

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: September 12, 2019

Q. Wang, Ed.  
X. Niu, Ed.  
ZTE Corporation  
Y. Xu  
CAICT  
March 11, 2019

Analysis for FlexE control model  
draft-wang-ccamp-flex-e-control-analysis-00

Abstract

This document gives some analysis about the control of FlexE.

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 <https://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 12, 2019.

Copyright Notice

Copyright (c) 2019 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 (<https://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
2. Terminology . . . . .	3
2.1. Requirements Language . . . . .	3
3. Analysis . . . . .	3
3.1. General Introduction of FlexE . . . . .	3
3.1.1. FlexE Group . . . . .	3
3.1.2. FlexE Client . . . . .	3
3.1.3. Adaptation function between FlexE Client and FlexE Group . . . . .	4
3.1.4. MAC Frame . . . . .	5
3.1.5. Adaptation between MAC frames and FlexE Client . . . . .	5
3.2. General requirements . . . . .	5
3.2.1. Configuration of FlexE group . . . . .	5
3.2.2. Allocate Resources for Client MAC flows . . . . .	6
3.3. FlexE Client Configuration Procedures . . . . .	6
3.4. Control Requirements Derived . . . . .	8
4. Summary . . . . .	9
5. Acknowledgements . . . . .	9
6. IANA Considerations . . . . .	9
7. Security Considerations . . . . .	9
8. References . . . . .	9
8.1. Normative References . . . . .	9
8.2. Informative References . . . . .	10
Authors' Addresses . . . . .	10

## 1. Introduction

OIF published the first version of FlexE Implementation Agreement in March 2016, aiming to provide a generic mechanism for supporting a variety of Ethernet MAC rates that may or may not correspond to any existing Ethernet PHY rate. SG15 in ITU-T has endorsed the OIF FlexE data plane and parts of [ITU-T G.872], [ITU-T G.709], [ITU-T G.798] and [ITU-T G.8023]. The Recommendations depend on or are based on the FlexE data plane.

This draft is intended to trigger discussion of the FlexE control architecture according to the analysis in section 2. What kind of architecture should we employed when configuring FlexE equipments, how to configure the FlexE group and FlexE client, and what kind of parameters do we need to take into consideration? The analysis is mainly based on the description in section 7 and 8 of [ITU-T G.8023], which is "Characteristics of equipment functional blocks supporting Flex Ethernet interfaces".

## 2. Terminology

### 2.1. 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].

## 3. Analysis

### 3.1. General Introduction of FlexE

The FlexE shim is built into the Ethernet PCS (physical coding sublayer). If a FlexE group is set up, a corresponding  $n \times 100\text{G}$  (or  $n \times 200\text{G}$ ,  $n \times 400\text{G}$ ) PCS module is created as well.

The difference between the FlexE and the traditional 100G Ethernet is that the traditional Ethernet PCS has a 1:1 relationship with the client MAC flow, while with FlexE one bonded huge PCS module can be used to transport more than one client MAC flow i.e., the relationship is 1:n.

#### 3.1.1. FlexE Group

A FlexE Group is consisted of 1 to  $m$  bonded Ethernet PHYs. The rate of the Ethernet PHY could be 100G, 200G or 400G. All PHYs in the group must operate at the same rate.

FlexE group is consisted of a number of PHYs, and each PHY is consisted of 66B blocks stream. Section monitoring overhead is added/extracted as one 66B block at the FlexE group source and destination (i.e., trail termination) to determine the status of the FlexE group (i.e., FlexE trail). Currently, only RPF (Remote PHY Fault) indication is used to report the state of FlexE group.

One FlexE group exists between two FlexE shim, there is no switching defined in FlexE. In addition, only one fault indication is defined, there is no other OAM function developed yet. Based on these analysis, we may understand FlexE is just an interface technology, and once a FlexE group is configured, it only functions as one Ethernet link, same as Ethernet PHY.

