

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: September 8, 2016

Q. Wang, Ed.  
ZTE  
March 7, 2016

RSVP-TE Signaling Extensions in support of Flexible Ethernet networks  
draft-wang-ccamp-flexe-signaling-00

Abstract

This draft describes the extensions to the Resource Reservation Protocol Traffic Engineering (RSVP-TE) signalling protocol to support Label Switched Paths (LSPs) in a GMPLS-controlled flexible Ethernet network.

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 September 8, 2016.

Copyright Notice

Copyright (c) 2016 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 . . . . .	2
1.1. FlexE Terminology . . . . .	3
1.2. Requirements Language . . . . .	3
2. Requirements for FlexE Signalling . . . . .	3
2.1. FlexE Group . . . . .	4
2.2. PHY Number . . . . .	4
2.3. Partial-rate . . . . .	4
2.4. Slot Assignment . . . . .	4
2.5. Granularity . . . . .	5
3. Protocol Extensions . . . . .	5
3.1. Traffic Parameters . . . . .	5
3.2. Generalized Label . . . . .	5
3.3. Flag extensions in Hop Attributes TLVs . . . . .	7
3.4. Signalling Procedure . . . . .	7
3.4.1. Procedure for FlexE over Ethernet PHYs . . . . .	7
3.4.2. Procedure for FlexE over ODU LSPs . . . . .	9
4. IANA Considerations . . . . .	11
5. Security Considerations . . . . .	11
6. Manageability Considerations . . . . .	11
7. References . . . . .	11
7.1. Normative References . . . . .	11
7.2. Informative References . . . . .	11
Author's Address . . . . .	12

## 1. Introduction

Flex Ethernet (FlexE) technology, which is defined by Optical Internetworking Forum (OIF), is a new kind of data plane technology and can be used to provides a generic mechanism for supporting a variety of Ethernet MAC rates that may or may not correspond to any existing Ethernet PHY rate. This includes MAC rates that are both greater than (through bonding) and less than (through sub-rate and channelization) the Ethernet PHY rates used to carry FlexE.

FlexE can provide a generic mechanism for supporting a variety of Ethernet MAC rates that may or may not correspond to any existing Ethernet PHY rate. In a more detailed description, FlexE employs more than one Ethernet PHYs as server layer and these Ethernet PHYs are bonded together as a FlexE group to carry FlexE client signal. The general capabilities supported by FlexE implementation includes:

Bonding of Ethernet PHYs, e.g., supporting a 200G MAC over two bonded 100GBASE-R PHYs.

Sub-rates of Ethernet PHYs, e.g., supporting a 50G MAC over a 100GBASE-R PHY.

Channelization within a PHY or a group of bonded PHYs, e.g., support a 150G and two 25G MACs over two bonded 100GBASE-R PHYs.

Note that hybrids are also possible, for example a sub-rate of a group of bonded PHYs, for example, a 250G MAC over three bonded 100GBASE-R PHYs.

In order to operate on the Ethernet PHYs, FlexE mechanism operates using a calendar which assigns slot positions on sub-calendars on each PHY of the FlexE group to each of the FlexE clients. The calendar has a granularity of 5G, and has a length of 20 slots per 100G of FlexE group capacity.

[FLEXE-FWK] defines a framework and the associated control plane requirements for the Generalized Multi-Protocol Label Switching (GMPLS) [RFC3945] based control of FlexE networks.

Based on the requirements described in FlexE framework document, this document defines additional requirements and protocol extensions to Resource Reservation Protocol-Traffic Engineering (RSVP-TE) [RFC3473] to set up LSPs in networks that support the FlexE.

### 1.1. FlexE Terminology

For terminology related to flexible Ethernet, please refer to [FLEXE-FWK] and FlexE Implementation Agreement.

### 1.2. Requirements Language

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].

## 2. Requirements for FlexE Signalling

The architecture for establishing LSPs in a FlexE network is described in [FLEXE-FWK].

