

Transforming from 100GE to 400GE: objectives, differences and new physical interfaces

app
note

EXFO

Transforming from 100GE to 400GE: objectives, differences and new physical interfaces



app
note

Introduction

Like previous Ethernet speed increases, 400G is designed to satisfy the ever-increasing demand for higher bandwidth. High-speed networks are under intense pressure from consumer markets as data and traffic volumes explode with the proliferation of cloud services, smartphones, social media apps, UHD 4k/8k video streaming and ongoing implementation of the Internet of Things. In response, webscale companies are seeking faster transmission rates that allow them to deliver content in increasingly efficient ways.

More and more data centers are being built to support high-capacity data networks. These centers are interconnected using 100 Gbit/s links but will soon migrate to 400 Gbit/s to increase capacity. According to the high-speed roadmap published by the Ethernet Alliance (Fig. 1), 400G Ethernet (400GE) is poised to become the next client rate in the Ethernet ecosystem as the industry ramps up to handle the massive demands of hyperscale data centers, service providers and business users.

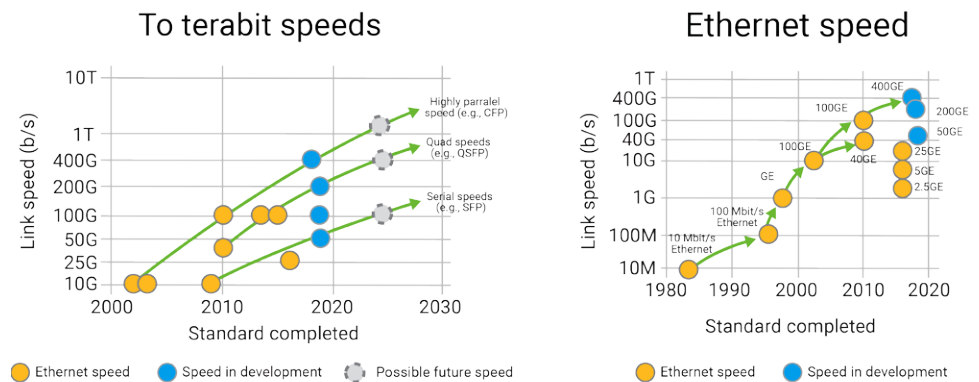


Figure 1: Ethernet Alliance roadmap

IEEE is developing a new standard for 400GE (IEEE 802.3bs) offering improved capabilities for key aggregation and high-bandwidth interconnect applications, including cloud-scale data centers, internet exchanges, colocation services, wireless infrastructure, service provider and operator networks and video distribution infrastructure.

This application note will highlight the key objectives of the new 400GE rate and the differences relative to 100GE. We will also take a look at the new supported physical interfaces.

400G Ethernet (400GE)

As was the case previously with the 100GE standard, the 400G task force has established a number of key objectives for the new 400GE rates: support MAC data rates of 200 Gbit/s and 400 Gbit/s, these are the new core router rates for transport applications; support a BER of 10^{13} or better at the MAC service interface; define new physical layer specifications for 200 Gbit/s and 400 Gbit/s operation and introduce a mandatory and improved Reed-Solomon FEC. This is a key difference over the 100GE objectives, where the FEC was optional. Another difference is the alignment of the different lanes: M:N gearbox conversion has been eliminated, and PCS, electrical and optical lanes are aligned and are multiples of each other.

The figure below illustrates the TX/RX PCS path structure.

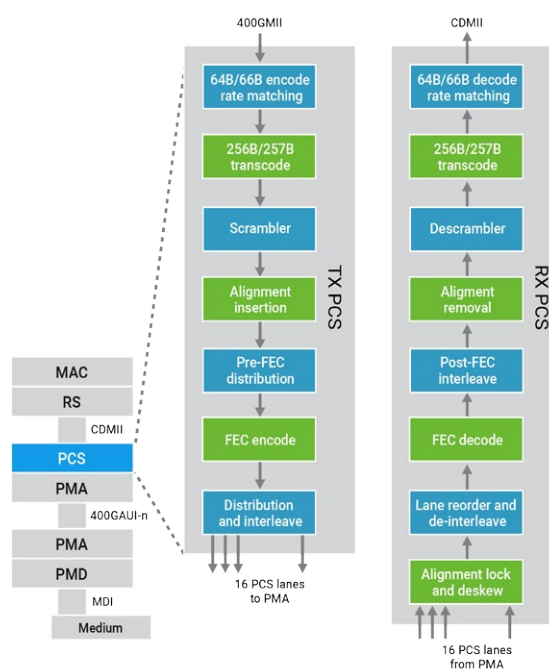


Figure 2: PCS layer

MAC layer traffic is fed directly to the encoders/transcoders and then to the alignment marker and the FEC encoder. Distribution is over 16 PCS lanes to the PMA, with PCS lanes running at a nominal rate of 26.5 Gbit/s.