#### 3.1.2. FlexE Client

A FlexE Client is an Ethernet flow based on a MAC data rate that may or may not correspond to any Ethernet PHY rate. The FlexE Client MAC rates supported by a FlexE Groups could be 10, 40, and  $m \times 25\text{ Gb/s}$ . The FlexE Client MAC rates supported by FlexE Groups may support all,

or only a subset of these FlexE Client rates, e.g.,  $m \times 25$  Gb/s. Each FlexE Client is presented to the FlexE Shim as a 64B/66B encoded bit-stream according to clause 82 of [IEEE 802.3].

According to the description in clause 8.1 of [ITU-T G.8023], there is no overhead defined for monitoring a FlexE client, so the trail for FlexE client in the equipment does not exist. The FlexE client trail termination function is a null function. Therefore, there is no need to model FlexE client as a layer.

### 3.1.3. Adaptation function between FlexE Client and FlexE Group

In order to distribute the FlexE client over PHYs of one FlexE group, a number of management information command should be sent to the adaptation function which performs the mapping of FlexE client over FlexE group.

According to the description in clause 7.2 of [ITU-T G.8023], the external management information command sent to the source adaptation function is listed below:

TxCC, TxCCA, TxCCB, TxCR, TxCA

TxGID, TxPHYMAP

The TxCC, TxCCA and TxCCB are used to configure the calendar for use, which could be type A or type B calendar configuration, and FlexE client number.

TxCR and TxCA are used to coordinate the switch of calendar configuration between the FlexE source and destination node.

The TxGID is used to configure the FlexE group identifier. The TxPHYMAP is used to configure the set of PHYs in the FlexE group.

The built-in function multiplexer performs the action of assigning the individual FlexE Client to specific calendar slots of the FlexE group.

At the destination side, the overhead should be extracted first to compare the GID and PID. The Demultiplexer function activates the FlexE Client and assigns the calendar slots of the FlexE group payload area to the individual FlexE client as defined by the client calendar information carried in the overhead.

#### 3.1.4. MAC Frame

Defined in IEEE.

#### 3.1.5. Adaptation between MAC frames and FlexE Client

It can be seen from the Figure 8-6 of [ITU-T G.8023] that the external management information commands used as input to the adaptation function are defined by [IEEE 802.3]. The [IEEE 802.3] process mainly includes the 64B/66B encoding, as well as MAC frame check sequence generation and frame counting. The FlexE client stream is generated at the determined FlexE Client MAC rate and 64B/66B encoded.

#### 3.2. General requirements

It can be inferred from section 2.1.2 and section 2.1.5 that process mainly involved when producing the FlexE Client from MAC frames is 64b/66b encoding, and this encoding has already been defined by [IEEE 802.3]. No extra overhead is added. Therefore, configuration for FlexE client layer is needed. Based on the above analysis, two general requirements for control/management of FlexE are considered in this draft.

Configuration of FlexE group

Allocation of one or more FlexE group calendar slot resources to a client MAC flow.

##### 3.2.1. Configuration of FlexE group

It can be concluded from the above analysis that external configuration tools should be involved to bring one FlexE group into service. The initial configuration commands can come from external management system, SDN controller etc.

A FlexE group must be configured first before any client signals are carried over it. When a new FlexE Group is brought into service, the initial configuration must be provisioned from both ends, and the initial configuration must be the same. The group is configured to consist of from 1 to n 100G FlexE Instances carried over from 1 to m PHYs of the same rate (100GBASE-R, 200GBASE-R, or 400GBASE-R). Each PHY tunnel may consist of multiple hops.

### 3.2.2. Allocate Resources for Client MAC flows

The FlexE client MAC flows are encapsulated in one or more FlexE calendar slots. Questions that may be raised when considering whether external control/management tools are needed.

How is the bandwidth (number of calendar slots) allocated to a MAC client?

How are the calendar slots assigned to each FlexE clients?

Does the external management/control system need to be involved?

According to the analysis in section 2.1.3, it can be inferred that the built-in multiplexer and demultiplexer function can work together to insert/extract FlexE Client stream from some calendar slots correctly, as well as the trail termination function for FlexE client is a null function. Therefore, only the bandwidth information of the FlexE Client is needed, the number of calendar slots required can be derived from the bandwidth information.

