Common Control and Measurment Plane I. Hussain

Internet-Draft R. Valiveti

Intended status: Informational Infinera Corp

Expires: August 17, 2019 Q. Wang

ZTE

L. Andersson

M. Chen

H. Zheng

Huawei

February 13, 2019

GMPLS Routing and Signaling Framework for Flexible Ethernet (FlexE)

draft-izh-ccamp-flexe-fwk-07

Abstract

This document specifies the GMPLS control plane requirements,

framework, and architecture for the FlexE technology. The document

also discusses interoperation between the GMPLS control plane for

FlexE and the control plane of any networking layer using the FlexE

technology as a server layer.

As different from earlier Ethernet data planes FlexE allows for

decoupling of the Ethernet Physical layer (PHY) and Media Access

Control layer (MAC) rates.

Study Group 15 (SG15) of the ITU-T has endorsed the FlexE

Implementation Agreement from Optical Internetworking Forum (OIF) and

included it, by reference, in some of their Recommendations.

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 August 17, 2019.

Hussain, et al. Expires August 17, 2019 [Page 1]

Internet-Draft FlexE Extensions February 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 . . . . . . . . . . . . . . . . . . . . . . . . 3

1.1. Requirements Language . . . . . . . . . . . . . . . . . . 5

1.2. Updates in the version . . . . . . . . . . . . . . . . . 5

2. Terminology . . . . . . . . . . . . . . . . . . . . . . . . . 6

3. FlexE Reference Model . . . . . . . . . . . . . . . . . . . . 7

4. GMPLS Controlled FlexE . . . . . . . . . . . . . . . . . . . 8

4.1. Interfaces in a FlexE network . . . . . . . . . . . . . . 9

4.2. Mapping of traffic in the data plane . . . . . . . . . . 9

4.3. The GMPLS Control Plane and the FlexE identifiers . . . . 10

4.4. Operational concerns . . . . . . . . . . . . . . . . . . 11

4.5. Pre-configured vs. Control Plane established LSPs in a

FlexE capable network . . . . . . . . . . . . . . . . . . 11

4.6. Signaling Channel . . . . . . . . . . . . . . . . . . . . 12

4.7. MPLS LSP over the FlexE Data Plane . . . . . . . . . . . 12

4.8. Configuring the data plane in FlexE capable nodes . . . . 14

4.8.1. Configure/Establish a FlexE Group/Link . . . . . . . 14

4.8.2. Configure/Establish a FlexE Client . . . . . . . . . 15

4.8.3. Advertise FlexE Groups and FlexE Clients . . . . . . 15

5. Framework and Architecture . . . . . . . . . . . . . . . . . 15

5.1. FlexE Framework . . . . . . . . . . . . . . . . . . . . . 16

5.2. FlexE Architecture . . . . . . . . . . . . . . . . . . . 16

5.2.1. Architecture Components . . . . . . . . . . . . . . . 16

5.2.2. FlexE Layer Model . . . . . . . . . . . . . . . . . . 16

5.2.2.1. FlexE Group structure . . . . . . . . . . . . . . 17

5.2.2.2. FlexE Client mapping . . . . . . . . . . . . . . 17

6. Control Plane . . . . . . . . . . . . . . . . . . . . . . . . 17

6.1. GMPLS Routing . . . . . . . . . . . . . . . . . . . . . . 18

6.2. GMPLS Signaling . . . . . . . . . . . . . . . . . . . . . 18

6.2.1. LSP setup with pre-configured FlexE infrastructure . 19

6.2.2. LSP setup with partially configured FlexE

infrastructure . . . . . . . . . . . . . . . . . . . 20

Hussain, et al. Expires August 17, 2019 [Page 2]

Internet-Draft FlexE Extensions February 2019

6.2.3. LSP setup with non-configured FlexE infrastructure . 21

6.2.4. Packet Label Switching Data Plane . . . . . . . . . . 21

7. Operations, Administration, and Maintenance (OAM) . . . . . . 23

8. Acknowledgements . . . . . . . . . . . . . . . . . . . . . . 23

9. IANA Considerations . . . . . . . . . . . . . . . . . . . . . 23

10. Security Considerations . . . . . . . . . . . . . . . . . . . 23

11. Contributors . . . . . . . . . . . . . . . . . . . . . . . . 23

12. References . . . . . . . . . . . . . . . . . . . . . . . . . 23

12.1. Normative References . . . . . . . . . . . . . . . . . . 23

12.2. Informative References . . . . . . . . . . . . . . . . . 24

Appendix A. Requirements . . . . . . . . . . . . . . . . . . . . 25

Authors' Addresses . . . . . . . . . . . . . . . . . . . . . . . 26

1. Introduction

This document specifies the GMPLS control plane requirements,

framework, and architecture for the FlexE technology. The FlexE

control plane requirements are found in an appendix.

Prior to FlexE Ethernet MAC rates were until constrained to match the

rates of the Ethernet PHY(s). FlexE, specified by the OIF, allows

MAC rates to be different from PHY rates. An OIF implementation

agreement [OIFFLEXE1] allows for complete decoupling of the MAC and

PHY rates. This has been further extended in [OIFFLEXE2].

SG15 in ITU-T has endorsed the OIF FlexE data plane and parts of

[G.872], [G.709], [G.798] and [G.8021]. The Recommendations depends

on or are based on the FlexE data plane.

