APPLICATION NOTE

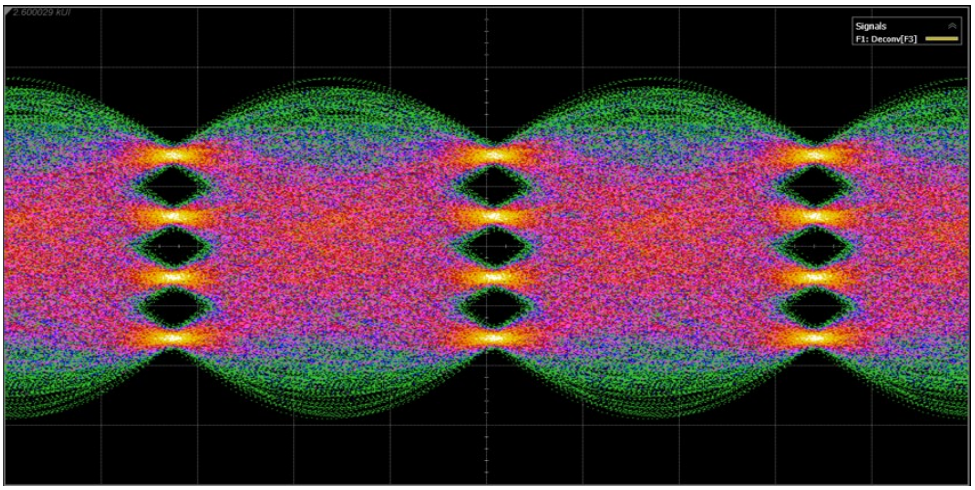
Data Center Ethernet Technology and Evolution to 224 Gbps

KEYSIGHT

Table of Contents

Introduction

As bandwidth demands for digital services (5G, Internet of Things etc.) continues to increase, data centers must upgrade infrastructure to keep pace. 800G and 1.6T research is underway, with single lane interface speeds up to 224 Gbps. This application note will focus on the latest evolution of high-speed Ethernet links in a modern data center, and the Keysight Technologies high-speed test solutions that are available for up to 224 Gbps interfaces.



Data Center Interconnects

Large-scale internet data centers are the fastest growing market for optical interconnection technology and innovation, with 70% of all internet traffic occurring inside the data center, driven by increasing machine-to-machine communication. A typical data center network structure based on CLOS[1] architecture (aka leaf-spine) is shown in Figure 1.

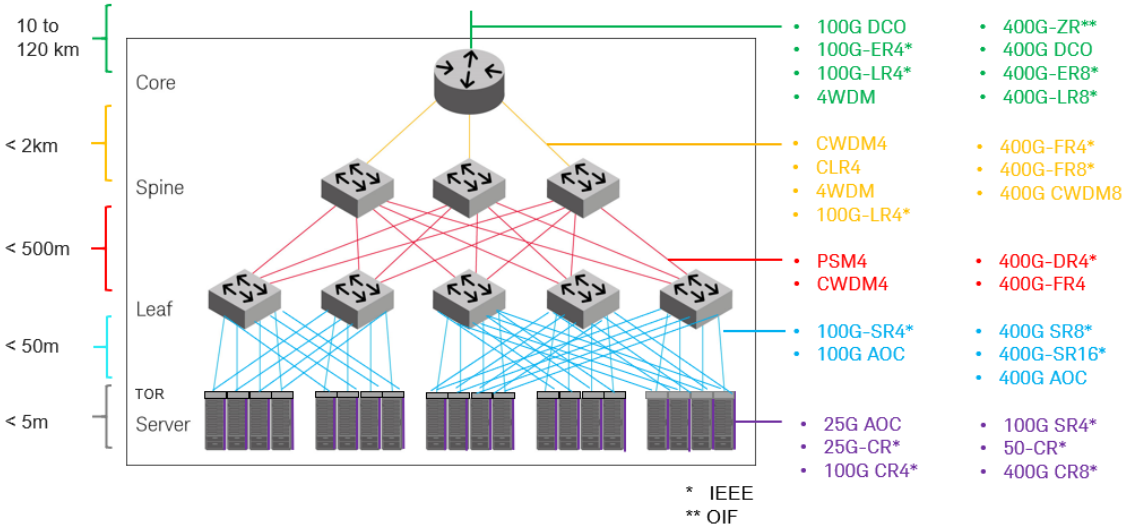


Figure 1. Network architecture of a typical hyperscale data center

A data center internal network usually has 3 to 4 levels from bottom to top.

Moving through the levels from the Server to the Core, the reach of each interconnection increases from a few meters to several kilometers, necessitating changes of technology and interface standards.

Server Cabinets / Top of Rack Switch (TOR): At the lowest level, individual server racks are connected to TOR switches at the top of the cabinet. Current data centers generally deploy 25G networks, with some artificial intelligence (AI) applications utilizing 50G speeds. Over the next few years 100G, 200G and 400G speed interconnection technology will be employed. Connection distances are short, being either within the cabinet or to adjacent cabinets and generally less than 5 meters. A typical interface technology used today is direct attach copper cable (DAC) or active optical cable (AOC). As speeds evolve to 400G and 800G the reach of DACs will be too short and active electrical cable (AEC) will be used instead.