The FlexE client physical port is an internal port which only perform the function of encapsulating upper layer packets into MAC frames, 64b/66b encoding. The bandwidth capability of these internal ports should be known by external management/control tools in order to multiplex and demultiplex the upper layer flow correctly.

### 3.3. FlexE Client Configuration Procedures

Example below is depicted to set up a 25G MPLS-TP P2P path (from MPLS-TP-1 to MPLS-TP-2) over one FlexE link[two FlexE Groups] between Route-1 and Router-2. This is to help understand why control of FlexE client is not needed. FlexE group is assumed configured in this case.

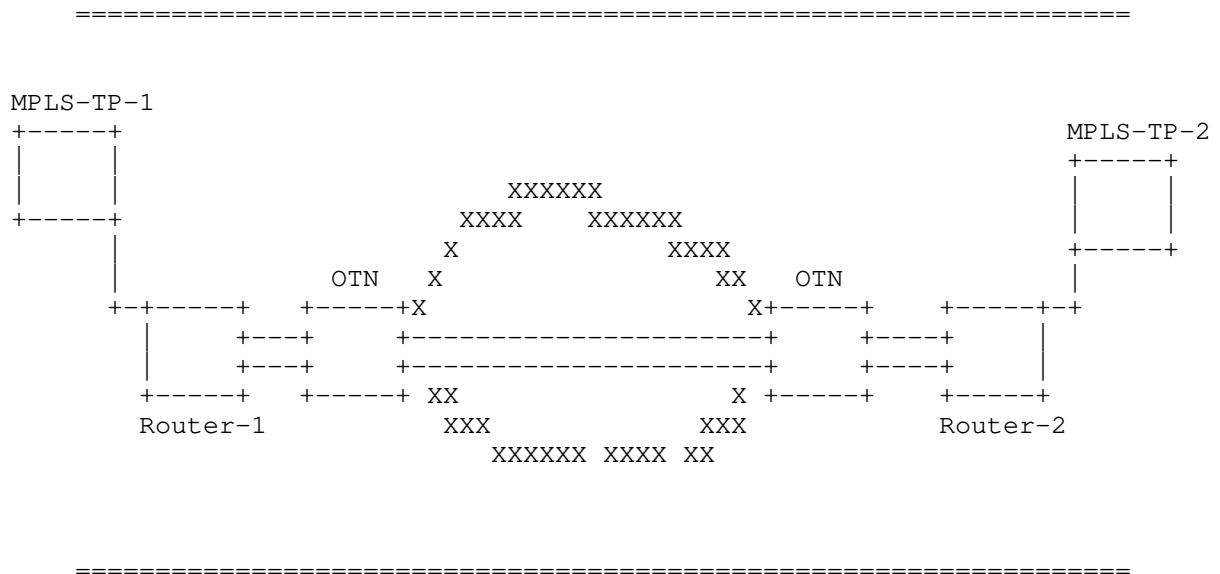


Figure 1: Network Scenario

Configuration needed for each node when setting up MPLS-TP path:

```
Outgoing MPLS label and output port, 25G --> MPLS-TP-1;
```

```
Incoming MPLS label and input port, 25G --> Router-1;
```

```
MPLS traffic over FlexE link, from Router-1 to Router-2;
```

```
Outgoing MPLS label and outgoing port, 25G --> Routing-2;
```

Incoming label and input port, 25G --> MPLS-TP-2

What else would be done by Router-1 and Router-2 if we want to put MPLS packets over FlexE link?

Router-1, perform the generally used data link (i.e., MAC) encapsulation procedures to transmit the packet to next hop Router-2, which includes:

encapsulation of the packets into MAC frames. The source and destination MAC address can be derived by searching Router-1's internal mapping table, i.e., (outgoing label, outgoing port, MAC address) table, then do the 64b/66b encoding for these MAC frames.

in addition, Router-1 would announce the required bandwidth required from Router-1 to Router-2 to local FlexE multiplexer, the local FlexE multiplexer then allocates 5\*5G calendar slots to the MPLS-TP client signal, and insert FlexE overhead onto each PHY.

Router-1's configuration is finished.