The FlexE LSP is a control-plane representation of a FlexE Connection and MUST be carried by Ethernet PHYs LSP or ODU LSP in the network. So in order to set up an end-to-end FlexE LSP, Ethernet PHY LSPs or ODU LSPs MUST be set up in advance. [FLEXE-FWK] gives a description of FlexE layer resource information needed to be reserved when signalling an end-to-end LSP. This section gives a detailed description of these resource information. Based on these information, source equipment and destination equipment will be able to map and demap the FlexE clients from the FlexE group properly.

### 2.1. FlexE Group

As defined in FlexE Implementation Agreement, the FlexE Group refers to a group of from 1 to n bonded 100GBASE-R Ethernet PHYs. Ethernet PHYs are bounded together and used as a whole by the FlexE LSP. FlexE LSPs between the same source and destination equipment SHOULD NOT have the same FlexE group number. Source equipment and destination equipment SHOULD be aware of the existing of FlexE group and which Ethernet PHYs are in which FlexE group.

This FlexE group information MUST be carried in the generalized label of signalling message during LSP establishment if FlexE group is needed.

### 2.2. PHY Number

The PHY number is a dynamic and logical number that is assigned through control plane or management plane, which is unique within the context of (source, destination), and has a one-to-one correlation with physical port. This information will also be carried in the FlexE overhead. Source equipment and destination equipment SHOULD negotiate a value for every Ethernet PHYs within one FlexE group.

The PHY number information MUST be carried in the generalized label of signalling message during LSP establishment. Besides the PHY number carried in the generalized label, RSVP\_HOP object MUST also be used to indicate the correlation between PHY number and physical port number. The sequence of the PHY numbers listed in the generalized label SHOULD be in accordance with the physical ports carried in RSVP\_HOP object.

### 2.3. Partial-rate

The partial-rate capability is usually supported by an OTN access equipment. If an equipment supports partial-rate, it means this equipment has the capability of discarding unavailable slots and transfers the left slots across OTN transport network. During the process of resource allocation, where the partial-rate would happen should be indicated.

### 2.4. Slot Assignment

The FlexE LSP transfers based on the slot positions, so the equipment SHOULD be able to tell which slot is assigned to which client according to the generalized label carried in the signalling message. This attribute SHOULD also take the unavailable slots information into consideration.

## 2.5. Granularity

Currently, only one kinds of 5G slot granularity is defined in OIF Flex Ethernet (FlexE) Implementation Agreement. During signalling process, this information can be inferred through the bandwidth parameters and slot number information within one Ethernet PHY.

## 3. Protocol Extensions

This section defines the extensions to RSVP-TE signalling for GMPLS [RFC3473] to support FlexE networks.

### 3.1. Traffic Parameters

In RSVP-TE, the SENDER\_TSPEC object in the Path message characterizes the traffic parameters for the data flow from the corresponding sender. The FLOWSPEC object in the Resv message indicates the actual resource reservation. As defined in [RFC3473], bandwidth encodings are carried in the SENDER\_TSPEC and FLOWSPEC objects, and these values are set in the Peak Data Rate field of Int-Serv objects, see [RFC2210]. Other bandwidth/service related parameters in the object are ignored and carried transparently. Signalling procedure of RSVP-TE used in FlexE network can also reuse the SENDER\_TSPEC object defined in [RFC2210] to describe the traffic parameters for the data flow of the sender.

### 3.2. Generalized Label

In the case of FlexE network, the GMPLS labels are used to locally represent the FlexE connections and its associated slots transferring. Parameters defined in section 3 are needed to be carried in the generalized label to represent the FlexE connections.

