Flex Ethernet: breaking the chains of physical bandwidth

app note

EXFO

Flex Ethernet: breaking the chains of physical bandwidth

app note

By Anabel Alarcon
Product Specialist
EXFO

Today, Ethernet is one of the fundamental technologies that helps sustain the transport of exponential quantities of information all over the world. This information is transported through different types of physical links (copper, fiber) and even through the air with wireless and Bluetooth® technologies allowing millions of users every day to interconnect without boundaries.

This seems so far from 1973 when Bob Metcalfe, co-inventor of Ethernet, wrote a memo describing Ethernet as only a system interconnecting a few computers inside a specific building to form a network. He was inspired by the poetic concept of "Ether" that was considered as the media that helped propagate electromagnetic waves.

Over the years, Ethernet has evolved without rest and its transport capacity has increased dramatically. With Ethernet's bandwidth capacity increasing over forty thousand times (Figure 1) since inception, it's a technology that has lead the evolution in the way we exchange information.

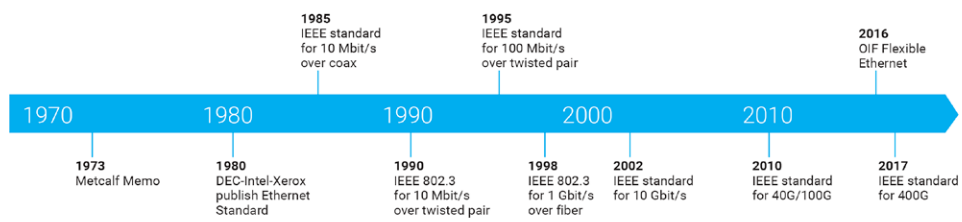


Figure 1. Introduction and evolution of Ethernet over time

In a very short time, designers, manufacturers and users had a big range of interface possibilities to choose from to interconnect network elements.

Every jump going forward has been associated with a different transport media, starting with coax cable, then twisted pair and finally the most popular today, fiber. In a very short period of time designers, manufacturers and users had a big range of interface possibilities to choose from to interconnect network elements.

A very important element that influenced the increase in Ethernet's bandwidth capacity was the development of the internet and the evolution of information technologies—generating the local and global expansion of networks to cover user needs. Ethernet networks grew from buildings, to cities, countries to across continents and they continue connecting the world. Service providers, hyperscale companies and data centers work every day to deploy new links that help them to move more and more information.

The challenges that interconnections face today include:

- Different types of physical interface (transceiver) technologies available on the market such as CFP4, QSFP28, and even advanced ones like CFP8, make it difficult for architects and designers to adapt to current and future requirements
- Mature information rates being widely deployed (e.g., 100G) and advanced new rates becoming standardized (Figure 2)
- Generation of new technology ecosystems
- Maximizing the utilization of existing infrastructure to fully profit from current investment before spending capital on new technology infrastructures
- Combination of transport technologies mixing Ethernet and coherent optical transmission techniques

FlexE breaks the chains imposed by physical interfaces, providing a solution to the challenges that hyperscale poses to the telecommunications and data center industries.

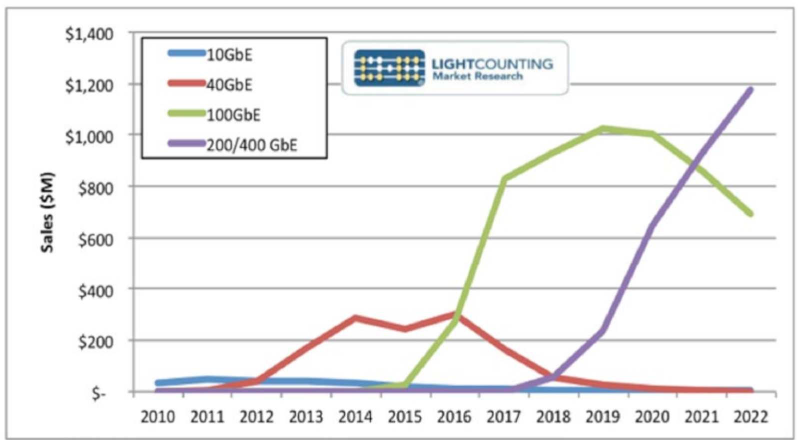


Figure 2. Market share across transportation rates

(Source: High-Speed Ethernet Optics, September 2017, LightCounting Market Research)

In the eye of this technology storm comes FlexE (Flex Ethernet). It breaks the chains imposed by physical interfaces, providing a solution to the challenges that hyperscale pose to the telecommunication and data center industries.