The FlexE implementation agreement includes support for:

a. MAC rates which are greater than the rate of a single PHY;

multiple PHYs are bonded to achieve this

b. MAC rates which are less than the rate of a PHY (sub-rate)

c. support for channelization within a single PHY, or over a group

of bonded PHYs.

The capabilities supported by the FlexE data plane are:

a. Support a large rate Ethernet MAC over bonded Ethernet PHYs, e.g.

supporting a 200G MAC over 2 bonded 100GBASE-R PHY(s)

b. Support a sub-rate Ethernet MAC over a single Ethernet PHY, e.g.

supporting a 50G MAC over a 100GBASE-R PHY

Hussain, et al. Expires August 17, 2019 [Page 3]

Internet-Draft FlexE Extensions February 2019

c. Support a collection of flexible Ethernet clients over a single

Ethernet PHY, e.g. supporting two MACs with the rates 25Gbps, and

one with rate 50G over a single 100GBASE-R PHY

d. Support a sub-rate Ethernet MAC over bonded PHYs, e.g. supporting

a 150G Ethernet client over 2 bonded 100GBASE-R PHY(s)

e. Support a collection of Ethernet MAC clients over bonded Ethernet

PHYs, e.g. supporting a 50G and 150G MAC over 2 bonded 100GBASE-R

PHY(s)

FlexE networks feature FlexE Ethernet interfaces, for more details

see Section 4.1.

From a control plane perspective, the FlexE Groups may be viewed as

logical links and FlexE Clients as logical sub-interfaces (or

channelized interfaces).

These logical point-to-point links may be realized in at least two

different ways:

a. direct point-to-point links with no intervening transport

network.

b. direct point-to-point links across a transport network transport

network.

c. Ethernet PHY(s) may be transparently transported via an Optical

Transport Network (OTN), as defined by ITU-T in [G.709] and

[G.798].

The OTN set of client mappings has been extended to support the

use cases identified in the OIF FlexE implementation agreement.

This document is a framework for the network control plane signaling

and routing extensions required to establish FlexE links (FlexE

Groups (PHY) and FlexE Clients (MAC)). FlexE Links may interconnect

customer edge devices (CE to CE), CE to provider edge devices (PE),

PE to PE, or devices at the edge to devices in the core (PE to P) or

devices in the core (P to P).

Any pair of neighbouring L2 and L3 device that support FlexE

interfaces may be interconnected P2P using a FlexE link (PHY and

MAC). Further a device that terminates a FlexE link MUST be able to

extract either the L2 or L3 payload and switch on the appropriate

level, i.e. Ethernet, MPLS or IP. It should be noted that any type

of switching is out of scope for the FlexE specification.

Hussain, et al. Expires August 17, 2019 [Page 4]

Internet-Draft FlexE Extensions February 2019

FlexE CE devices may typically be L3 routers or other devices that

use FlexE at layers 1 and 2 to provide point-to-point connectivity

between each other.

Thus this draft considers the cases in which the two peer FlexE

devices are:

o interconnecting two parts of a customer network (CE to CE).

o at the edge of the customer network (CE) and the close edge of the

provider network (PE to CE).

o opposite edges of the FlexE capable network (PE tom PE).

o at the edge of the FlexE network PE interconnected to a provider

device (PE to P).

o interconnecting two provider devices (P to P).

This list of deployment cases will help identify the GMPLS control

plane (i.e. routing and signaling) extensions that may be required

to support establishment of FlexE services.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",

"SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and

"OPTIONAL" in this document are to be interpreted as described in BCP

14 [RFC2119] [RFC8174] when, and only when, they appear in all

capitals, as shown here.

1.2. Updates in the version

This section will be removed before posting.

1. Following a suggestion from Daniele the FlexE Control Plane

Requirements has been moved to an appendix.

2. There are still some of the comments from Daniele that might need

to be addressed, but we have had a pretty large overlap in

comments, so the intention is that all should be addressed.

3. The terms Ethernet Interface and Ethernet sub-Interface has been

re-introduced in relation to FlexE Group and FlexE Client

respectively.

Hussain, et al. Expires August 17, 2019 [Page 5]

Internet-Draft FlexE Extensions February 2019

4. Except for some spelling corrections Section 5 to Section 7 are

virtually untouched, though it is likely that some of the changes

in the earlier parts of the document will have to be reflected

into those sections also.

2. Terminology

a. CE (Customer Edge): the group of functions that support the

termination/origination of data received from or sent to the

network. Sometimes the term CE device is used.

b. controller: a joint term for any entity that may set up a LSP,

FlexE Group or FlexE Client, e.g. a control plane, centralized

controller, YANG model or management system.

c. crunch: the term crunch in the context of OTN networks and FlexE

links is used when e.g. unavailable calendar slots are not

transported across the OTN network, but are removed at the

ingress and recreated at the egress.

d. Ethernet PHY: an entity representing Physical Coding Sublayer

(PCS), Physical Media Attachment (PMA), and Physical Media

Dependent (PMD) layers.

e. FlexE Calendar: The total capacity of a FlexE Group is

represented as a collection of slots which have a granularity of

5Gbps. The calendar for a FlexE Group composed of n 100G PHYs is

represented as an array of 20n slots (each representing 5Gbps of

bandwidth). This calendar is partitioned into sub-calendars,

with 20 slots per 100G PHY. Each FlexE Client is mapped into one

or more calendar slots (based on the bandwidth the FlexE Client

flow will need).

f. FlexE Channelized sub-Interface, the channelized Ethernet sub-

interface realized by the FlexE Client.

g. FlexE Client: An Ethernet flow based on a MAC data rate that may

or may not correspond to any Ethernet PHY rate.

h. FlexE Group: A FlexE Group is composed of from 1 to n Ethernet

PHYs. In the first version of FlexE each PHY is identified by a

number in the range {1-254}.

i. FlexE Interface, the Ethernet interface realized the FlexE Group

j. FlexE Shim: the layer that maps or demaps the FlexE Client flows

carried over a FlexE Group.