The following is the GENERALIZED\_LABEL object format that is used with the TDM Switching Type:

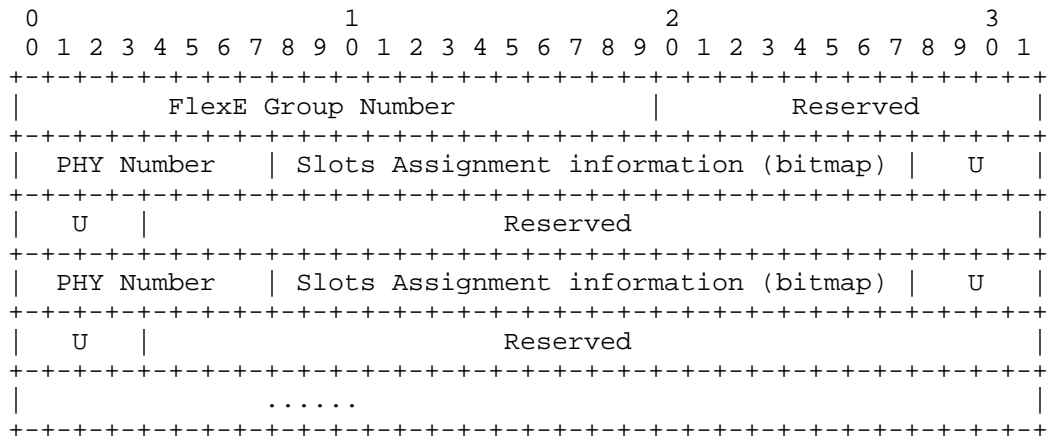


Figure 1: FlexE Label

FlexE Group Number: 20 bits

The FlexE Group refers to a group of from 1 to n bonded 100GBASE-R Ethernet PHYs.

PHY number: 8 bits

The PHY number is a dynamic and logical number that is assigned through control plane or management plane, which is unique within the context of (source, destination), and has a one-to-one correlation with physical port. This information will also be carried in the FlexE overhead.

Slots Assignment information (bitmap) : 20 bits

This attribute is used to indicate slots assignment information, including slots assigned for the client, which are indicated by the bit set to "1" and slots unused, which are indicated by the bit set to "0".

U field: 8 bits

This field is used to indicate the number of unavailable slot. As defined in OIF FlexE Implementation Agreement, unavailable slots are always at the end of the sub-calendar configuration for the respective PHY, so this draft uses a specific field to describe them.

Note: the number of the Ethernet PHY used by FlexE can be referred from the number of the "PHY number" in the generalized label.

### 3.3. Flag extensions in Hop Attributes TLVs

The Attribute Flags TLV defined in [RFC5420] is carried in an ERO Hop Attributes subobject. Flags set in the Attribute Flags TLV [RFC5420] carried in an ERO Hop Attributes subobject SHALL be interpreted in the context of the received ERO. Only a subset of defined flags are defined as valid for use in Attribute Flags TLV carried in an ERO Hop Attributes subobject. Invalid flags SHALL be silently ignored.

A bit in the Attribute Flags TLV is assigned to indicate the partial-rate mux and demux. This ERO Hop Attributes subobject MUST come in pairs. The node which do the partial-rate mux MUST check the existence of partial-rate demux flag in the ERO Hop Attributes subobject of the path message. If it does not exist, path will not be set up successfully.

### 3.4. Signalling Procedure

This section gives description of FlexE layer model in different cases. Figure 2 depicts a FlexE layered network scenario. In this network, B and E are FlexE capable nodes, C and D are OTN ODUflex/ODU4 switch capable nodes. Node B, C are mainly used to encapsulate the client layer signal into the server layer, while node D, E are mainly used to extract the client layer signal from the server layer signal.