FlexE is an implementation agreement by the OIF (Optical Internetworking Forum). Developed in 2016, it describes the mechanism to transport a range of different Ethernet MAC rates not based on any current Ethernet PHY rate, breaking the constrains of transporting traffic limited to specific interface capacities.

An objective of this new implementation agreement is to improve and maximize the interconnection between network elements (routers) and transport gear. A second objective is to help network elements reach the bandwidth being currently handled today by transport elements that use coherent technologies.

Let's look at the elements that allow this flexible physical mapping structure to work:

FlexE client: A FlexE client is defined as an Ethernet data stream associated to a rate that could (or not) belong to an existing Ethernet PHY rate. (see Figure 3)

FlexE clients defined by the agreement include Ethernet rates already standardized (10G, 40G, 100G) as well as new rates (e.g., 200G and 400G) and finally multiples of 25G(n x 25G) such as 125 Gbit/s etc. Each FlexE client has a separate MAC address.

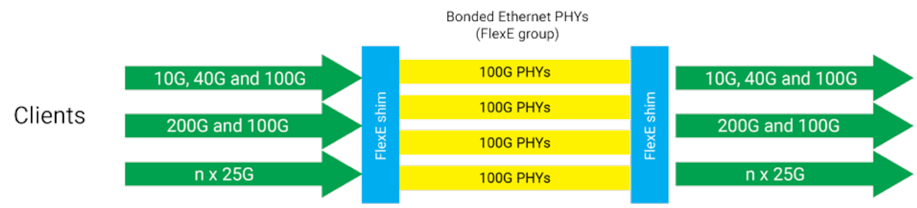


Figure 3. FlexE clients

FlexE group: The OIF defines the FlexE group as the "n" Ethernet PHYs bonded on the link where each group should have at least one Ethernet PHY. The first version of the implementation agreement defined 100G PHYs, but upcoming versions may include additional PHYs based on new standard rates.

The PHYs belonging to the FlexE group must have a number that identifies them within the group and also must be interconnected by the same two FlexE shims.

FlexE shim: Defined as the layer located between the MAC and the PCS inside the IEEE 802.3 stack (see Figure 4), FlexE shim is responsible for associating the different clients transported over a FlexE group.

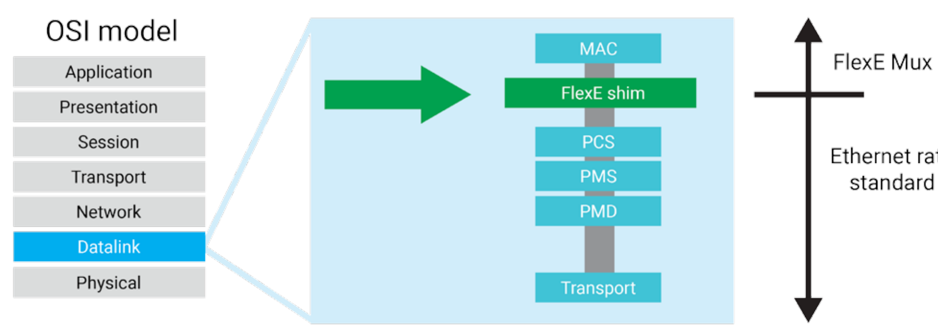
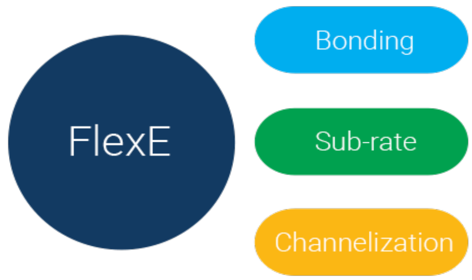


Figure 4. Characteristics of FlexE shim

This configuration allows multiplexing to occur at the MAC level leaving the rest of the layers to continue to be used as defined by the specific rate standard.