TOR to Leaf Switch: The second level is the connection from TOR switches to Leaf switches. This distance ranges up to about 50 meters, using 100G interconnection technology now and moving to 200G and 400G speeds and in a few years to 800G. Typically optical modules such as 100GBASE-SR4 or 200GBASE-SR4 combined with multi-mode optical fiber are used today along with NRZ (Non-Return to Zero) signaling. For this level and the higher-level interconnections, the move to 200G and 400G also changes the signaling to PAM4 (Pulse Amplitude Modulation 4 level).

1. CLOS Topology https://howdoesinternetwork.com/2019/clos-topology

Leaf to Spine: The leaf to spine connection may be within campus, or adjacent campus, with connection distance up to 500 meters. Using similar interface rates as TOR to Leaf, 100G moving to 200/400G now and 800G around 2023. With the longer reach, the technology moves to single mode fiber and often several parallel fibers utilizing modules such as 100G-PSM4, 100G-CDWM4 and moving to 200GBASE-DR4 and 400GBASE-DR4.

Spine to Core: As the reach increases further up to 2 kilometers, the cost of fiber starts to be a consideration, and so wavelength division multiplexing technology is often used to send data via several different optical wavelengths on one fiber, today using modules such as 100GBASE-LR4, 100G-CWDM4, 400GBASE-ER4/-LR4/-FR4 etc.

Data Center Interconnect (DCI): This is generally a connection between several adjacent data centers for load balancing or disaster recovery backup. The distance may range from tens of kilometers to around a hundred kilometers. Over this longer distance dense wavelength division multiplexing is employed and, more recently, coherent communication is being used in preference to direct detect technologies. Telecom operators have deployed 100G coherent technology for many years in long-distance (hundreds of kilometers) applications. Speed increases into 200, 400, 800G technology is also ongoing. For DCIs, since the transmission distance is not as far as the telecom applications and is mainly point-to-point, coherent transmission is feasible using pluggable module technology with smaller size and power consumption, such as 400G-ZR.

The Development of Data Center Interconnection Technology

As shown in Figure 1 there are several different electrical and optical interconnection technologies in use in the data center, and they are continuously evolving.

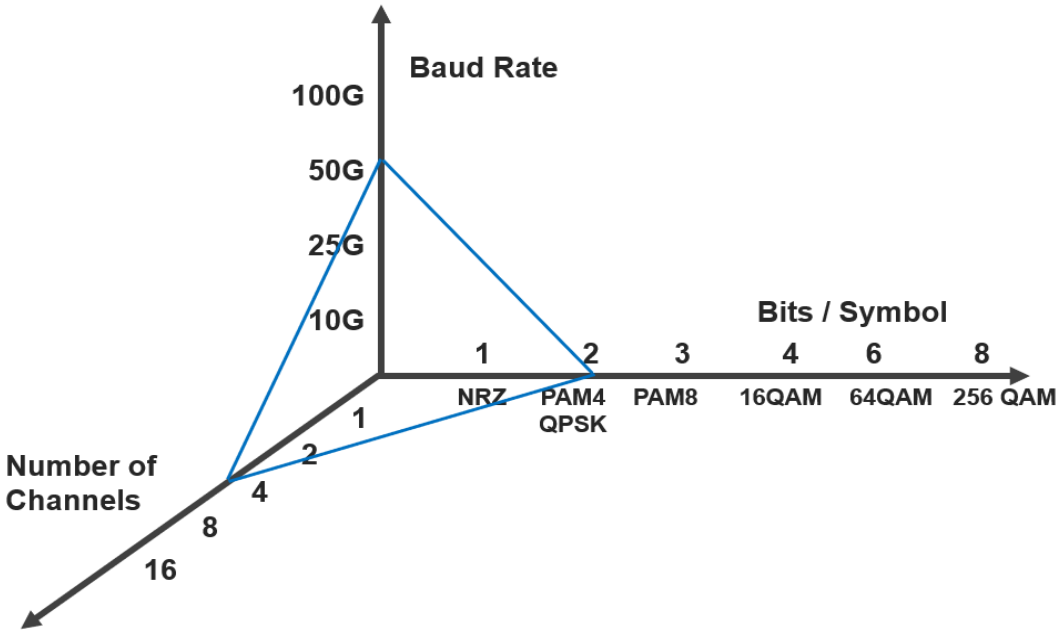


Figure 2 Technology for improving the speed of Ethernet interconnections in the data center

The speed of each interface can be achieved by more than one method or interface standard, each having different trade-offs between performance, reach, power consumption and cost.

There are three technical directions to increase the speed of the interconnection interface (Figure 2):

The first method is to directly increase the data or baud rate of the channel, e.g. the development from 155 Mb/s to 622 Mb/s in the SDH/SONET era, or 100 Mb/s Ethernet ports all the way to 10 Gb/s Gigabit Ethernet ports. Often the baud rate improvement required can be ahead of the available technology at the time so other methods have been used.