Router-2's configuration would be different, as its function is to extract the FlexE client signal from the calendar slots according to the information carried in the FlexE overhead and route the packets based on the MPLS label. As there is no switching of calendar slots in FlexE, the calendar slots allocated to certain client signal would not change.

It can be inferred from the above procedures that the configuration of a MPLS client is not impacted by the use of FlexE comparing to traditional Ethernet.

### 3.4. Control Requirements Derived

- a. Using control plane method to configure FlexE group, which may include the configuration of group number, PHY number and correlation between logical PHY number and physical port number.
- b. Configuration of "logical" PHY, which includes initial configuration of bandwidth (number of calendar slots); change of bandwidth while in service.
- c. External control command should be provide to trigger the switch of calendar slots.
- d. Interworking between 5G slot granularity capable node and 25G slot granularity node.
- e. Configuration of aware case. As calendar slots is not visible to external management/control tools. Bandwidth information of the Ethernet PHYs selected can be used to infer the unavailable slot information, as the unavailable slots are placed at the end of each relevant sub-calendar (the highest numbered slots). [RFC6002] defines a new switching type DCSC (Data Channel Switching Capable) to describe the switching of whole digital channel presented on single channel interface, which can describe the Ethernet terminated at the physical (port) level and all traffic received on a port is switched to a physical port at the LSP egress. The FlexE digital channel presented at each PHY interface can be described with DCSC switching type. However, FlexE aware case may not be depicted with DCSC, as described in section 17.12 of [ITU-T G.709], each FlexE (sub) groups (i.e.,



each 100G FlexE signals) are crunched, padded and interleaved in FlexE, the bandwidth of digital channel links presented in one tunnel may varies. Therefore, a new switching type DCSC-flexE is needed in aware case to depict the end-to-end tunnel with bandwidth varies at each digital channel links. The configuration of FlexE instance and unavailable slot information can be derived through the bandwidth of each hop.

Different kinds of alarms should be taken into consideration when modelling FlexE technology, which may include PHY failed, skew exceed threshold, inconsistent configuration between two ends.

#### 4. Summary

According to the analysis in section 2, the main control/management requirement for FlexE technology is to configure the FlexE group. Once a FlexE group is configured and the capability information of the internal FlexE client ports associated with this FlexE group is known, use of the FlexE technology is the same as that in traditional Ethernet.

#### 5. Acknowledgements

#### 6. IANA Considerations

This memo includes no request to IANA.

#### 7. Security Considerations

None.

#### 8. References

##### 8.1. Normative References

[ITU-T\_G709]

ITU-T, "ITU-T G.709: Optical Transport Network Interfaces; 07/2016", <http://www.itu.int/rec/T-REC-G.709-201606-P/en>, July 2016.

[ITU-T\_G798]

ITU-T, "ITU-T G.798: Characteristics of optical transport network hierarchy equipment functional blocks", August 2018.

## [ITU-T\_G8023]

ITU-T, "ITU-T G.8023: Characteristics of equipment functional blocks supporting Ethernet physical layer and Flex Ethernet interfaces", , 2016.

## [ITU-T\_G872]

ITU-T, "ITU-T G.872: The Architecture of Optical Transport Networks; 2017", <http://www.itu.int/rec/T-REC-G.872/en>, January 2017.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

## 8.2. Informative References

## [I-D.izh-ccamp-flex-e-fwk]

Hussain, I., Valiveti, R., Pithewan, K., Wang, Q., Andersson, L., Zhang, F., Chen, M., Dong, J., Du, Z., zhenghaomian@huawei.com, z., Zhang, X., Huang, J., and Q. Zhong, "GMPLS Routing and Signaling Framework for Flexible Ethernet (FlexE)", draft-izh-ccamp-flex-e-fwk-00 (work in progress), October 2016.

## Authors' Addresses

Qilei Wang (editor)  
ZTE Corporation  
Nanjing  
CN

Email: wang.qilei@zte.com.cn

Xiaobing Niu (editor)  
ZTE Corporation  
Beijing  
CN

Email: niu.xiaobing@zte.com.cn

Yunbin Xu  
CAICT  
Beijing  
CN

Email: xuyunbin@caict.ac.cn