OIF FlexE White Paper

Bonding, Sub-Rating and Channelization for Ethernet

Abstract:

The OIF FlexE Interoperability Agreement enables new ways to flexibly use network bandwidth by supporting bonding, sub-rating and channelization of links. This white paper describes the key features and properties of FlexE.

About the OIF:

The OIF facilitates the development and deployment of interoperable networking solutions and services. Members collaborate to drive Implementation Agreements (IAs) and interoperability demonstrations to accelerate and maximize market adoption of advanced internetworking technologies. OIF work applies to optical and electrical interconnects, optical component and network processing technologies, and to network control and operations including software defined networks and network function virtualization. The OIF actively supports and extends the work of national and international standards bodies. Launched in 1998, the OIF is the only industry group uniting representatives from across the spectrum of networking, including many of the world's leading service providers, system vendors, component manufacturers, software and testing vendors. Information on the OIF can be found at http://www.oiforum.com.

For additional information contact:

OIF, 5177 Brandin Ct

Fremont, CA 94538 USA

510-492-4040 info@oiforum.com www.oiforum.com

Table of Contents

Table of Figures

Introduction

The OIF published the FlexE Interoperability Agreement (IA) in 2016. FlexE enables a new generation of equipment to support many new Ethernet connection types, allowing operators to utilize bandwidth in flexible and fine-tuned ways.

The **key features** that the FlexE IA supports are:

- **Bonding of Multiple Links** – this allows an operator to create a larger link out of multiple slower links.
- **Sub-rating of Links** – this allows an operator to only use a portion of a link.
- **Channelization of Links** – this allows one link to carry several lower-speed or sub-rated links from different sources.

The **key properties** that FlexE has are:

- FlexE is backwards compatible with existing physical layer solutions. A FlexE compatible interface can leverage standards-based optical modules, electrical cables and backplanes.
- FlexE reuses many mechanisms from Ethernet. Much of the new functionality is achieved by adding a calendar that interacts with the existing Ethernet 64b66b mechanism, allowing bandwidth to be allocated with 5 Gb/s granularity.
- FlexE is defined to make use of standards-defined physical lanes.
- FlexE is backwards compatible with the existing transport infrastructure. A FlexE compatible interface can be connected to a piece of transport gear that is not aware of FlexE.
- The optional use of FlexE-aware transport equipment provides additional functionality such as matching client and line rates.
- FlexE can utilize the entire aggregated link, creating a compelling an alternative to traditional Link Aggregation (LAG) solutions, which use 70-80% of a link. FlexE has deterministic performance, whereas LAG does not.
- FlexE has low added latency as compared to Ethernet.
- FlexE has a set of features to support its use in transport networks.
- The OIF task group that developed FlexE has active liaison relationships with the IEEE.802.3 and the ITU.

FlexE Key Features

Bonding of Multiple Links

FlexE supports bonding of multiple links into a higher speed link. It does so in a way that has good properties as compared to previous methods, as detailed in a later section.

A specific example of the use of this feature is shown in the figure below. Equipment A and B support a 500 Gbps FlexE connection by bonding together five 100 GE links.

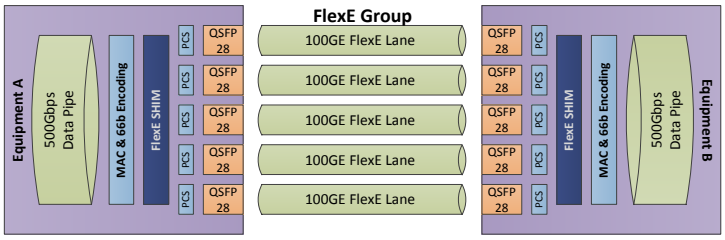

Figure 1 Bonding of links [Image Xilinx 2]

Sub-rating of Links

FlexE supports the use of just a portion of a link. There are many scenarios where the ideal bandwidth of a link that is purchased is lower or different than the rate of the physical interface, particularly when some form of channelized transport infrastructure is being connected to. FlexE's sub-rating mechanism allows this to be used.

An example of this is shown in the figure below. A 300 Gb/s link is desired by the user, but the coherent optical interface needs four 75 Gbps inputs in order to send at 150 Gbps per wavelength over two wavelengths over the desired reach. FlexE allows four 100 GE lanes to carry only 75 Gbps each.

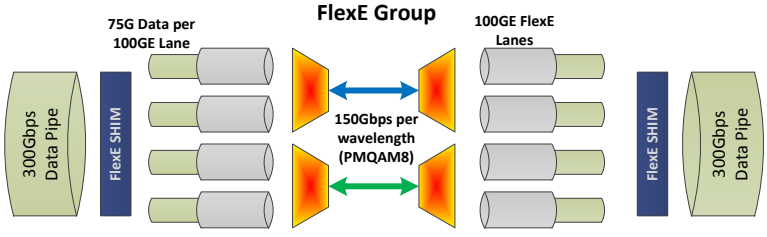

Figure 2 Sub-rating Links [Image Xilinx 2]

A similar use of FlexE would be to just use 75 Gbps of bandwidth out of one 100GE link if that were all that was needed in the application.