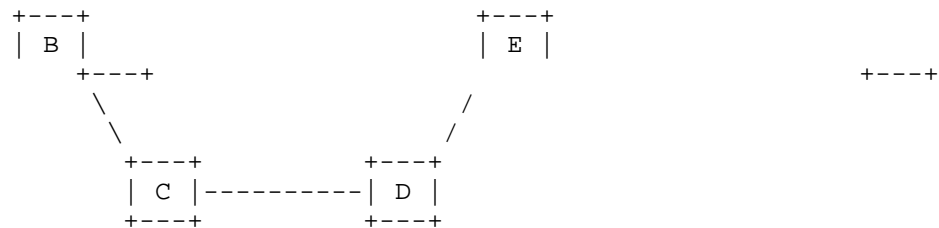


Figure 2: FlexE Layer Network

#### 3.4.1. Procedure for FlexE over Ethernet PHYs

In this case, node C and node D are unaware of the FlexE signal.

Suppose node B receives a FlexE client path set up request from node B to node E and the bandwidth of this path is 150G. There are three Ethernet PHYs between B and C, D and E. Also there exist enough ODU4 connections between node C and node D to bear the traffic from node B. Following is the signalling procedure.

a. Node B intends to send the RSVP-TE path message to set up an end-to-end path towards the destination node E. The bandwidth requirement is 150G.

b. Before node B begins to transfer the FlexE client path message, it will first assert a new FlexE path needs to be set up from node B to node E to carry the Ethernet client traffic by comparing the switching capability carried in the FlexE client path message and the switching capability that node B can support. Node B first blocks the Ethernet client path message, then initiate another new path message to set up the FlexE path from node B to node E. The requested switching capability of the FlexE path is set to TDM and the encoding type is set to "FlexE-LSP".

Before node B send the FlexE path message towards the next hop, node B first check if there are enough Ethernet PHYs to carry the FlexE traffic. Then node B will start the set up of two Ethernet PHYs LSP from node B to node E. The Ethernet PHY path messages are then sent towards the next hop node C. Two Ethernet PHY LSPs will be set up.

c. When node C receives the Ethernet PHY path messages from node B, it will first assert two new ODU4 path needs to be set up from node C to node D to carry the Ethernet PHY traffic by comparing the switching capability carried in the path message and the switching capability that node C can support. Node C first blocks the Ethernet PHY path message, then initiate another new path messages to set up the ODU4 path from node C to node D.

Considering the set up of ODU4 path is not the focus of this draft, procedure of setting up ODU4 paths are not going to be described in this draft. Node C will receive the Resv message from node D and confirm the successful set up of ODU4 paths. The ODU4 LSPs behave as the server layer of Ethernet PHY paths.

d. Node C will then continue sending the Ethernet PHY path messages towards node E to finish the set up of Ethernet PHY LSPs. The Ethernet PHY LSPs from node B to node E behave as the a link with only one hop.

e. After node B receives the Resv message from node E and confirm the successful set up of Ethernet PHY paths, node B will continue sending the FlexE path message towards the next hop node E. The bandwidth requirement is 150G.

f. Node E receives the FlexE path message from node B. Considering there are already two Ethernet PHYs from node B to node E, node E determines to set up the FlexE path over the two Ethernet PHYs by



carrying the assigned FlexE group number, dynamic PHY number and slot positions for the Ethernet client in the generalized label.

Besides the generalized label, RSVP\_HOP object MUST also be used to indicate the correlation between PHY number and physical port number. The sequence of the PHY numbers listed in the generalized label SHOULD be in accordance with the physical ports number carried in RSVP\_HOP object.

Node E then send the assembled Resv message toward the FlexE path source, node B, to finish the set up of FlexE path.

g. After node B receive the Resv message from node E and confirm the successful set up of FlexE path, node B will continue sending the blocked FlexE client path message towards the destination node E to finish the set up of the client path.

FlexE LSP is demonstrated as an Ethernet client link once it was set up. Unused bandwidth on the FlexE LSP can be used further by another FlexE client.

#### 3.4.2. Procedure for FlexE over ODU LSPs

In this case, node C and node D are aware of the FlexE signal.