The second method is to increase the number of channels. This has the advantage of keeping the baud rate constant but does bring additional cost and complexity to the interface design. For example, the Ethernet interface transition from 10 Gb/s to 40 Gb/s employed a 4 x 10 Gb/s channel approach in preference to a single-channel 40 Gb/s link and its high cost of implementation. The same approach was used to move to 100G Ethernet, using initially 10 x 10 Gb/s channels, and later 4 x 25 Gb/s channels, which has become the mainstream 100G Ethernet interface implementation. For the electrical interfaces this approach always means more channels for the devices and circuit boards, introducing crosstalk as a new design consideration. For the optical interface, multiple channels can be implemented as parallel multi-mode or single-mode fibers for short-distance transmission and by using wavelength division multiplexing (WDM) on a single fiber for longer-distance transmission. The type and number of available fibers per port is usually set by the existing fiber infrastructure, governing the deployment of higher speed

optical interfaces in an existing datacenter. Generally, 4 or 8 wavelengths are used in WDM (5 nm spacing) or CWDM (20 nm spacing). Some leading-edge research is studying few-mode multi-core fibers where multiple cores are made from a single fiber to realize spatial multiplexing transmission.

The third method is to use a more complex modulation method. Data rates up to 25 Gb/s have used NRZ signaling. When the industry put forward the technical requirements for 400G Ethernet, it was challenging to increase the data rate to 53 Gb/s at that time, especially in the electrical domain, pushing the limits of device bandwidth, packaging, and PCB design. Increasing the number of channels increases the space requirements and power consumption of solutions and does not help in the overall cost/bit reduction target for data centers. As a result, PAM4 complex modulation was proposed over NRZ. PAM4 4 level modulation enables each data symbol to carry 2 bits / symbol, doubling the interface data rate for the same number of channels and baud rate.

Complex modulation techniques are commonly used in the field of long-distance coherent optical communication. For example, 100G coherent communication generally uses QPSK modulation, where one symbol can carry two bits; while 400G coherent communication uses 16-QAM modulation, with 4 bits per symbol. In wireless communications and new coherent implementations 256-QAM is used with 8 bits / symbol throughput.

All of the methods above have been used at different times to increase data throughput in the data center. Figure 3 summarizes the development of electrical and optical port data rates for both the 100G and 400G Ethernet standards. In addition to the data rate increase, the transceiver form factor and internal architecture has also evolved to reduce size and power consumption.

The earliest 100G transceivers were packaged in CXP or CFP, with 10 lanes for both electrical and optical ports. With the subsequent development of a gearbox, 10 x 10 Gb/s electrical lanes could be converted to 4 x 25 Gb/s optical lanes. This increased complexity and power consumption in the transceiver until further development allowed for 4 x 25 Gb/s electrical lanes and the transceiver architecture could be simplified back to just a CDR for re-timing the signals. This is the most common approach today, and together with the QSFP28 form factor is a good balance of cost, power consumption, size, and performance.

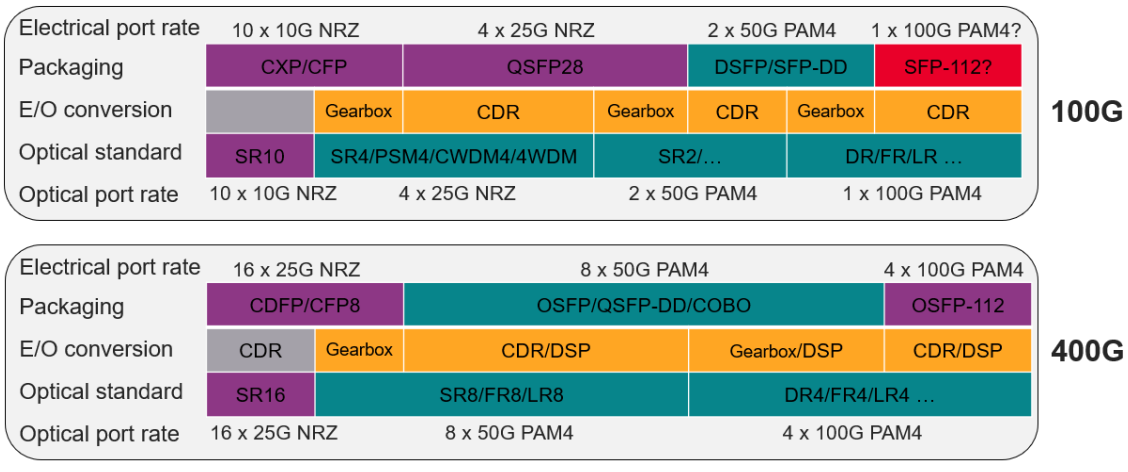


Figure 3. The development of 100G/400G Ethernet electrical and optical port data rates

Further development of 100G is possible, switching to PAM4 signaling with either 2 x 50 Gb/s or 1 x 100Gb/s lanes per port.

The technical development of 400G Ethernet follows a similar process. Initial development centered on CDFP or CFP8 packaging and 16 lanes of NRZ signaling. However, this was never commercially viable from a cost/complexity standpoint. Today the most common implementation is 8 x 50G PAM4 (i.e. 25 GBd) at the electrical port and a combination of either 8 x 50G PAM4 or 4 x 100G PAM4 (50 GBd) at the optical port, using QSFP-DD or OSFP form factors. Again, to produce 4 x 100G optical lanes from 8 x 50G electrical lanes requires a gearbox which while increasing complexity and power consumption is offset by needing only 4 lasers and 4 detectors.

The optimum configuration would be to have the same number of electrical lanes as optical lanes, eliminating the need for a gearbox. As the industry moves to 800GE and 100G electrical lanes, the 400G technology will advance to 4 x 100G PAM4 electrical and optical lanes.