The clients are distributed sequentially into the calendar inside the FlexE shim and then distributed to each physical PHY PCS lane where each PHY is transported independently over the network. The FlexE shim is responsible for calendar management, synchronization and PHY mapping.

As part of the implementation agreement, the illustration below shows the main actions that a network element can do with FlexE:



Bonding: Bonding creates high capacity data pipes by assigning several Ethernet PHYs to transport MAC rates greater than 100G. (see Figure 5)

This is a very good alternative to layer 3 aggregation technologies like LAG, and ECMP that are being used to aggregate 100G links, but use former hashing algorithms that can't completely saturate the link and are difficult to control.

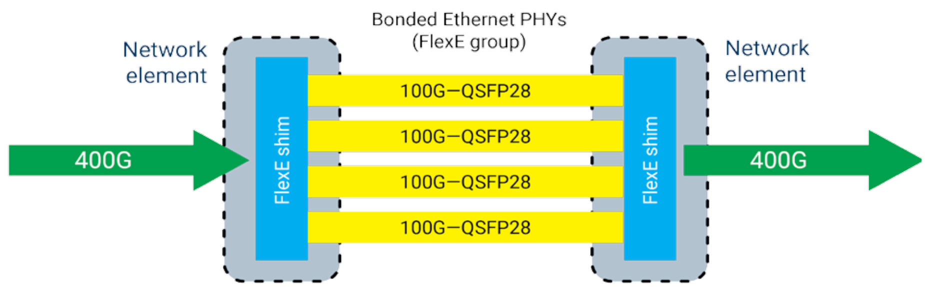


Figure 5. 400G bonding

Sub-rating: An action that allows network elements to sub-divide physical interfaces in order to transport lower data pipes over partially filled Ethernet PHYs. (see Figure 6)

Some applications of sub-rating are attempting to mimic transmission technologies such as coherent optical transmission. The purpose is to transport specific rates per wavelength that are not Ethernet standardized.

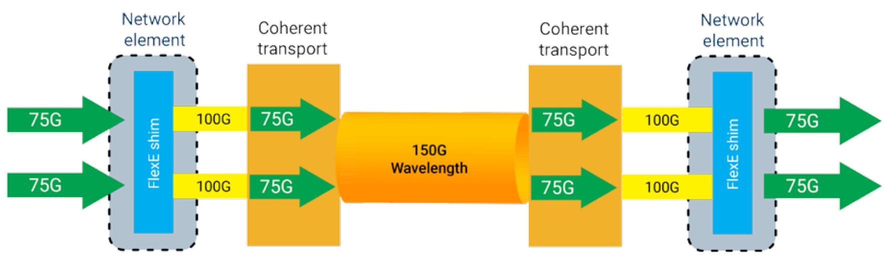


Figure 6. Sub-rating interop with coherent technologies

Channelization: Providing network elements with the capability to create specific transport channels for multiple data pipes though associated Ethernet PHYs, on the same or even different directions is called channelization. (see Figure 7)

Potentially this could be a layer 2 substitute for current technologies like LSP on MPLS or even VLANs inside switching networks, in order to create specific paths for Ethernet traffic.

A FlexE topology can be divided into three types: terminated, aware and unaware.