Suppose node B receives a FlexE client path set up request from node B to node E and the bandwidth of this path is 150G. There are three Ethernet PHYs between B and C, D and E. Also there exist 180G ODUFlex connections between node C and node D to bear the traffic from node B. The number of unavailable slots between B and C is 4, between D and E is 5. Following is the signalling procedure.

a. Node B intends to send the RSVP-TE path message to set up an end-to-end path towards the destination node E. The bandwidth requirement is 150G.

b. Before node B begins to transfer the FlexE client path message, it will first assert a new FlexE path needs to be set up from node B to node E to carry the FlexE client traffic by comparing the switching capability carried in the FlexE client path message and the switching capability that node B can support. Node B's computation element will compute an end-to-end partial-rate supported path B-C-D-E. Node B then first blocks the Ethernet client path message, then initiate another new path message to set up the FlexE path from node B to node E. The requested switching capability of the FlexE path is set to TDM and the encoding type is set to "Partial-rate FlexE-LSP".

c. Before node B send the FlexE path message towards the next hop, node B first check if there are enough Ethernet PHYs to carry the FlexE traffic. Then node B will start the set up of two Ethernet PHYs LSP from node B to node E with the G-PID field set to "Partial-rate FlexE-LSP". The Ethernet PHY path messages are then sent towards the next hop node C. Two Ethernet PHY LSPs will be set up.

d. When node C receives the Ethernet PHY path messages which are used to support partial-rate FlexE-LSP from node B, it will first check if itself support partial-rate capability as a result of the "Partial-rate FlexE-LSP" carried in G-PID. Node C then sends the Resv message to node B to finish the set up of Ethernet PHYs LSP if it confirms the support of partial-rate capability.

e. When node B receives the Ethernet PHYs resv message from node C, it will first finish the set up of Ethernet PHYs LSP, then continue sending the FlexE path message towards next hop node C. These Ethernet PHY LSPs behave as a one hop FlexE link in the FlexE network.

f. When node C receives the FlexE path message, node B first block the FlexE path message, then checks if there are ODUFlex LSP available to carry the FlexE traffic. If no, node B first initiate the ODUFlex LSP set up and the bandwidth requested does not take the unavailable slots into consideration.

Considering the set up of ODUFlex path is not the focus of this draft, procedure of setting up ODUFlex paths are not going to be described in this draft. Node C will receive the Resv message from node D and confirm the successful set up of ODUFlex paths. The ODUFlex LSPs behave as a one hop FlexE link in the FlexE network.

g. When node C receives the ODUFlex Resv message from node D. It will first finish the set up of ODUFlex LSP, then continue sending the FlexE path message towards next hop node D. These ODUFlex LSPs behave as a one hop FlexE link in the FlexE network and this link carries all FlexE slots except unavailable slots.

h. When node D receives the FlexE path message, it will repeat the procedure in c, d, e. Node E will receive the FlexE path message sent from node B and it will construct the Resv message sent back to node B to finish the resource reservation along the path.

i. After node B receives the FlexE Resv message, it first finish the path set up and then continue the blocked FlexE client signalling procedure.

#### 4. IANA Considerations

In this draft, two new encoding types "FlexE-LSP" and "Partial-rate FlexE-LSP" are assigned for FlexE LSP.

#### 5. Security Considerations

TBD

#### 6. Manageability Considerations

TBD

#### 7. References

##### 7.1. Normative References

- [FlexE] Stephen, J. and David. R. Stauffer, "FlexE Implementation Agreement", 2016.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <<http://www.rfc-editor.org/info/rfc3209>>.
- [RFC3471] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, DOI 10.17487/RFC3471, January 2003, <<http://www.rfc-editor.org/info/rfc3471>>.
- [RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <<http://www.rfc-editor.org/info/rfc3473>>.

##### 7.2. Informative References

- [RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, DOI 10.17487/RFC3945, October 2004, <<http://www.rfc-editor.org/info/rfc3945>>.

Author's Address

Qilei Wang (editor)  
ZTE  
Nanjing  
CN

Email: wang.qilei@zte.com.cn