Hussain, et al. Expires August 17, 2019 [Page 6]

Internet-Draft FlexE Extensions February 2019

k. LMP: Link Management Protocol

l. LSP: Label Switched Path

m. OIF: Optical Internetworking Forum

n. OTN: Optical Transport Network

o. PE: Provider Edge (device) the term is used for the functions

needed at the edge of a provider network or the device to which

these functions are allocated.

p. P: Provider (device), the term is used for the functions needed

in the core of a provider network or the device to which these

functions are allocated.

q. SG15: ITU-T Study Group 15 (Transport, Access and Home).

r. TE: Traffic Engineering

s. TED: Traffic Engineering Database

3. FlexE Reference Model

The figure below gives a simplified FlexE reference model.

Hussain, et al. Expires August 17, 2019 [Page 7]

Internet-Draft FlexE Extensions February 2019

+-----+

| P |

...........+-+-+-+.............

n x PHY . n x | | m x .

. crunched | | crunched .

+----+ +-----+ PHYs | | PHYs +-----+ +----+

| CE +------+ PE1 +------------+ +--------------+ PE2 +----+ CE |

+----+ +-----+ +-----+ +----+

. .

+----+ p x PHY . . +----+

| CE +-----------------------------------------------------+ CE |

+----+ . . +----+

. OTN Network .

. .

...............................

+----+ q x PHY +----+

| CE +-----------------------------------------------------+ CE |

+----+ +----+

Legend: m, n, p and q indicate how many PHYs

there are in a FlexE Group

Figure 1: FlexE Reference Model

The services offered by Flexible Ethernet are essentially the same as

for traditional Ethernet, connectionless Ethernet transport. In

essence the FlexE interfaces and links may be viewed as any other

Ethernet interfaces or links. However, it is possible to capture

additional TE information in the Traffic Engineering Data Base

showing unique characteristics of FlexE channelized interfaces and

links. This makes it possible for the control plane to strategically

use FlexE networks to support advanced TE.

4. GMPLS Controlled FlexE

The high level goals for using a GMPLS control plane for FlexE can be

summarized as:

o Set up a FlexE Group

o Set up a FlexE Client

o Advertise the TE information of FlexE Groups and FlexE Clients

Hussain, et al. Expires August 17, 2019 [Page 8]

Internet-Draft FlexE Extensions February 2019

o Set up of a higher layer LSPs that require to be (or would have

significant benefits to) be run over a FlexE network infrastructure.

o Decoupling PHY and MAC bandwidth opens up some interesting

features for networks that features FlexE links. By establishing

several FlexE Clients with bandwidth that are part of the

bandwidth of the FlexE Group, it is possible to create channels

between two nodes.

o By controlling the mapping a user packets (or frames) to these

channels it is possible to create bandwidth that are dedicated for

special purposes, and that can't be infringed on by packets (or

frames) that does not satisfy this mapping.

4.1. Interfaces in a FlexE network

FlexE Ethernet interfaces are realized by the means of a basic

building block. The same building block is used for a single PHY and

when the PHYs are bonded. The building block consists of two FlexE

Shim functions (see Section 5.2.2.2) and a logical point to point

link. The FlexE Shim functions are located at each end of the

logical point to point link. This link carries the Ethernet PHY

signals between the two FlexE Shim Functions.

4.2. Mapping of traffic in the data plane

An example of which data plane mappings takes palace when an upper

layer, e.g. IP or MPLS, send packets over a FlexE interface is

shown in Figure 2.

Hussain, et al. Expires August 17, 2019 [Page 9]

Internet-Draft FlexE Extensions February 2019

IP packet IP Layer

||

\/

MPLS encapsulation MPLS

||

\/

FlexE Client ID FlexE

||

\/

calendar mapping FlexE

||

\/

FlexE Group FlexE

Figure 2: Traffic Mapping

In the mapping steps indicated in Figure 2 only one step in the

mapping is visible by each layer.

o the MPLS layer knows from the IP address, which MPLS label stack

to encapsulate the IP packet in

o the MPLS layer also know which MPLS label(s) that maps to which

FlexE Client

o the FlexE layer also knows from the FlexE Client Identifier, which

calendar slots the packet will be transferred over

o the FlexE layer knows which FlexE Group a certain set of calendar

slots belongs too

4.3. The GMPLS Control Plane and the FlexE identifiers

This section lists some of the procedures and actions on FlexE

Interface Identifiers that a GMPLS Control plane need to perform.

Also, a centralized controller, YANG model or a management system

that are used to establish interfaces and links need to perform the

same actions.

The FlexE Group Identifier and the FlexE Client Identifier, included

in the overhead of each frame sent over a FlexE Interface or sub-

Interface, indicates a particular Group or Client.

