restart.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on January 10, 2011.

Copyright Notice

Copyright (c) 2010 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 (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

- [1. Introduction](#) [3](#)
- [2. Conventions and Terminology used in this document](#) [3](#)
- [3. Motivation and problem statement](#) [3](#)
- [4. LDP graceful restart capability encoding](#) [5](#)
- [5. PW graceful restart negotiation and procedure](#) [7](#)
 - [5.1. Procedures for restarting PE](#) [8](#)
 - [5.2. Procedures for helper PE](#) [8](#)
- [6. Compatibility considerations](#) [9](#)
- [7. IANA Considerations](#) [9](#)
- [8. Acknowledgement](#) [9](#)
- [9. Normative references](#) [9](#)
- [Authors' Addresses](#) [10](#)

1. Introduction

LDP graceful restart mechanism defined in [[RFC 3478](#)] can be applied to pseudowire graceful restart. This document introduces the LDP graceful restart capability to allow LSR to negotiate FEC element graceful restart capability with its neighbor. And then it describes an optimized pseudowire graceful restart (GR) mechanism that helps to minimize the negative effects on pseudowire traffic during the process of pseudowire graceful restart.

[[RFC 4447](#)] defines extension to LDP for pseudowire setup and maintenance. [[RFC 3478](#)] defines graceful restart mechanism for LDP, in which LSRs are enhanced with the ability to preserve MPLS forwarding state during control plane restart. After control plane restart and LDP session re-established, the LDP control plane can be restored. The mechanism defined in [[RFC 3478](#)] is for MPLS LSP (FEC for IP prefix).

Because pseudowire defined in [[RFC 4447](#)] use LDP as the signaling mechanism, the LDP graceful restart mechanism defined in [[RFC 3478](#)] can also be applied to pseudowire FEC, but there can be some optimization for the graceful restart procedure defined in [[RFC 3478](#)] when applying for pseudowire. The purpose of this draft is to specify the procedures and extensions to [[RFC 3478](#)] for pseudowire control plane graceful restart.

2. Conventions and Terminology used in this document

The term "LSP-FEC" in this document means Prefix FEC defined in [RFC 5036 section 3.4](#), and the corresponding FEC element type value is 0x02.

The term "PW-FEC" in this document means Pwid or Generalized Pwid FEC defined in [RFC 4447 section 5.1](#), the Pwid FEC element type value is 0x80, the Generalized Pwid FEC element type value is 0x81.

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](#)].

3. Motivation and problem statement

According to the mechanism described in [[RFC 3478](#)], if PE support pseudowire graceful restart, PE indicates that it is capable of supporting LDP Graceful Restart, as defined in [[RFC 3478](#)], by including the Fault Tolerant (FT) Session TLV as an Optional

Parameter in the LDP Initialization message. The format of the FT Session TLV is defined in [RFC 3479]. The L (Learn from Network) flag MUST be set to 1, which indicates that the procedures in [RFC 3478] are used. The rest of the FT flags are set to 0 by a sender and ignored on receipt. But there is no indication for PE to indicate to its neighbor of supporting pseudowire graceful restart.

If PE pseudowire graceful restart capability share the same capability with LSP-FEC (means Prefix FEC defined in [RFC 5036 section 3.4](#)), it will have some problem as described below.

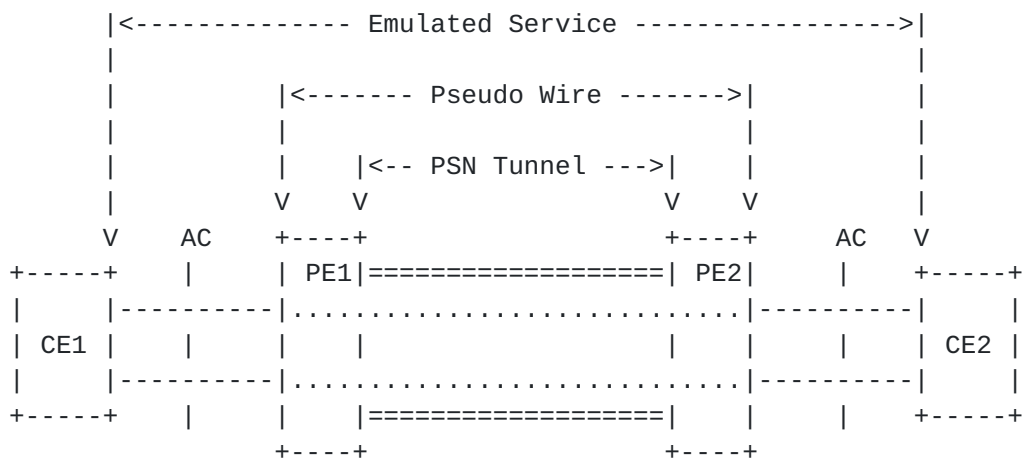


Figure 1

For the scenario in figure 1, the two PEs are connected and share one LDP session with LSP-FEC and PW-FEC. And PE1 supports graceful restart for both LSP-FEC and PW-FEC, while PE2 only supports graceful restart for LSP-FEC. After the negotiation as described in [RFC 3478], PE1 will assume PE2 also supports graceful restart for both LSP-FEC and PW-FEC, while PE2 will assume PE1 also only supports graceful restart for LSP-FEC.