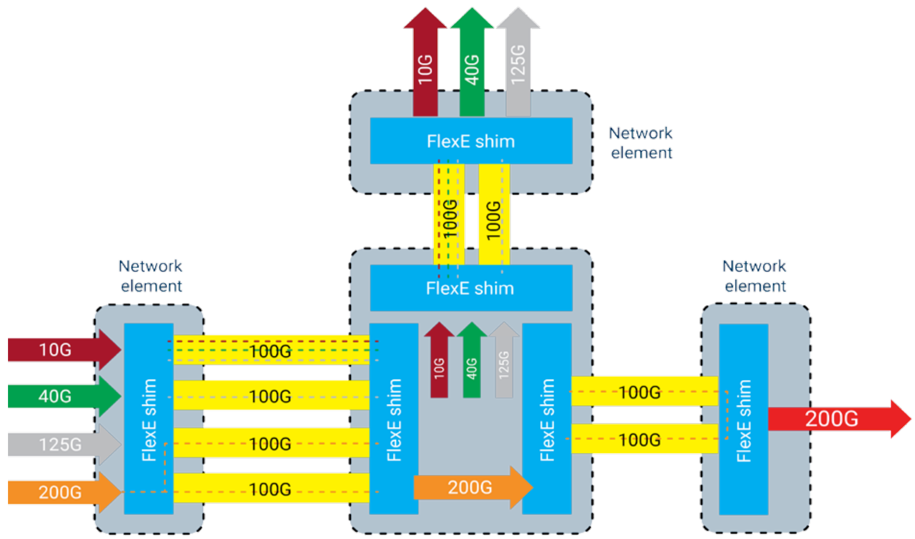


Figure 7. FlexE channelization example

Types of elements within the FlexE topology

Depending on the application implemented, there are three types of network elements within a FlexE topology.

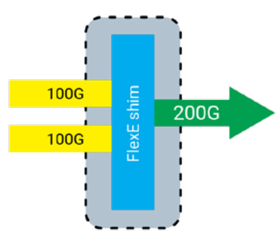
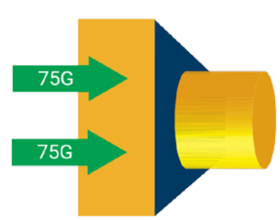
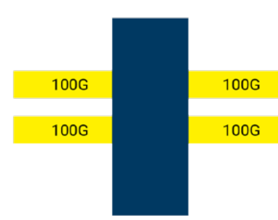
FlexE terminated	FlexE aware	FlexE unaware
An element that has a FlexE shim that terminates the data pipe.	An element that does not have a FlexE shim that terminates the data pipe and can match it to the line rate.	An element that does not have a FlexE shim to terminate the pipe but is capable of transporting traffic.
		

Table 1: Network elements inside the FlexE topology

FlexE applications

FlexE may still seem like a very new technology with still-unproven potential. However, the applications where this technology could prove important are growing every day. Here are some examples:

Network equipment manufacturers (NEMs): The first implementation of FlexE will directly involve network elements. Various NEMs are working hard to deliver equipment capable of generating different data pipes to transport FlexE clients, and most importantly, accomplish FlexE tasks like bonding, sub-rating and channelization. Today, NEMs already require test equipment with FlexE features to help them produce the next generation of routers and switches.

Data centers: The exponential growth of internet content and IoT is pushing data center infrastructures to the limit—forcing the deployment of more fibers faster and with increased efficiency. FlexE's bonding capability will allow data centers to easily increase bandwidth leaving the datacenter—efficiently aggregating and saturating 100GE PHYs now and 400GE PHYs as they become available. This will replace existing tools like LAG or ECMP.

Subsea cables: In their quest for bandwidth, data centers have acquired subsea cables supporting coherent optical transmission technologies. These technologies are able to transport non-standardized Ethernet rates; for example, 8QAM modulation, which is a type of coherent optical transmission that handles 150G per wavelength, a non-standardized rate. Some applications of sub-rating are attempting to match these types of line rates, allowing data centers to take full advantage of their current infrastructure.

Service providers: Inside service provider networks, a very different set of technologies coexist to transport customer traffic. Some networking schemes interchange VLANs between switches or generate an intricate set of tunnels (MPLS) between network elements to connect several sections of the network. These technologies add headers to the frames and bandwidth control is complex. FlexE channelization provides an alternative to manage current and future infrastructure better. Ethernet data pipes (channels) can be created easily, with specific bandwidth characteristics and direction inside the network.

5G networks: Due to its granularity, FlexE is not only providing alternatives for high-speed interconnectivity, it is being considered as a solution to the connectivity challenges of 5G. FlexE could address the 5G bandwidth problems by using bonding and channelization to efficiently aggregate low rate CPRI/eCPRI Ethernet clients (e.g., 10G or 25G) or potentially aggregate new standards (e.g., 50G) from the backhaul to the 5G core. In such circumstances, traffic coming from each residential user or enterprise line would become a FlexE client once they enter the network.