When the Control Plane, a centralized controller, a YANG model or a

management system sets up a FlexE Interface at least the bandwidth

Hussain, et al. Expires August 17, 2019 [Page 10]

Internet-Draft FlexE Extensions February 2019

has to be included in the setup message. The FlexE system wherein the interface belongs return the

FlexE Group Identifier in the response message.

When a channelized sub-interface is set up, the party initiates

the setup includes the Interface (FlexE Group) Identifier over which

the sub-Interface will be established, and the bandwidth requested

for the sub-interface. The FlexE system wherein the sub-interface belongs returns the FlexE Client

Identifier.

The identifiers received by the party that initiate a setup of an

FlexE Interfaces are used, by a controller, to set up FlexE sub-

interfaces.

The identifiers received by the party that initiate a setup of an

FlexE sub-Interfaces are used, e.g. to map an MPLS label to the

correct FlexE sub-interfaces.

4.4. Operational concerns

When operating a link in a FlexE network it is likely that an

operator would like to split the FlexE Interface into some sub-Interfaces

used for best effort traffic and some sub-Interfaces dedicated for

special purposes. An example would be when there is a 100 Gbit/s

FlexE are split in to five 10 Gbit/s sub-interfaces and one 50 Gbit/s

sub-interface. The 50 Gbit/s sub-interface could be used best effort

traffic, the five 10 Gbit/s could be used for dedicated traffic.

In such cases it is conceivable that packets/frames that have a

matching key will be put on a specific sub-Interface, while traffic

that do not have a matching key will be put on the best effort sub-

interface.

4.5. Pre-configured vs. Control Plane established LSPs in a FlexE

capable network

The FlexE infrastructure may be established in three different ways

o The FlexE Groups and FlexE Client may be pre-configured

o Only the FlexE Groups may be pre-configured, while the setup of

the FlexE Client is triggered by the request to setup a MPLS LSP.

o The setup of both FlexE Group and FlexE Client may be triggered by

the request to setup an MPLS LSP.

In the case the FlexE Groups and FlexE Clients are preconfigured the

FlexE capable nodes need to have the ability to announce the

preconfigured FlexE Client and/or FlexE Groups as if they were LSPs.

Hussain, et al. Expires August 17, 2019 [Page 11]

Internet-Draft FlexE Extensions February 2019

4.6. Signaling Channel

In the type of equipment for which FlexE was first specified an out

of band signaling channel is not commonly available. If that is the

case, and the GMPLS FlexE control plane will be used, the FlexE Group

will have to setup by e.g. a management system and a FlexE Client on

that FlexE Group (also configured) will have to allocated as a

signaling channel.

Further details of the setup of the FlexE Groups, FlexE Clients and

MPLS LSPs over a FlexE infrastructure will be found in Section 6.2.

4.7. MPLS LSP over the FlexE Data Plane

FlexE is a true link layer technology, i.e. it is not switched, this

means that the FlexE Groups and FlexE Clients are terminated on the

next-hop node, and that the switching needs to take place on a higher

layer.

The FlexE technology can be used to establish link layer connectivity

with high and deterministic bandwidth. However, there is no way

described in the FlexE specification to, in a deterministic way,

allocate certain traffic to a specific FlexE Client. Control of the

FlexE link layer by a GMPLS control plane can achieve this.

A GMPLS controlled FlexE capable node may be thought of using the

traditional model of a node with a separation between control and

data plane.

Hussain, et al. Expires August 17, 2019 [Page 12]

Internet-Draft FlexE Extensions February 2019

+------------------+

| +------------+ |

| | GMPLS | |

| | Control | |

| | Plane | |

| +------------+ |

| ^ |

| | |

| v |

| +------------+ |

| | FlexE | |

| | Data | | ^

| | Plane | |

| +------------+ |

+------------------+

Figure 3: GMPLS controlled FlexE Node

The GMPLS control plane will speak extended standard GMPLS protocols

with its neighbours and peers.

Node A Node B Node Z

+--CP

+-|-----------+ |------------------+ ~ +---------+

| | | | | | |

| | +------+ | | +--------------+ | | +-----+ |

LSP | +->| v | | | | ....x..... | | | | ^ | |

| | | . | | | | . . | | | | . | |

| | +--.---+ | | +--.--------.--+ | | +--.--+ |

FlexE | +->| o | | | | o | | o | | | | o | |

Client | | | o | | | | o | | o | | | | o | |

| | +--o---| | | +--o--+ +--o--+ | | +--o--| |

FlexE | +->| U | | | | U | | U | | | | U | |

Group | | U | | | | U | | U | | | | U | |

| +--U---| | | +--U--+ +--U--+ | | +--U--+ |

|-------U-----+ +----U--------U----+ +----U----+

U U U U

UUUUUUUUUUUUUUUUUUUUU UUUUUUU ~ UUUUUUUU

Legend ... = LSP

ooo = FlexE Client

UUU = FlexE Group

Figure 4: GMPLS controlled network with FlexE infrastructure

Hussain, et al. Expires August 17, 2019 [Page 13]

Internet-Draft FlexE Extensions February 2019

Figure 4 describes how an MPLS LSP is mapped over a FlexE Client and

FlexE Group.

4.8. Configuring the data plane in FlexE capable nodes

In Figure 4 we show an LSP, a FlexE Client and a FlexE Group, the LSP

is there because while the FlexE Channel and Group are not switched,

switching in our example takes place on the LSP level. This section

will discuss establishment of FlexE Clients and Groups, and mapping

of the LSP onto a FlexE Client.