Channelization of Links

An important feature of FlexE is the ability to allow FlexE aware transport equipment to carry client traffic in a channelized manner. This feature can be thought of as the transport side of the sub-rating story.

In the figure below, five separate clients share a 500 Gb/s FlexE link. Those five different customers are running at 100 Gb/s (blue), 50 Gb/s (yellow), 25 Gb/s (purple), 200 Gb/s (orange) and 150 Gb/s (green).

FlexE brings some of the functionality of the Optical Transport Network (OTN) into the Ethernet world. FlexE allows a service provider to offer deterministic, Ethernet-oriented pipes of flexible width to their clients.

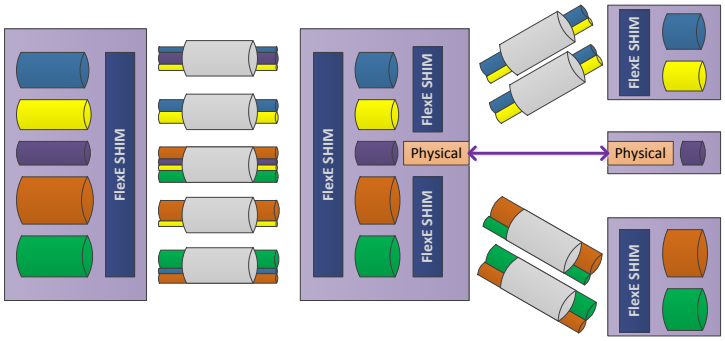

Figure 3 Channelized Links [Image Xilinx 2]

FlexE Key Properties

Backwards compatible with existing physical layer solutions

FlexE is backwards compatible with existing physical layer solutions. A FlexE compatible interface can use existing standards-based optical modules, electrical cables and backplanes.

Reuses Mechanisms from Ethernet

FlexE reuses many mechanisms from Ethernet. FlexE adds a calendar mechanism that interacts with Ethernet's existing 64b66b mechanism. Ethernet itself has to cope with a similar bonding situation because it must bond multiple lanes together to form an aggregated link. For example, 100GE is typically created from four 25 Gb/s

lanes, bonded together by the PCS layer. FlexE adds an additional calendar mechanism that allows bandwidth to be allocated on a 5 Gb/s basis.

Uses standards-defined physical lanes

FlexE is defined to make use of standards-defined lanes of 25 Gb/s lanes. This leverages the work of other standards development organizations. It also supports reuse of major portions of the many implementations of those standards.

Backwards compatible to Transport Infrastructure

FlexE is backwards compatible with existing transport infrastructure. A FlexE compatible interface can be connected to a piece of transport gear that is not aware of FlexE.

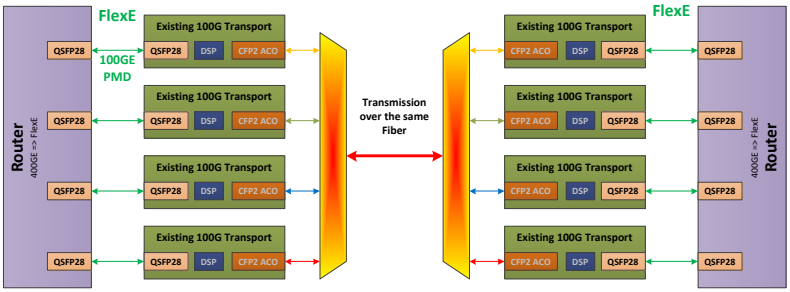

Figure 4 Backwards Compatible to Transport [Image Xilinx 2]

Options for Flex-Aware Transport

Optional FlexE-Aware transport equipment can provide additional functionality such as matching client and line rates.

In the following example, the next generation transponder uses another OIF specification, the CFP2 Analog Coherent interface to Optics (ACO) to an optical module, to allow it to run at a line rate that is changeable by the system. The Digital Signal Processor (DSP) for the link is on the front card, and the optical module has an analog interface instead of a digital interface. The FlexE block in the transponder can then reshuffle the FlexE mapping to correspond to the line rate of the optics link.

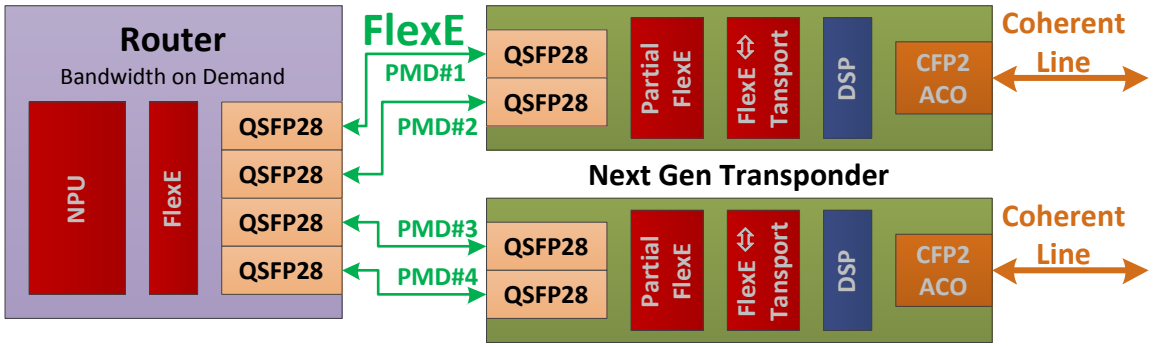

Figure 5 Options for Flex-Aware Transport [Image Xilinx 2]