Moving to 800G ethernet

Development has commenced for next Ethernet speed class - 800G. The first generation of 800G will employ 112 Gbps per lane, enabling 200/400/800G links, and the second generation will introduce 224 Gbps per lane for up to 1.6T links. Standards organizations have all commenced projects or working groups for 800G e.g.

- OIF Common Electrical I/O (CEI)-112G and Common Electrical I/O (CEI)-224G
- IEEE 802.3 Beyond 400 Gb/s Ethernet Study Group
- 800G Pluggable MSA (Multi-Source Agreement)
- OSFP MSA 200G/lane Electrical Signaling Group

Figure 4 shows the technical roadmap of 800G optical interconnection as described in a white paper issued by the 800G-MSA standard organization. The 800G Ethernet optical ports will mainly have two implementation methods. One will use 8 parallel single-mode fibers (e.g. 800G-PSM8) for short-distance connections up to about 100 meters at 100G per fiber. The other will use either 4 parallel single-mode fibers (e.g. 800G-PSM4 or 800G-DR4) or 1 single-mode fiber with WDM (e.g. 800G-FR4) for medium-distance connections up to about 2 km at 200G per fiber or per wavelength. The IEEE802.3db working group is also discussing the feasibility of using 8 channel multimode fiber to achieve short distance (tens of meters) 800G transmission, but currently it is limited by the bandwidth of commercially available VCSEL lasers. The feasibility of its industry acceptance remains to be seen.

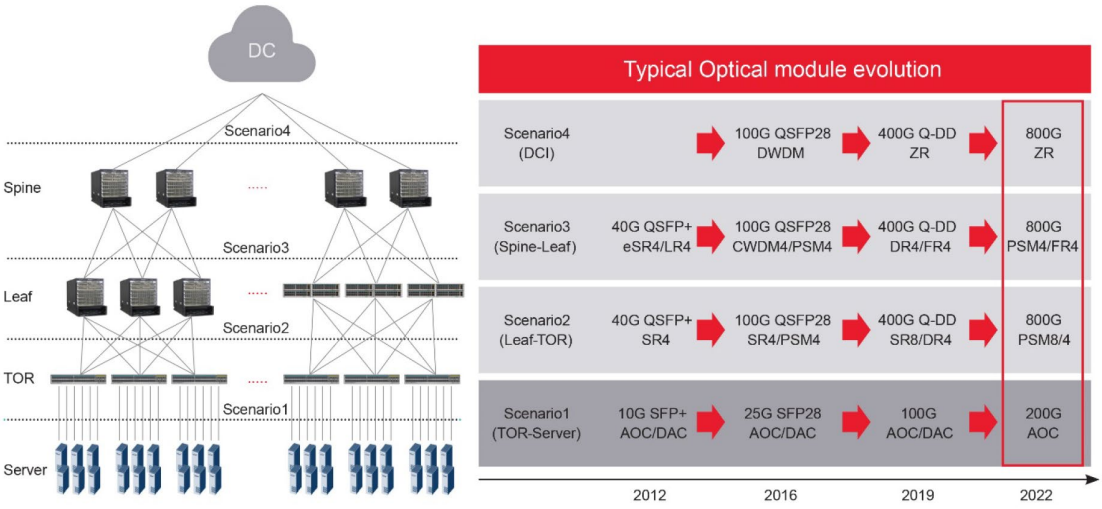


Figure 4. The technical roadmap of optical interconnection organized by the 800G Pluggable MSA [2]

The implementation of a typical 800G optical interconnection is shown in Figure 5, as either an 8 channel optical port solution or a 4 channel optical port solution. The yellow blocks in the figure indicate the addition of 800G Ethernet to the currently available 400G Ethernet technology.

The 800G-PSM8 or -DR8 implementation is not very different from 400G-DR4, except that the number of optical channels has doubled. There is already mature chip and industry supply-chain support. At the electrical interface, the previous 400G Ethernet used 8 channel 50 Gbps connections. For 800G, the rate of the electrical interface between the 800G Ethernet switch chip and the optical module must be increased to 100 Gbps per channel so that 8 channels can provide 800 Gbps overall. 800G Ethernet optical modules can also use QSFP-DD or OSFP packages commonly used for 400G optical modules, but the performance needs to be improved to support higher electrical interface rates. Since the end of 2020, chip manufacturers have released CDR/DSP chips supporting single-channel 100 Gbps and higher-performance QSFP-DD and OSFP packages. Therefore, the technology is available to implement 8 x 100 Gbps electrical and optical ports to achieve 800G connections. If power consumption and cost can be controlled, this solution could be commercialized relatively quickly.