The establishment of a LSP over a FlexE system is very similar to how

this is done in any other system. Building on information gathered

through the routing system and using the GMPLS signaling to establish

the LSP.

4.8.1. Configure/Establish a FlexE Group/Link

Consider the setup of a FlexE Group between node A and B,

corresponding to the row of U's from node A to B in Figure 4. The

FlexE Group is considered to consist of n PHYs, but does not have any

FlexE Clients defined from start.

When this is done by the GMPLS control plane, two conditions need to

be fulfilled (1) there need to be a data channel defined between node

A and B; and (2) a FlexE capable IGP-TE protocol needs to be running

in the network.

Node A will send an RSVP-TE message to node be with the information

describing the FlexE Group to be setup. This information might be

thought of as the "FlexE Group Label" (or part of the FlexE label).

It will contain at least the following information:

o A FlexE Group Identifier (FGid).

o The number of active FlexE Channels (numFC), where 0 indicates

that zero clients are active.

o Number of PHYs that the FlexE Group is composed of, for each PHY

\* PHY identifier

\* PHY bandwidth

\* slot granularity/number of slots

\* available and unavailable slots

Hussain, et al. Expires August 17, 2019 [Page 14]

Internet-Draft FlexE Extensions February 2019

When node B receives the RSVP-TE message it checks that it can setup

the requested FlexE Group. If the check turns positive, node send an

acknowledgment to node A and the FlexE Group is setup.

A more detailed description of how to setup a FlexE Group, will be

included in the draft dealing with signaling in detail.

4.8.2. Configure/Establish a FlexE Client

Consider the situation where a FlexE Group is already established (as

described in Section 4.8.1) and an m G FlexE Client is needed.

Similar to the establishment of the FlexE Group, node A will send a

RSV-TE message to node B.

This RSVP-TE message include at least the following information:

o FlexE Group Identifier

o FlexE Client Identifier

o from which PHYs the slots will allocated, i.e. slots might come

from more than one PHY.

o Information per PHY

\* PHY bandwidth

\* slot granularity

\* available/unavailable slots

\* allocated slots

A more detailed description of how to setup a FlexE Channel, will be

included in the draft dealing with signaling in detail.

4.8.3. Advertise FlexE Groups and FlexE Clients

Once the FlexE Group and FlexE Clients are configured they can be

advertised into the routing system as normal routing adjacencies,

including the FlexE specific TE information.

5. Framework and Architecture

This section discusses FlexE framework and architecture. Framework

is taken to mean how FlexE interoperates with other parts of the data

communication system. Architecture is taken to mean how functional

groups and elements within FlexE work together to deliver the

Hussain, et al. Expires August 17, 2019 [Page 15]

Internet-Draft FlexE Extensions February 2019

expected FlexE services. Framework is taken to mean how FlexE

interacts with it environment.

5.1. FlexE Framework

The service offered by Flexible Ethernet is a transport service very

similar (or even identical) to the service offered by Ethernet.

There are two major additions supported by FlexE:

o FlexE is intended to support high bandwidth and FlexE can offer

granular bandwidth from 5Gbits/s and a bandwidth as high as the

FlexE Group allows.

o As FlexE Groups and clients are setup as a configuration activity,

by a centralized controller or by a GMPLS control plane the

service is connection oriented.

5.2. FlexE Architecture

5.2.1. Architecture Components

This section discusses the different parts of FlexE signaling and

routing and how these parts interoperate.

The FlexE routing mechanism is used to provide resource available

information for setup of higher layer LSPs, like Ethernet PHYs'

information, partial-rate support information. Based on the resource

available information advertised by routing protocol, an end-to-end

FlexE connection is computed, and then the signaling protocol is used

to set up the end-to-end connection.

FlexE signaling mechanism is used to setup LSPs.

MPLS forwarding over a FlexE infrastructure is different from

forwarding over other infrastructures. When MPLS runs over a FlexE

infrastructure it is possible that there are more than FlexE Client

that meet the next-hop requirements, often it is possible to use any

suitable FlexE Clientfor a hop between two nodes. If the mapping

between a MPLS encapsulated packet and the FlexE Client, this mapping

need to be explicit when the LSP is set up, and the MPLS label will

be used to find the correct FlexE Client.

5.2.2. FlexE Layer Model

The FLexE layer model is similar Ethernet model, the Ethernet PHY

layer corresponds to the "FlexE Group", and the MAC layer corresponds

to the "FlexE Client".

Hussain, et al. Expires August 17, 2019 [Page 16]

Internet-Draft FlexE Extensions February 2019

As different from earlier Ethernet the combination of Flexe Group and

Client allows for a huge freedom when it comes to define the

bandwidth of an Ethernet connectivity.

5.2.2.1. FlexE Group structure

The FlexE Group might be supported by virtually any transport

network, including the Ethernet PHY. While the Ethernet PHY offers a

fixed bandwidth the FlexE Group has been structured into 5 Gbit/s

slots. This means that the FlexE Group can support FlexE Clients of

a variety of bandwidths.

The first version is defined for 20 slots of 5 Git/s over a 100 Gbit/

s PHY. The 100 Gbit/s PHYs can be bonded to give higher bandwidth.