Can Utilize Full bandwidth of Aggregated Link

FlexE can utilize the entire aggregated link. It forms a good alternative to traditional Link Aggregation (LAG) solutions, which can only use 70-80% of a link under typical conditions.

LAG uses a hashing algorithm to place traffic from a given source-destination pair onto just one of the links that are being aggregated together. Because of random nature of much Internet traffic, the right mix of traffic often does not show up to fill all of the links to capacity at any given moment.

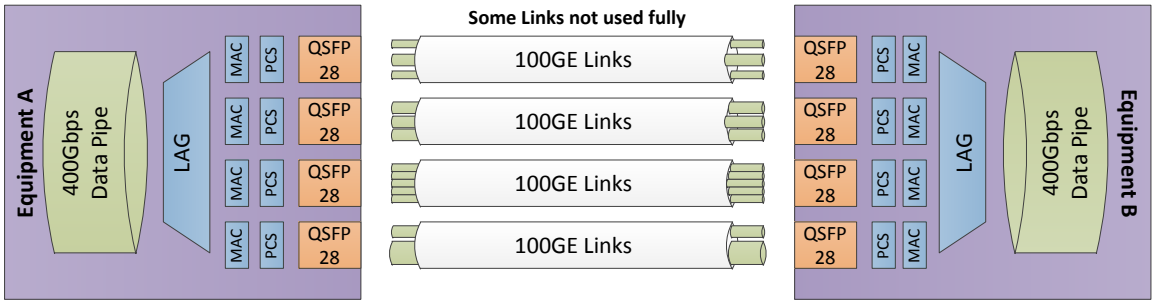

Figure 6 Aggregation via LAG [Image Xilinx 2]

FlexE uses a calendar-based Time Division Multiplexing (TDM) technique similar to the way that Ethernet's own PCS layer spreads traffic over the physical layer channels of a given link to spread the traffic over the channels. FlexE produces a deterministic bandwidth and thus can use the link fully.

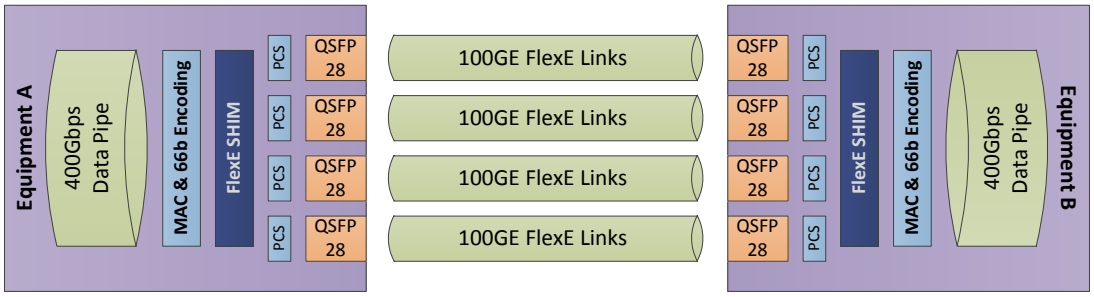

Figure 7 Aggregation via FlexE [Image Xilinx 2]

Low Added Latency

FlexE has low added latency as compared to what Ethernet normally has because it extends the mechanisms that are already present in Ethernet's Protocol Coherence Sublayer.

Transport Friendly

FlexE has set of features that allow its use in transport networks. These features include redundant calendars, continuous communication of the calendars, and an overhead communications channel..

Shared Knowledge

The OIF task group that developed FlexE has active liaison relationships with both the IEEE.802.3 and the ITU. Some individuals attend more than one of these bodies. These social bonds help ensure FlexE's bright future.

Summary

The OIF FlexE Interoperability Agreement enables the next generation of equipment to support many new Ethernet connection types, allowing operators to utilize bandwidth in flexible, dynamic and fine-tuned ways. Supported features include bonding of multiple links to form a higher speed link, sub-rating a link to allow only a portion of a link to be used, and channelization of links to allow one link to carry several links.

FlexE is particularly optimized to allow cloud-scale data center users to make the best use of bandwidth that they purchase from service providers or install on their own.

References

[1] Faisal Dada, Xilinx: "Implementing End to End Flexible Optical Transport with FPGAs and Transport Modems Utilizing FlexE, Optical Connections, Issue 6, Q1 2016, FTTH Conference Edition

[2] Xilinx, FlexE Introduction & Applications, Powerpoint file, provided by Faisal Dada