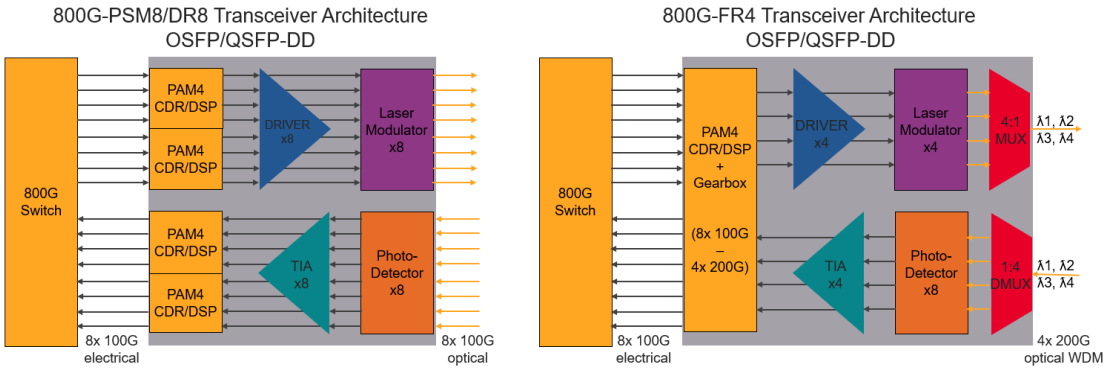


Figure 5. Typical implementations of 800G optical interconnection

2. Reference: Enabling the Next Generation of Cloud & AI using 800Gb/s Optical Modules. https://www.800gmsa.com/documents/white-paper

The 800G-FR4 4 channel solution brings a new challenge, because the optical module also needs a gearbox to convert 8 channels of 100 Gbps electrical signals into 4 channels of 200 Gbps electrical signals. This then drives 4 optical modulators that can support the corresponding 200 Gbps rate. It is necessary to define a new optical port standard and devices that can handle the higher bandwidth such as DSP, driver, modulator, TIA, etc. At present, this technology is still in the pre-research stage, with modulation mode, baud rate, system bandwidth, link budget, bit error rate, FEC mode, etc. all being explored and discussed.

Data rate terminology

There are various nomenclatures used for the Ethernet data rates particularly with the move to 400G and now 800G. Some common terms are 100G/lane, 200G/lane, 100G/lambda, 112 Gbps, 224 Gbps etc.

For 800G first generation, the IEEE 802.3 standard data rate is 53.125 GBd PAM4 which is 106.25 Gbps. Whereas the OIF CEI specifies a data rate range from 38 to 58 GBd PAM4 which is 76 to 116 Gbps but has settled on 112 Gbps as the descriptive title of the interfaces.

For the remainder of this document, we will use the serial rate descriptions 112 Gbps and 224 Gbps respectively to describe the 800G first and second generations, both optical and electrical.

Technical challenges of 800G ethernet

The speed of the previous 400G electrical lanes is doubled to 112 Gbps for the first generation of 800G Ethernet and goes to 4 times (224 Gbps) in the second generation. Interface chips, DSP chips, packaging, connectors etc. all will need performance improvements or new designs to work at the higher speeds. There are already some corresponding standards in the industry that define 112 Gbps electrical ports, such as the OIF CEI-112G family of implementation agreements and the IEEE 802.3ck standard. Both standards bodies define electrical link requirements for various distances or reaches.

The standards formulated by the OIF for different scenarios have been initially released, ranging from short distance chip to chip connections such as CEI-112G-MCM/XSR, medium distance chip to module CEI-112G-VSR, and up to longer distances through backplanes and copper cable such as CEI-112G-MR/LR. The IEEE 802.3ck working group is defining complementary standards for chip to chip (C2C), chip to module (C2M), backplanes (KR) and copper cable assemblies (CR).

IEEE and OIF standards are highly leveraged from each other, though there are some differences in parameter definitions and measurement methodology. OIF CEI-112G-VSR, -MR and -LR have an equivalent clause in IEEE 802.3ck as shown in Table 1.

OIF-CEI standard	IEEE 802.3ck standard	Max. channel loss
CEI-112G-MCM	-	6 dB
CEI-112G-XSR		10 dB
CEI-112G-VSR	100GAUI-1 C2M[3]	16 dB
CEI-112G-MR	100GAUI-1 C2C[3]	20 dB
CEI-112G-LR	100GBASE-KR1[3]	28-30 dB
-	100GBASE-CR1[3]	24 dB

Table 1. OIF-CEI-112G and IEEE 802.3ck equivalent clauses

Taking the CEI-112G-VSR or 100GAUI-1 C2M interface as an example, the loss of the channel from host IC to module IC can be up to 16 dB maximum at the Nyquist frequency which is very high for chip packaging, PCB materials and connector performance requirements. To overcome this problem, new optical module packaging and electrical connections have been developed.

A new type of QSFP-DD800 package, which can support 8 channel 112 Gbps electrical signal transmission is shown in Figure 6. The QSFP-DD800 package is compatible with the 400G QSFP-DD package. In addition, an OSFP package is also available. The OSFP package is larger which has advantages in situations where more lasers need to be included or greater power consumption needs to be supported.