5.2.2.2. FlexE Client mapping

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 Shim is

the layer that maps or demaps the FlexE Client flows carried over a

FlexE Group. As defined in [OIFFLEXE1], MAC rates of 10, 40, and any

multiple of 25 Gbit/s are supported. This means that if there is a

100 Gbit/s FlexE Group between A and B, a FlexE Client of 10, 25, 40,

50, 75 and 100 Gbit/s can be created.

However, by bonding, for example 5 PHYs of 100 Git/s to a single

FlexE Group, FlexE Clients of 500 Gbit/s can be supported.

6. Control Plane

This section discusses the procedures and extensions needed to the

GMPLS Control Plane to establish FlexE LSPs.

There are several ways to establish FlexE Groups, allocate slots for

FlexE Clients, and setup higher layer LSPs. A configuration tool, a

centralized controller or the GMPLS control plane can all be used.

To create the FlexE GMPLS control plane Groups, FlexE Clients and

higher layer LSPs, extensions to the following protocols may be

needed:

o "RSVP-TE: Extensions to RSVP for LSP Tunnels" (RSVP-TE) [RFC3209]

o "Link Management Protocol" (LMP) [RFC4204]

o "Path Computation Element (PCE) Communication Protocol" (PCEP)

[RFC5440]

Hussain, et al. Expires August 17, 2019 [Page 17]

Internet-Draft FlexE Extensions February 2019

o IS-IS Extensions for Traffic Engineering (ISIS-TE) [RFC5305]

o "OSPF Extensions in Support of Generalized Multi-Protocol Label

Switching (GMPLS)" (OSPF-TE) [RFC4203]

o "North-Bound Distribution of Link-State and Traffic Engineering

(TE) Information Using BGP" (BGP-LS) [RFC7752]

A FlexE control plane YANG model will also be needed.

Section 6.2 and Section 6.1 discusses the role of the GMPLS control

plane when primarily setting up LSPs.

When discussing the signaling and routing procedures we assume that

the FlexE Group has been established prior to executing the

procedures needed to establish an LSP. Technically it is possible to

establish FlexE Group, allocate FlexE Client slots and LSP with a

single exchange of GMPLS signaling messages.

6.1. GMPLS Routing

To establish an LSP the Traffic Engineering (TE) information is the

most critical information, e.g. resource utilization on interfaces

and link, including the availability of slots on the FlexE Groups.

The GPMPLS routing protocols needs to be extended to handle this

information. The Traffic Engineering Database (TED) will keep an

updated version of this information.

The FlexE capable nodes will be identified by IP-addresses, and the

routing and traffic engineering information will be flooded to all

nodes within the routing domain using TCP/IP.

When an LSP over the FlexE infrastructure is about to be setup, e.g.

R1 - R4 - R5 in Figure 5 the information in the TED is used verify

that resources are available. When it is conformed that the LSP is

established the TED is updated, marking the resources used for the

new LSP as used. Similarly, when a LSP is taken down the resources

are marked as free.

6.2. GMPLS Signaling

As described in Section 4 the state of the FlexE infrastructure may

effect the actions needed to setup an LSP in a FlexE capable network.

The FlexE infrastructure maybe be:

1. fully pre-configured

Hussain, et al. Expires August 17, 2019 [Page 18]

Internet-Draft FlexE Extensions February 2019

2. partially pre-configured, i.e. the FlexE Group may be pre-

configured, but not the FlexE Clients

3. not pre-configured, i.e. the setup of FlexE Group and FlexE

Client will be triggered because of the request to setup an LSP.

Figure 5 will be used to illustrate the different cases.

+----+

| R1 +---------------------+

+----+ |

|

+----+ +--+--+ +----+

| R2 +------------------+ R4 +-------------------------+ R5 |

+----+ +--+--+ +----+

|

+----+ |

| R3 +---------------------+ PHY R1 to R4 100 Gbit(s

+----+ PHY R2 to R4 100 Gbit(s

PHY R3 to R4 100 Gbit(s

PHY R4 to R5 200 Gbit(s

Figure 5: FlexE LSP Example

The text in Section 6.2 is not a specification of the GMPLS signaling

extensions for FlexE capable network, it is a description to

illustrate the expected features of such a protocol. Nor do we

discuss failure scenarios.

6.2.1. LSP setup with pre-configured FlexE infrastructure

In this first example, referencing Figure 5, one 100 Gbit/s FlexE

Group is configured between R1 and R4, between R2 and R4, and between

R3 and R4. Between R4 and R5 there is a 200 Gbit/s FlexE Group.

Over each 100 Gbit/s FlexE Group there are four 5 Gbit/s, two 20

Gbit/s and one 40 Gbit/s FlexE Clients configured. Over the 200 Git/

s FlexE Group there are eight 5 Gbit/s, four 20 Gbit/s and two 40

Gbit/s FlexE Clients configured.

One of the 5 Gbit/s FlexE Clients on each FlexE Groups are used as

signaling channel.

Hussain, et al. Expires August 17, 2019 [Page 19]

Internet-Draft FlexE Extensions February 2019

To establish the for example a 200 Mbit/s MPLS LSP the normal GMPLS

request/response procedures are followed. R1 sends the request to

R4, R4 allocate resources on one of the FlexE Clients, forward the

request to R5. R5 responds to R4 indicating the label and the FlexE

Client the traffic should be sent over, R4 does the same for R1.

The only difference between the standard signaling and what happens

here is that there the assigned label will be used to find the right

FlexE Client.