The 400GAUI-n lanes are electrical lanes that feed the optical transceiver, where “n” can be 16 lanes or 8 lanes. The 400GAUI-8 will support 50G lanes for Ethernet and OTN 400G support.

Supported media defined in the IEEE 802.3bs 400G standard that can connect the 400GAUI-n are described in Table 1 below.

| Reach | Rates/ Optical lane | Media | IEEE 802.3bs | Modulation |
|------------|------------------------|--------------|-----------------|-----------------|
| SR (100 m) | 26.5625 Gbit/s | Parallel MMF | 400GBASE-SR16 | 16 x 25G-λ NRZ |
| DR (500 m) | 26.5625 Gbit/s | Parallel SMF | 400GBASE-DR4 | 4 x 100G-λ PAM4 |
| FR (2 km) | 53.125 Gbit/s | Duplex SMF | 400GBASE-FR8 | 8 x 50G-λ PAM4 |
| LR (10 km) | 53.125 Gbit/s | Duplex SMF | 400GBASE-LR8 | 8 x 50G-λ PAM4 |

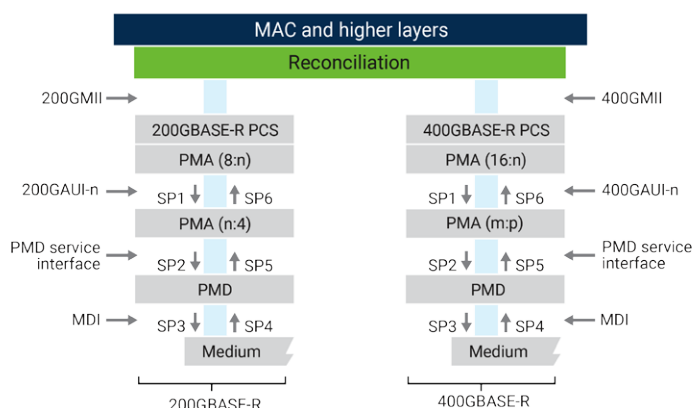
Table 1: Client media interfaces

The 400G PCS lane skew, where we measure the difference between the times of the earliest PCS lane and latest PCS lane for the transition of the alignment marker sync bits.

Skew variation may be introduced due to variations in electrical, thermal or environmental characteristics. The skew points at 400G are 4 times those at 100G: see Table 2 below for the maximum skew per skew point (SP).

| Skew points | Maximum skew (ns) ^a | Maximum skew for 200GBASE-R or 400GBASE-R PCS lane (UI) |
|----------------|--------------------------------|---|
| SP1 | 29 | ≈ 770 |
| SP2 | 43 | ≈ 1142 |
| SP3 | 54 | ≈ 1434 |
| SP4 | 134 | ≈ 3559 |
| SP5 | 145 | ≈ 3852 |
| SP6 | 160 | ≈ 4250 |
| At PCS receive | 180 | ≈ 4780 |

Table 2: PCS layer



The maximum skew and skew variation at physically instantiated interfaces are specified at skew points SP1, SP2 and SP3 for the transmit direction and SP4, SP5 and SP6 for the receive direction. 200GE and 400GE skew monitoring will be more challenging than with 100GE.

Forward error correction (FEC)

In addition to a more stringent skew, 400GE incorporates stronger FEC: RS (544,514) FEC is part of the bj FEC solution, with an estimated gain of about $\approx 8\text{dB}$. Key 400G applications are highly sensitive to latency, a drawback when using an FEC mechanism where coding and decoding increases the transmission delay. Unlike 100GE, 400GE uses a mandatory FEC. This means the link won't be error-free: correctable FEC errors are expected where the FEC will compensate for signal degradation and optimize link quality. Some FEC thresholds (Fig. 3) can be set to validate the link in pre-FEC mode, which in some cases can reroute user traffic without dropping it. The signaling used is similar to local/remote fault signaling, allowing for improved interoperability and the ability to show link health prior to packet errors. If the pre-FEC threshold is exceeded, it is treated the same as a sync header error threshold with local degrade (LD) and remote degrade (RD) signaling. It will be critical to validate FEC implementation.