3. Also includes the 200GAUI-2, 400GAUI-4 and 200GBASE, 400GBASE variants

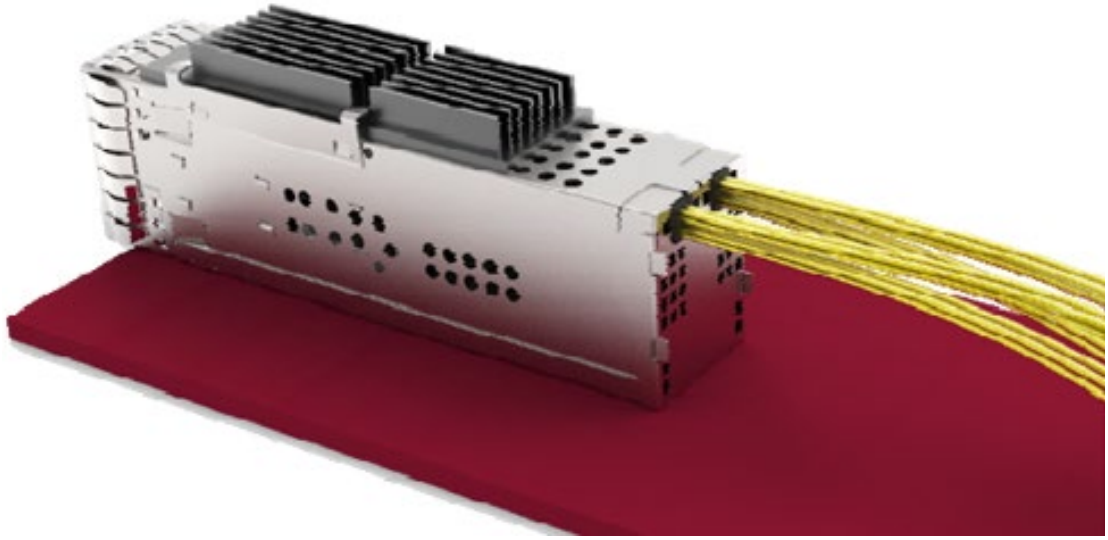


Figure 6. QSFP-DD800 package [4]

Also important is the PCB loss, which basically scales with increasing baud rate, therefore is about 1.5 to 2 times that of 400G modules and restricts the usable channel length at higher baud rates. Some alternative designs employ cable assemblies in preference to PCB material to minimize the loss. Figure 7 shows two examples where cable assemblies are used to realize the electrical connection between the switch chip and the optical module slot: Samtec's Flyover technology and Molex's BiPass solutions. The insertion loss of these cable assemblies can be about half that of a comparative length printed circuit board trace. Some companies are also investigating methods of directly attaching coaxial cables to the switch chip.

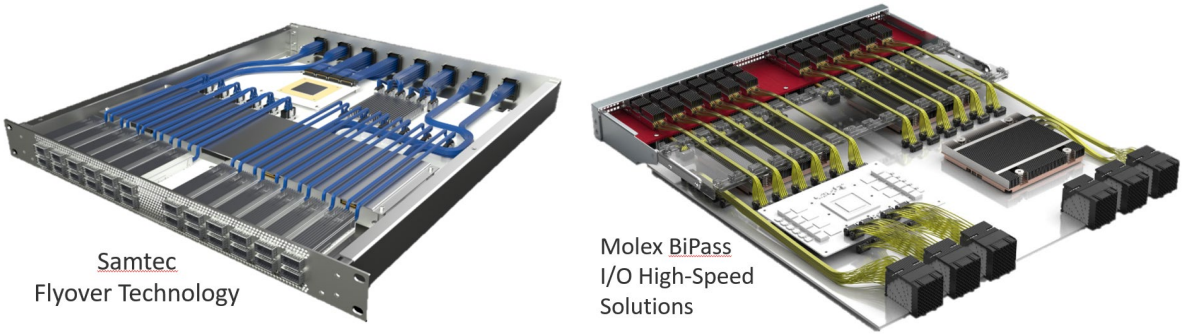


Figure 7. Examples of high-speed cable assemblies [5]

In future developments, the electrical output interface of the chip will be replaced by using co-packaged optics (CPO) technology. CPO will directly integrate laser engines on switch chips and computing chips, providing an optical connection interface. At present, there are still many technical problems to be solved, such as power consumption, heat dissipation, integration of different materials, etc

4. Reference: www.qsfp-dd.com
5. Reference: www.samtec.com and www.molex.com(Drawing courtesy of and copyright Molex, LLC.)

Another key factor for 800G Ethernet technology is the data throughput of the switch chip. Currently 25.6T switching capacity chips supporting 100 Gbps electrical interfaces are available, with the next generation of 51.2T chips coming in the next 1 to 2 years.

For the optical port, 112 Gbps technology has been mature and commercialized for 400G Ethernet (e.g. 400GBASE-DR4) so there are few challenges in technical implementation for the first generation of 800G. The real challenge comes from the single-channel 224 Gbps optical port technology that is required to achieve 800G connections with 4 channels (e.g. 800G-FR4) per Figure 5.

At present there are no published standards for 224 Gbps transmission, though research is under way and working groups started. One topic is modulation and whether to continue with PAM4 or go to higher order schemes such as PAM6 (~2.6 bits per symbol), PAM8 (3 bits per symbol) or even PAM16 (4 bits per symbol). Other possible approaches are discrete multi-tone (DMT) or partial-response PAM4. The advantage of continuing to use PAM4 modulation is that there has been much research on this technology, but the disadvantage is that the requirements for device bandwidth on the entire link are relatively high. PAM4 at 224 Gbps has a baud rate of 112 GBd, so the bandwidth of the entire system including modulator, detector, ADC, and DAC must be about 80GHz. A major design challenge at this baud rate is reflections, which become a dominant waveform impairment, requiring complex and power-intensive equalization schemes.