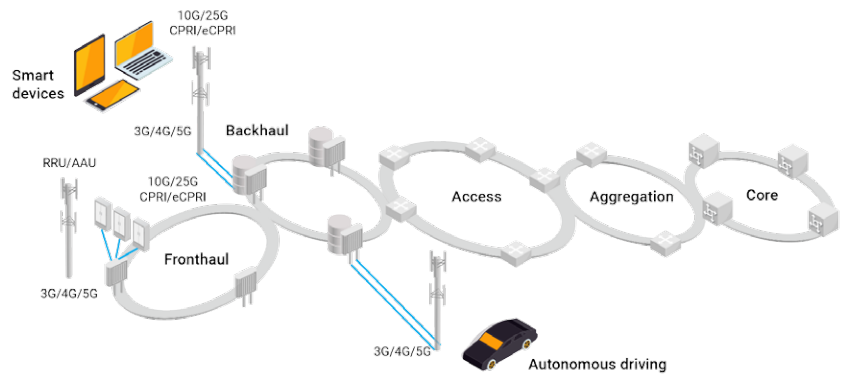


Figure 9. Example of a 5G FlexE implementation

Due to its flexibility and granularity, FlexE is being considered as a solution to the connectivity challenges of future 5G networks.

> Compliant with the OIF FlexE implementation agreement, EXFO's latest Power Blazer module includes four QSFP28 ports and is a robust testing solution.

FlexE testing

Once again, as the network test, monitoring and analytics experts, EXFO is taking a leading role in understanding and creating solutions that move the latest technology innovations forward. Understanding the new 400G Ethernet ecosystem has enabled us to design and introduce solutions that permit our customers to fully test FlexE capabilities provided by new network and transport elements. Our FTBx-88400NGE Power Blazer 400G multiservice tester is the most compact solution on the market and includes basic and advanced capabilities for lab and field implementations, in addition to our FlexE BERT application. This solution is a powerful tool for network equipment manufacturers, service providers and data centers in their quest to deploy 400G.

Compliant with the OIF FlexE implementation agreement, this latest Power Blazer module includes four QSFP28 ports and is a robust testing solution for today's challenges, with capabilities that include:

- Mapping a wide range of Ethernet rates
- Advanced features with BER analysis to stress the data pipe per client
- Monitoring a variety of alarms and errors, per PHY, group and client
- Providing visibility on FlexE shim characteristics

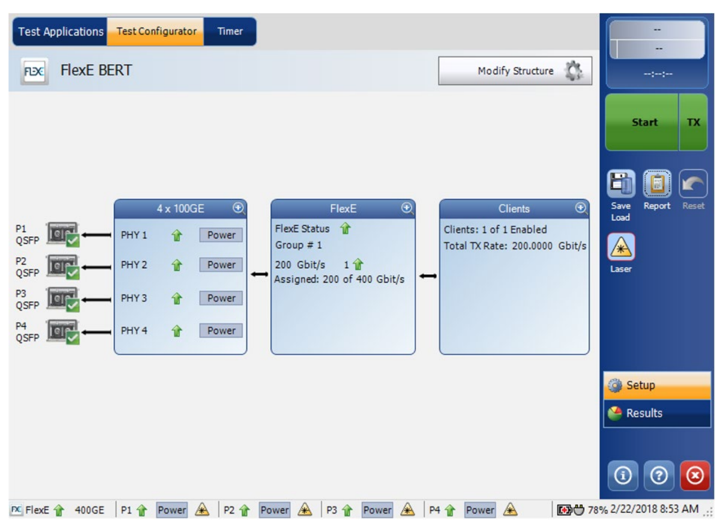


Figure 10. EXFO's FlexE BERT application

Learn more

Read about our solution:
FTBx-88400NGE Power Blazer - 400G multiservice test solution

Get our poster:
400G technical poster

Conclusion

FlexE is a very powerful tool that is breaking the chains that have held Ethernet networks captive to physical interfaces. It is opening the door for new client rates; generating new models that adapt more efficiently to the constant changes in technology standards and the transceiver evolution; and generating future-proof designs ready for upcoming standards.

FlexE is a very interesting alternative to aggregating and saturating ports across current high-speed infrastructure. Aligning client rates with transport technologies, maximizes current network interconnections and minimizes deployment times. FlexE is also a flexible and reliable alternative to low-rate transport infrastructure and could prove to be an important component in taking technologies such as 5G from lab to live.

APPNOTE358.1EN 18/03