Threshold crossing

FEC excessive threshold

- If the pre-FEC BER exceeds this **mandatory** threshold for longer than the *Interval*, link fault signaling is generated

FEC degrade threshold

- If the pre-FEC BER is higher than this **optional** threshold for longer than the *Interval*, a link degrade signal is generated

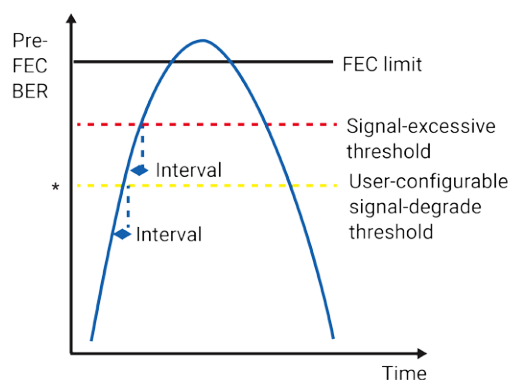


Figure 3: IEEE P802.3bs 400GbE Task Force

* The degraded SER has 2 thresholds; the first one is to declare a degraded SER, and the second is to clear the alarm. Both thresholds can be configured to a specific value.

Physical interfaces

The IEEE 802.3bs standard also lists proposed 400G interfaces. Network equipment manufacturers (NEMs) are already providing interface cards with CFP8 support. Today's CFP8s are based on an 8-lane or 16-lane electrical design (8 x 50 Gbit/s) (16 x 25 Gbit/s) with a reach of up to 10 km. Two types of PMDs are available: an NRZ modulation for 16 x 25 Gbit/s and a PAM-4 modulation for the 8 x 50 Gbit/s and 4 x 100 Gbit/s versions. The CFP8 supports eight times and four times the bandwidth-density of CFP and CFP2 form factors, respectively.

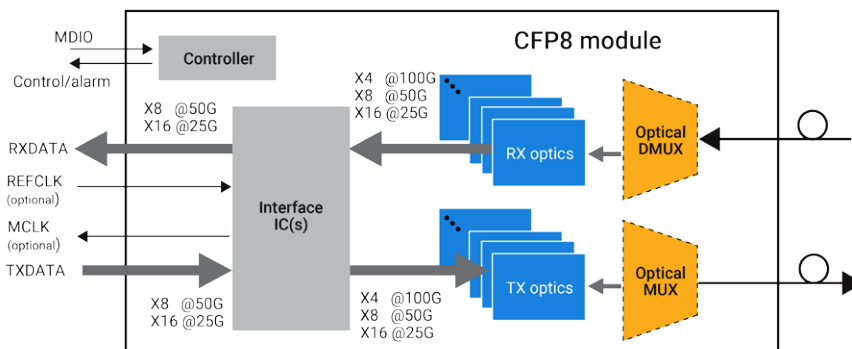


Figure 4: CFP8 functional block diagram

The following optical 400G interfaces are considered:

- **100 m MMF: 400GBASE-SR16:** 400 Gbit/s transmission over sixteen lanes (i.e., 32 fibers total)
- **500 m SMF: 400GBASE-DR4:** 400 Gbit/s transmission over four lanes (i.e., 8 fibers total)
- **2 km SMF: 400GBASE-FR8:** 400 Gbit/s transmission over an 8 wavelength division multiplexed (WDM) lane (i.e., 2 fibers total)
- **10 km SMF: 400GBASE-LR8:** 400 Gbit/s transmission over an 8 WDM lane (i.e., 2 fibers total)

The following optical 200G interfaces are considered:

- **500 m SMF: 200GBASE-DR4:** 200 Gbit/s transmission over four lanes (i.e., 8 fibers total)
- **2 km SMF: 200GBASE-FR4:** 200 Gbit/s transmission over a 4 WDM lane (i.e., 2 fibers total)
- **10 km SMF: 200GBASE-LR4:** 200 Gbit/s transmission over a 4 WDM lane (i.e., 2 fibers total)



Conclusion

High-speed networks are under intense pressure, where exploding traffic growth is pushing the industry to introduce high Ethernet rates, such as 400G Ethernet.

The EXFO 400G FTBx-88400NGE Power Blazer multiservice test module is an all-in-one 400G testing solution with basic and advanced capabilities for the lab and the field. Compatible with the IEEE 802.3bs 400G standard with support for the latest CFP8 MSA, the FTBx-88400NGE is a prevailing solution that allows users to test and validate their 400G designs. The module comes equipped for the most advanced 400GE tests, including lane skew measurement, Ethernet frame testing, BER thresholds, service disruption with pass/fail verdicts and 400GE filtering and capture. The module also supports the latest FEC implementation.



FlexE, FlexO, OTUCn, 25GE and 50GE testing capabilities are also available with the FTBx-88400NGE Power Blazer.