6.2.2. LSP setup with partially configured FlexE infrastructure

In the second example, also referencing Figure 5, the FlexE Groups

are setup in the same way as in the first example, however only one 5

Gbit/s FlexE Client per FlexE Group are established by configuration.

This FlexE Client will be used for signaling.

When preparing to send the request that a 5 Gbit/s MPLS LSP shall be

set up R1 discovers that there are no feasible FlexE Client between

R1 aand R4. R1 therefore sends the request to establish such a FlexE

Client, when receiving the request R4 allocates resources for the

FlexE Client on the FlexE Group. There may be different strategies

for allocating the bandwidth for this FlexE Client. Such strategies

are out of scope for this document. R1 then sends the information

about the FlexE Client to R1, and both ends establish the FlexE

Client.

When the FlexE Client between R1 and R4 is established, R1 proceeds

to send the request for an MPLS LSP to R4. R4 will discover that a

feasible FlexE Client is missing between R4 and R5. The same

procedure s for setting up the FlexE Client between R1 and R4 is

repeated for R4 and R5. When there is a feasible FlexE Client

available the signaling to set up the MPLS LSP continues as normal.

The label allocated for the MPLS LSP will be used to find the correct

FlexE Client.

When a FlexE Clients is set up in this way they can be announced into

the routing system in two different ways. First, they can be made

generally available, i.e. it will be free to use for anyone that want

to set up LSPs over the FlexE Group between R1 and R4 and between R4

and R5. Second, the use of the FlexE Clients may be restricted to

the application that initially did set up the FlexE Client.

Hussain, et al. Expires August 17, 2019 [Page 20]

Internet-Draft FlexE Extensions February 2019

6.2.3. LSP setup with non-configured FlexE infrastructure

This example also refers to Figure 5 as different from the earlier

example no FlexE Group or FlexE Client configuration is done prior to

the first request for an MPLS LSP over the FlexE infrastructure.

To make the set up of LSPs in a FlexE network where no FlexE Groups

or FlexE Clients have been configured two conditions need to be

fulfilled. First an out of band signaling channel must be available.

Second the FlexE Capabilities must be announced in to the IGP and/or

centralized controller.

If these two conditions are fulfilled, the set up of an MPLS LSP

progress pretty much as in the partially configured network. The

difference is that the set up of both the FlexE Group and FlexE

Client are triggered by the request to set up an MPLS LSP.

As in the partially configured case FlexE Clients can be announced

into the routing system in two different modes, either they are

generally available. It or they are reserved for the applications

that first established them.

6.2.4. Packet Label Switching Data Plane

This section discusses how the FlexE LSP data plane works. In

general it can be said that the interface offered by the FlexE Shim

and the FlexE Client is equivalent to the interface offered by the

Ethernet MAC.

Figure 6 below illustrates the FlexE packet switching data plane

procedures.

Hussain, et al. Expires August 17, 2019 [Page 21]

Internet-Draft FlexE Extensions February 2019

R1 R3 R4

............. ...................... ...........

. +-------+ . . +----------------+ . . +-----+ .

. | LSP | . . | LSP \ / LSP | . . | LSP | .

. | a | . . | a \/ b | . . | b | .

. +-------+ . . +----------------+ . . +-----+ .

. | ETH | . . | ETH | | ETH | . . | ETH | .

. | i/f | . . | i/f | | i/f | . . | i/f | .

. +-------+ . . +-----+ +-----+ . . +-----+ .

. | FlexE | . . |FlexE| |FlexE| . . |FlexE| .

. | trsp | . . |trsp | |trsp | . . |trsp | .

. +---+---+ . . +--+--+ +--+--+ . . +--+--+ .

......|...... .....^..........|..... .....^.....

| | | |

+--------------------+ +--------------------+

Figure 6: LSP over FlexE Data Plane

The data plane processes packets like this:

o The LSP encapsulating and forwarding function in node R1 receives

a packet that needs to be encapsulated as an MPLS packet with the

label "a". The label "a" is used to figure out which FlexE

emulated Ethernet interfaces the label encapsulated packet need to

be forwarded over.

o The Ethernet interfaces, by means of FlexE transport, forwards the

packet to node R3. Node R3 swaps the label "a" to label "b" and

uses "b" to decide over which interface to send the packet.

o Node R3 forwards the packet to node R, which terminates the LSP.

Sending MPLS encapsulated packets over a FlexE Client is similar to

send them over an Ethernet 802.1 interface. The critical differences

are:

o FlexE channelized sub-interfaces guarantee a deterministic

bandwidth for an LSP.

o When a application that originally establish a FlexE Client

reserve it for use by that application only, it is possible to

create uninfringeable bandwidth end-to-end for an MPLS LSP.

o FlexE infrastructure allows for creating very large end to end

bandwidth

Hussain, et al. Expires August 17, 2019 [Page 22]

Internet-Draft FlexE Extensions February 2019

7. Operations, Administration, and Maintenance (OAM)

To be added in a later version.

8. Acknowledgements

9. IANA Considerations

This memo includes no request to IANA.

Note to the RFC Editor: This section should be removed before

publishing.

10. Security Considerations

To be added in a later version.