After PE2 restarts its control plane, PE1 detects that its LDP session with a neighbor went down, and knows that the neighbor is capable of preserving its LSP-FEC and PW-FEC forwarding state across the restart (as was indicated by the FT Session TLV in the Initialization message received from the neighbor), then PE1 retains the label-FEC (the FEC includes LSP-FEC and PW-FEC) bindings received via that session, and tries to re-establish LDP communication with the neighbor following the usual LDP procedures. But the fact is that PE2 will not retain PW-FEC in the data plane, and remove the label and PW-FEC binding. The traffic black hole will be caused on PE1 until recovery time out, because PE1 will continue forwarding PW

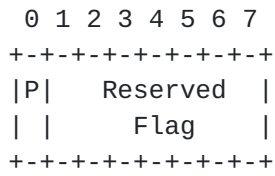


Figure 3

The most significant bit is defined as the Preserve Forwarding State (P) bit, which can be used to indicate whether the forwarding state for FEC that were advertised with the given FEC element type has indeed been preserved during the previous LDP restart. When set (value 1), the bit indicates that the forwarding state has been preserved.

The remaining bits are reserved and MUST be set to zero by the sender and ignored by the receiver.

FEC element type: One octet FEC Element Type that specifies the FEC Element Type that supports graceful restart. Please see [section 3.4.1 of \[RFC5036\]](#) and [\[RFC4447\]](#).

Several FT FEC capability TLVs can be appended behind FT Session TLV, and carried in LDP initial message, and each FT FEC capability TLV will carry a FEC element type which indicates the FEC that support graceful restart. See figure 4.

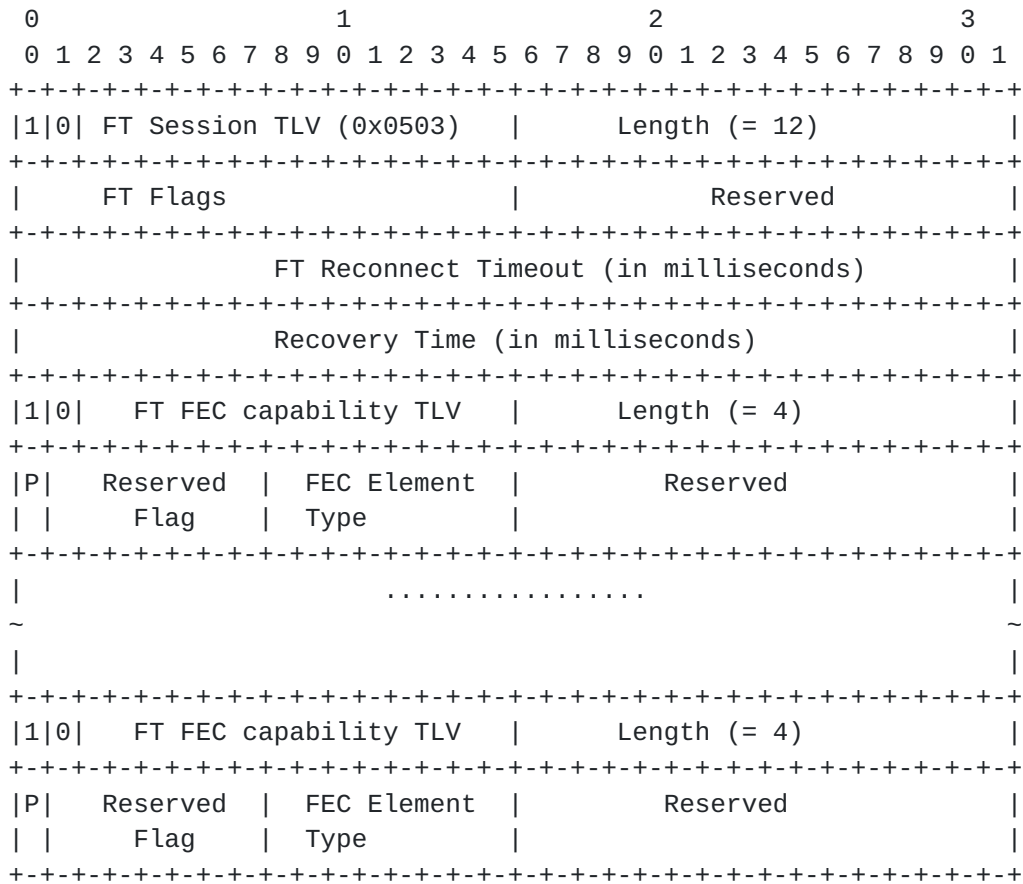


Figure 4

Note: in order to be compatible with the definition in [RFC 3479](#), if there is not FT FEC capability TLV carried in LDP initialization message, the default FEC that support graceful restart SHOULD FEC type of prefix. If there is at least one FT FEC capability TLV presented in LDP initialization message, only the FEC type that carried in FT FEC capability TLV SHOULD be considered to support graceful restart.