Since PAM6 and PAM8 enable more bits per symbol the baud rate for 224 Gbps is reduced to 87 GBd and 75 GBd respectively. However, the signal-to-noise ratio will be higher, and sensitivity to noise and jitter also increased, as well as more complex SerDes design required to handle the additional signal levels. At present, PAM4 technology has just begun to enter commercial use after years of R&D work in the industry. While sticking with PAM4 modulation at 112 GBd seems the most logical step, new modulation schemes may be adopted for 224 Gbps links. Maintaining a high degree of flexibility is needed for both technology choices and test & measurement equipment investment.

The use of higher-speed optical interfaces also requires the introduction of new bit error ratio standards, new FEC (forward error correction) coding methods, and more complex receiver equalizers. For 100G Ethernet, FEC is not mandatory but the move to 400G Ethernet and PAM4 modulation makes it essential for optical links to achieve zero error ratio (1E-12 or 1E-15 in practice). In 400G Ethernet, with a link BER of less than 2.4E-4 and even distribution of errors, reliable communication can be ensured with RS(544,514) Reed-Solomon FEC (aka KP4). For 800G Ethernet, and a single-channel speed of 224 Gbps the link BER will be higher (e.g. 1E-3) and will require a more powerful FEC error correction method, either new encoding or adding additional FEC on top of the KP4 encoding. With the additional overhead from a more complex FEC scheme, the link end-to-end delay and chip power consumption will also increase. Link delay (latency) is critical for high-performance computing applications like AI, therefore, the final FEC method selection will be based on tradeoffs between BER, FEC capability, latency, and power consumption.

Another very important factor that affects the bit error rate of the system is the equalization ability of the receiver. Because the multi-level signal has a small eye opening and will be further affected by dispersion and loss through optical fiber transmission, it takes complex equalization to open the eye at the receiver. IEEE 802.3bs defined a reference equalizer with 5-tap FIR filter for 400G Ethernet optical compliance testing. In practice, however, the equalizers actually implemented by many chip companies are much more complicated than this reference equalizer, e.g. 10 or 20 tap. In the development of 224 Gbps technology, more complex equalization will be needed, which will also increase the power consumption of

the system DSP chip. At the same time, more complex reference equalizers must be developed for the test and measurement instruments used to measure the optical signal parameters.

As we have discussed, there are still many technical challenges for realizing single-channel 224 Gbps optical signal transmission with direct modulation – the modulation method, signal baud rate, device bandwidth, signal-to-noise ratio requirements, system error rate, FEC, equalizer mechanism, etc. These are key technologies that need to be developed in the upcoming months.

Test Solutions

112 Gbps test solutions

IEEE 802.3bs and 802.3cd standards cover the transmitter and receiver compliance testing for 56 Gbps and 112 Gbps optical interfaces. These technologies are mature. Current standards creation in the OIF CEI-112G program and IEEE 802.3ck define the compliance requirements for electrical lanes at 112 Gbps.

Transmitter measurements are performed with a high bandwidth oscilloscope, measuring key parameters such as TDECQ, extinction ratio (ER), jitter, eye opening etc. Receiver measurements employ a stressed receiver input test, where a 'worst-case' signal is created using a BERT and adding inter-symbol interference (ISI), jitter, and interference to the signal then measuring the BER performance of the receiver in the presence of the stressed signal. This stressed signal is first calibrated using a high bandwidth oscilloscope.

For example, let's take a look at the test of the 112 Gbps electrical interface used for the connection of the switch and the optical module. The electrical interface standard is defined in detail in the OIF CEI-112G-VSR or the IEEE 802.3ck 100GAUI-1 C2M interface.

Figure 8 shows a typical setup to test the Host transmitter output signal quality. One difference between the IEEE 802.3 and OIF-CEI standards is the data rate, which is 106 Gbps (53 GBd PAM4) in the IEEE standard, and a data range up to 112 Gbps (56 GBd PAM4) in the OIF CEI standard. For transmitter signal quality compliance testing, a high-bandwidth sampling oscilloscope or a real-time oscilloscope is used, with a bandwidth of at least 50 GHz. The scope should be capable of modelling a reference receiver bandwidth of 4th order Bessel-Thomson or 4th order Butterworth response. Test fixtures - host compliance board (HCB) and module compliance board (MCB) are used, which together with the long/short channel emulation means the transition times are not particularly steep. However, if you are doing performance research and evaluation of the device or chip, it is recommended to use an oscilloscope with a bandwidth of at least 80GHz.

In addition to the requirements for oscilloscope bandwidth and noise floor, many analysis algorithms are involved in the test of PAM4 electrical signals. For example, it is necessary to perform clock recovery on the signal according to the standard loop bandwidth requirement (about 4 MHz) and accumulate data to form an eye diagram. The reference receiver model should include CTLE equalization (up to 12 dB gain) and 4-tap DFE equalization, and a means to determine the optimum CTLE and DFE settings, as well as the algorithms to measure eye opening, SNDR, etc. as defined in the standard.