11. Contributors

Khuzema Pithewan, Infinera Corp, kpithewan@infinera.com

Fatai Zhang, Huawei, zhangfatai@huawei.com

Jie Dong, Huawei, jie.dong@huawei.com

Zongpeng Du, Huawei, duzongpeng@huawei.com

Xian Zhang, Huawei, zhang.xian@huawei.com

James Huang, Huawei, james.huang@huawei.com

Qiwen Zhong, Huawei, zhongqiwen@huawei.com

Yongqing Zhu China Telecom zhuyq@gsta.com

Huanan Chen China Telecom chenhuanan@gsta.com

12. References

12.1. Normative References

[G.709] ITU, "Optical Transport Network Interfaces

(http://www.itu.int/rec/T-REC-G.709-201606-P/en)", July

2016.

Hussain, et al. Expires August 17, 2019 [Page 23]

Internet-Draft FlexE Extensions February 2019

[G.798] ITU, "Characteristics of optical transport network

hierarchy equipment functional blocks

(http://www.itu.int/rec/T-REC-G.798-201212-I/en)",

February 2014.

[G.8021] ITU, "Characteristics of Ethernet transport network

equipment functional blocks", November 2016.

[G.872] ITU, "Architecture of optical transport networks", January

2017.

[OIFFLEXE1]

OIF, "FLex Ethernet Implementation Agreement Version 1.0

(OIF-FLEXE-01.0)", March 2016.

[OIFFLEXE2]

OIF, "FLex Ethernet Implementation Agreement Version 2.0

(OIF-FLEXE-02.0)", June 2018.

[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>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC

2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,

May 2017, <https://www.rfc-editor.org/info/rfc8174>.

12.2. Informative References

[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,

<https://www.rfc-editor.org/info/rfc3209>.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in

Support of Generalized Multi-Protocol Label Switching

(GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005,

<https://www.rfc-editor.org/info/rfc4203>.

[RFC4204] Lang, J., Ed., "Link Management Protocol (LMP)", RFC 4204,

DOI 10.17487/RFC4204, October 2005,

<https://www.rfc-editor.org/info/rfc4204>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic

Engineering", RFC 5305, DOI 10.17487/RFC5305, October

2008, <https://www.rfc-editor.org/info/rfc5305>.

Hussain, et al. Expires August 17, 2019 [Page 24]

Internet-Draft FlexE Extensions February 2019

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation

Element (PCE) Communication Protocol (PCEP)", RFC 5440,

DOI 10.17487/RFC5440, March 2009,

<https://www.rfc-editor.org/info/rfc5440>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and

S. Ray, "North-Bound Distribution of Link-State and

Traffic Engineering (TE) Information Using BGP", RFC 7752,

DOI 10.17487/RFC7752, March 2016,

<https://www.rfc-editor.org/info/rfc7752>.

Appendix A. Requirements

This section summarizes the signaling and routing requirements for a

FlexE control plane, with respect to establishing FlexE Groups, FlexE

Clients and MPLS LSPs that require support from an FlexE

infrastructure.

Req-1 The FlexE control plane SHALL support the creation of FlexE

Groups.

\* A FlexE Groups consist one or more 100GE Ethernet PHY(s).

In the first version of FlexE the number of PHYs are in the

range of 1 to 254.

\* This requirement can be met by several methods, e.g.

routing and signaling protocols, a centralized controller

or a management system.

Any such method need to have network access to the FlexE

shims at each of the Ethernet PHY(s) termination points.

Req-2 The FlexE control plane SHALL have the ability to delete a

FlexE Group.

Req-3 The FlexE control plane SHALL have the ability to initiate an

administratively lock or unlock of a FlexE Group.

\* This ability is needed e.g. for executing the next

requirement.

Req-4 When a FlexE Group has been administratively looked is SHALL

be possible to add PHYs to an operational FlexE Group.

Req-5 When a FlexE Group has been administratively looked is SHALL

be possible to remove PHYs from an operational FlexE Group.

Hussain, et al. Expires August 17, 2019 [Page 25]

Internet-Draft FlexE Extensions February 2019

Req-6 The FlexE control plane SHALL support the ability to collect,

advertise and discover information about FlexE capable nodes,

including the TE information the FlexE Groups and FlexE

Clients the nodes support.

Note: In essence correct, but something is backward. Need to

think.

Req-7 The FlexE control plane SHALL allow the addition (or removal)

of one or more FlexE clients to a FlexE Group. The addition

(or removal) of a FlexE Client flow SHALL NOT affect the

services of the other FlexE Client signals.

Req-8 The FlexE control plane SHALL, though this MAY not be possible

in all network scenarios, support FlexE Client flow resizing

without affecting any existing FlexE Clients within the same

FlexE Group.

Req-9 The FlexE control plane SHALL support establishment of MPLS

LSPs that requires the support of a FlexE infrastructure.

Authors' Addresses

Iftekhar Hussain

Infinera Corp

169 Java Drive

Sunnyvale, CA 94089

USA

Email: IHussain@infinera.com

Radha Valiveti

Infinera Corp

169 Java Drive

Sunnyvale, CA 94089

USA

Email: rvaliveti@infinera.com

Qilei Wang

ZTE

Nanjing

CN

Email: wang.qilei@zte.com.cn

Hussain, et al. Expires August 17, 2019 [Page 26]

Internet-Draft FlexE Extensions February 2019

Loa Andersson

Huawei

Stockholm

Sweden

Email: loa@pi.nu

Mach Chen

Huawei

CN

Email: mach.chen@huawei.com

Haomian Zheng

Huawei

CN

Email: zhenghaomian@huawei.com

Hussain, et al. Expires August 17, 2019 [Page 27]