

Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: January 9, 2020

Q. Wang, Ed.
X. Niu, Ed.
ZTE Corporation
Y. Xu
CAICT
July 8, 2019

Analysis for FlexE control
draft-wang-ccamp-flex-e-control-analysis-02

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 January 9, 2020.

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 | 4 |
| 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 Mode for FlexE client | 6 |
| 3.2.2. Configuration of FlexE group | 6 |
| 3.2.3. Allocate Resources for FlexE Client | 7 |
| 3.3. Control Requirements Derived | 7 |
| 4. Summary | 8 |
| 5. Acknowledgements | 8 |
| 6. IANA Considerations | 8 |
| 7. Security Considerations | 8 |
| 8. References | 8 |
| 8.1. Normative References | 8 |
| 8.2. Informative References | 9 |
| Authors' Addresses | 9 |

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 requirements, which can be found in section 2. What kind of model should we employed when configuring FlexE capable equipments, how to configure the FlexE group and FlexE client, and what kind of parameters do we need to take into consideration when configuring FlexE group and FlexE client. The analysis is based on the description in section 7 and 8 of [ITU-T G.8023] and FlexE IA 2.0.

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 with multiple FlexE client ports could be 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, with each MAC flow mapped into one FlexE client.

3.1.1. FlexE Group

A FlexE Group is consisted of from 1 to n 100G FlexE instances, which are carried over from 1 to m 100G, 200G or 400G Ethernet PHYs. All PHYs in the group must operate at the same rate.

FlexE group is consisted of a number of FlexE instances, and each instance 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 in ITU-T terminology). Currently, only RPF (Remote PHY Fault) indication is used to report the state of FlexE group.

The FlexE group exists between two FlexE shim, there is no slot switching defined in FlexE. Only one fault indication is defined, there is no other OAM function developed yet. Based on these analysis, we should be able to understand that FlexE is just an interface technology, and once a FlexE group is configured, it only functions as one Ethernet link, similar to 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 10Gb/s, 40Gb/s, or $m \times 25\text{Gb/s}$. The FlexE Client MAC rates supported by FlexE Groups may support all, or only a subset of these FlexE Client rates. Each FlexE Client is presented to the FlexE Shim as a 64B/66B encoded bit stream according to clause 82 of [IEEE 802.3]. FlexE clients have the semantics of an Ethernet PHY. There is no new layer network. Both FlexE group and FlexE client are processed at Ethernet PHY layer. From the network management perspective, the FlexE client can be created and the calendar slots information of one FlexE group can be allocated to one FlexE client. The FlexE client could be generated internally within a system, or created from a traditional Ethernet PHY. What kind of FlexE clients will be created depends on the operator's needs.

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, modelling FlexE client as a network layer is not correct.

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, slots allocated for a specific FlexE client 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. If

200G and 400G are used, the 100G FlexE instance should be used in the case of PHYMAP, as current version of [ITU-T G.8023] only cover the scope of 100G PHY.

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

At the destination side, the Demultiplexer function could use activate the FlexE Client and assigns the calendar slots of the FlexE group payload area to the individual FlexE client according to external configuration or the client calendar information carried in the overhead. Expected group ID, PHYMAP and calendar allocation information are needed sometimes to help verify the correctness of FlexE configuration.

3.1.4. MAC Frame

Defined in IEEE.

3.1.5. Adaptation between MAC frames and FlexE Client

The external management information commands used as input to the adaptation function are defined by [IEEE 802.3], according to the description in [ITU-T G.8023]. 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 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 during this process. Therefore, configuration for mapping MAC frames into FlexE client from external management system is not needed.

Based on the above analysis, two high-level requirements for control/management of FlexE are considered in this draft.

Configuration mode

Configuration of FlexE group

Creation of FlexE client and allocation of one or more FlexE group calendar slot resources to a FlexE client.

3.2.1. Configuration Mode for FlexE client

There are two different configuration modes for bring one FlexE client into service. The first one is static model, which is to use external management system to configure the FlexE client and resources allocated for the FlexE client at source and destination FlexE shims. In this case, the CR/CA mechanism does not work. Verification of configuration consistency at FlexE source and destination site by comparing the inband FlexE overhead with the configuration at FlexE destination are needed; The other one is MASTER/SLAVE mode, which is to use the FlexE overhead to coordinate the resource configuration between FlexE source and destination, the external resource configuration information is only sent the source node.

3.2.2. 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 could be 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 for both ends, and the initial configuration must be the same for both direction. The group is configured to be 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). A PHY number may correspond to the physical port ordering on equipment, but the FlexE Shim at each end of the group must identify each PHY in the group using the same PHY number, and each 100G FlexE Instance with the same 100G FlexE Instance number. In certain cases, it may be desirable not to populate all 100G FlexE instances on a 200G or 400G PHY, and these so-called unequipped FlexE instance should also be configured. Unequipped instances must always be the highest numbered instance(s) on a PHY of the FlexE Group, and there must always be at least one equipped 100G FlexE Instance on every PHY.

If aware case is needed to be considered, unavailable slot information should be configured at FlexE aware node to discard unavailable slot first, so as to put the rest of available slots onto the lower rate physical port.

3.2.3. Allocate Resources for FlexE Client

The FlexE client MAC flows are encapsulated in one or more FlexE calendar slots.

According to the analysis in section 3.2.1, there are two different configuration modes. For the first one, static mode, after the FlexE group is configured, the FlexE client resource allocation information are sent both to FlexE source and destination to help create the FlexE client. A number of expected configuration parameters are sent to FlexE destination to help verify the correctness of configuration at both sides. Information sent can be found in [draft-xiaobn-ccamp-flexe-yang-mod]. For the Master/slave mode, the FlexE client resource allocation information are only sent to the FlexE source site. The FlexE source site first create the FlexE clients, and then the built-in multiplexer at the FlexE source site allocates the calendar slots to a specific FlexE client according to the input from external management system, and insert these configuration information into the FlexE overhead. When these overheads arrives at the destination site, the demultiplexer function at the destination site extracts FlexE overhead first and get the information of calendar slot allocation information. Based on these information, the FlexE destination site finish the configuration of FlexE clients. In order to verify the correctness of the resource configuration, the expected FlexE group ID, PHY number and instance number information, FlexE client number and slot allocation information for a specific FlexE client should also be configured to FlexE destination site.

The FlexE client 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 be used by the upper layer (e.g., MPLS-TP) flow correctly.

3.3. Control Requirements Derived

- a. Using external control/management system to configure FlexE group, which may include the configuration of group number, PHY number and instance number, as well as correlation between logical PHY number and physical port number. A number of expected configuration parameters are also needed to help verify the consisten between FlexE source and destination.
- b. Using external control/management system to create the FlexE client, which include the FlexE client number, FlexE client type and slots allocation information. Different configuration mode for FlexE client are needed.

- c. External control command could 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 unequipped instance, unavailable slots, which include the number of unequipped instance and number of unavailable slots on each instances

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 and FlexE client. Once a FlexE group is configured and the FlexE client ports is created, slots allocation is configured, 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.
- [I-D.xiaobn-ccamp-flex-e-yang-mod]
NIU, X., Wang, Q., Xu, Y., and S. Munagapati, "A YANG Data Model for Flex Ethernet (FlexE)", draft-xiaobn-ccamp-flex-e-yang-mod-01 (work in progress), May 2019.

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