5. PW graceful restart negotiation and procedure

A PE will indicate that it is capable of supporting PW graceful restart by including Fault Tolerant (FT) FEC type TLV appended to FT session TLV as an Optional Parameter in the LDP Initialization message. By negotiating the FEC element type in FT FEC capability TLV, the PE can ensure which kind of FEC element type support graceful restart by both endpoints of PE. PW graceful restart function will be only enabled locally when both local and remote PE carry the FEC element type value 128/129 defined in [[RFC 4447](#)].

When one of the PE restarts its L2VPN/PW control plane, the LDP session between the two PEs will be DOWN. When PW graceful restart is supported, the two PE SHOULD preserve its PW-FEC forwarding state to continue traffic forwarding. The PE restarting its L2VPN/PW control plane will be restarting PE, and PE detecting LDP communication down will be helper. See the following sections for detail procedures.

5.1. Procedures for restarting PE

After a PE restarts its control plane, the PE MUST check whether it was able to preserve its PW-FEC forwarding state from prior to the restart. If not, then the PE sets the P bit to 0 in the FT FEC capability TLV the PE sends to its peer.

The PE restarting its L2VPN/PW control plane will only preserve the FEC forwarding state only for the FEC that supports graceful restart at both end points. If the PE knows that the neighbor is only capable of preserving its LSP-FEC forwarding state across the restart, by negotiating through FT FEC capability TLV, it will delete the PW-FEC forwarding state immediately, but it will preserve the LSP-FEC forwarding state. If the PE knows that the neighbor is capable of preserving both LSP-FEC and PW-FEC forwarding state across the restart, by negotiating through FT FEC capability TLV, it will preserve both LSP-FEC and PW-FEC forwarding state.

The other procedures will follow the description in [section 3.1](#) and 3.2 of [[RFC 3478](#)].

5.2. Procedures for helper PE

The PE detecting that its LDP session with a neighbor went down, knows that the neighbor is only capable of preserving its LSP-FEC forwarding state across the restart, by negotiating through FT FEC capability TLV, it will delete the PW-FEC and label binding immediately, but it will retains the LSP-FEC and label bindings received via that session and marks them as "stale ". If the PE knows that the neighbor is capable of preserving both LSP-FEC and PW-FEC forwarding state across the restart, by negotiating through FT FEC capability TLV, it will retains the label-FEC (includes both LSP-FEC and PW-FEC) bindings received via that session (rather than discarding the bindings), but marks them as "stale".

The next procedure will follow the description in [section 3.3](#) in [[RFC 3478](#)].

6. Compatibility considerations

Currently some products have already support PW graceful restart based on [[RFC3478](#)], and PW graceful restart shares same capability with LDP prefix based LSP graceful restart. For the PE not supporting the feature described in this draft will ignore the FT FEC capability TLV advertised by neighbor. And for the PE supporting the feature described in this draft will consider the neighbor only support LSP-FEC graceful restart, and assume the neighbor not supporting PW graceful restart function. Then only graceful restart for LSP-FEC will be enabled for session between the two peers. The PE can disable the function described in this draft, so as to be fully compatible with its neighbor.

7. IANA Considerations

This document creates a new type TLV of FT FEC capability TLV that is to be managed by IANA. This document requires allocation of FT FEC capability TLV type value: 0x050C.

8. Acknowledgement

[Editor's note] Will be added in future.

9. Normative references

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3478] Leelanivas, M., Rekhter, Y., and R. Aggarwal, "Graceful Restart Mechanism for Label Distribution Protocol", [RFC 3478](#), February 2003.
- [RFC3479] Farrel, A., Brittain, P., and P. Matthews, "Fault Tolerance for the Label Distribution Protocol (LDP)", [RFC 3479](#), February 2003.
- [RFC4447] Martini, L., "Pseudowire Setup and Maintenance using LDP", [RFC 4447](#), April 2006.
- [RFC5036] Andersson, L., Minei, I., and B. Thomas, "LDP Specification", [RFC 5036](#), October 2007.

Authors' Addresses

Kan Hu
ZTE Corporation
68, Zijinghua Road
Nanjing, 210012, China

Email: hu.kan@zte.com.cn

Jian Luo
ZTE Corporation
68, Zijinghua Road
Nanjing, 210012, China

Email: luo.jian@zte.com.cn

Bo Wu
ZTE Corporation
68, Zijinghua Road
Nanjing, 210012, China

Email: wu.bo@zte.com.cn

Lizhong Jin
ZTE Corporation
889, Bibo Road
Shanghai, 201203, China

Email: lizhong.jin@zte.com.cn