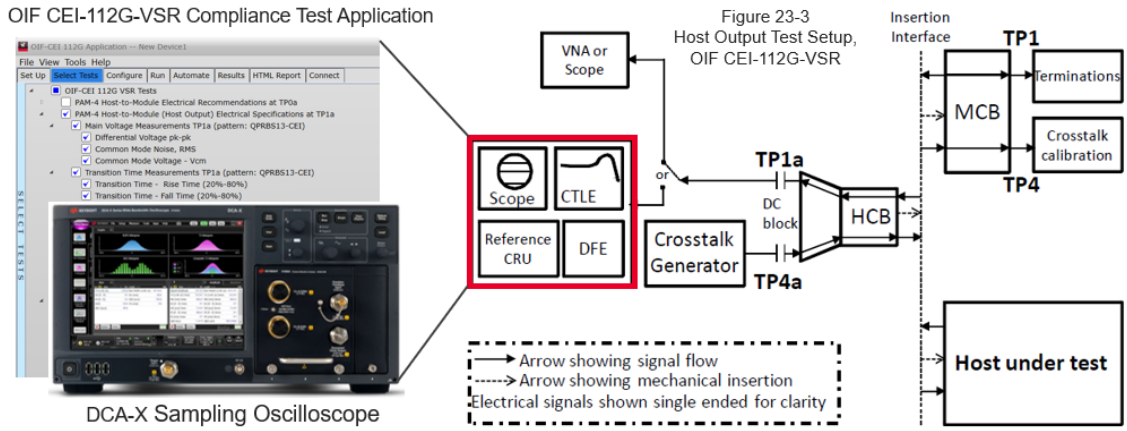


Figure 8. 112 Gbps electrical Host output signal quality test

This is a complex procedure to perform manually and without making measurement or setup errors. Keysight provides transmitter compliance applications for both the IEEE 802.3ck and OIF-CEI-112G standards that guide you through the required measurements and automate the equipment and equalization settings. These compliance applications are available on the DCA-X sampling and UXR-series real-time oscilloscopes.

Tx test solution	Standard covered	Reach/Link	Available on
N1091CKCA	IEEE 802.3ck	C2C, C2M, KR, CR	DCA-X
N109212CA	OIF-CEI-112G	VSR, MR, LR	DCA-X
D90103CKC	IEEE 802.3ck	C2C, C2M, KR, CR	UXR-series
D9050CEIC	OIF-CEI-112G	VSR, MR, LR	UXR-series

Figure 9 shows a typical setup to test module receiver performance. Before running the test, a calibrated stressed signal must be generated. This is accomplished by connecting the BERT pattern generator through a mated MCB-HCB pair to a high bandwidth oscilloscope. Noise, sinusoidal jitter, bounded-uncorrelated jitter (BUJ), unbounded uncorrelated gaussian jitter (UUGJ), ISI and crosstalk are added to the data signal while measuring signal parameters such as vertical eye closure (VEC), signal to noise and distortion ratio (SNDR), eye height etc. until the stressed signal complies with the definition in the standard. Once calibrated, the module under test is connected to the MCB and BER and jitter tolerance testing performed. BER can be measured either by looping the signal through the module under test and back to the BERT analyzer or using the module-under-test's internal BER counters.

Calibration of the stressed signal is complex, requiring optimization of de-emphasis and CTLE, DFE equalization settings and uses many of the same measurements used during transmitter testing. An iterative process of measuring and adjusting stress parameters makes the process labor intensive and time-consuming.

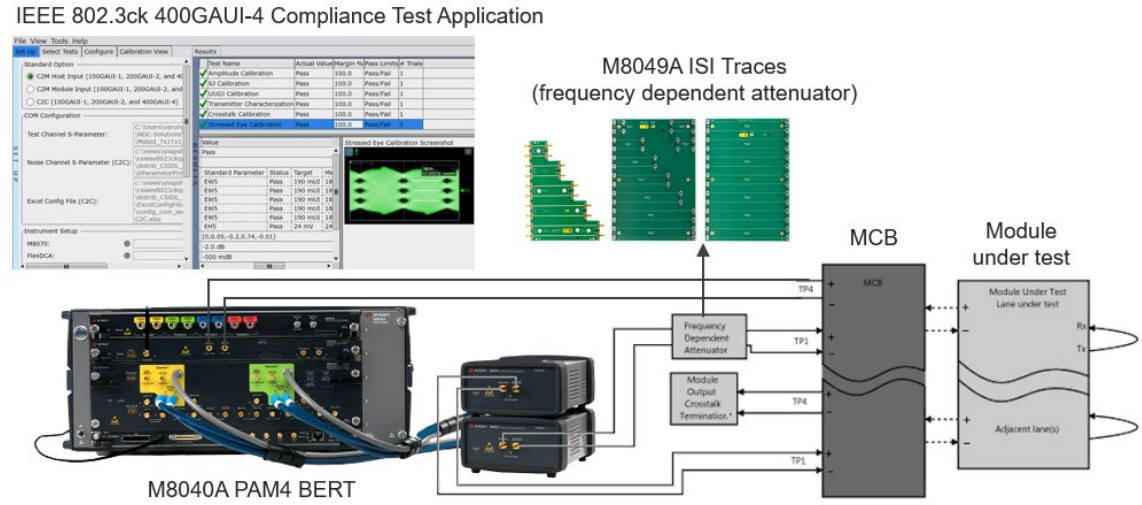


Figure 9. 112 Gbps electrical Module input stressed BER test

Receiver stress test efficiency is best achieved through automation of the entire setup including both calibration and test. Keysight provides receiver compliance applications for both the IEEE 802.3ck and OIF-CEI-VSR standards. These applications step through various calibration routines to achieve a compliant stressed signal and the required device tests. Application note 3121-1220.EN "Conformance Testing of 800G Ethernet Links for the Data Center" describes the Tx and Rx testing in detail.

Rx test solution	Standard covered	Reach/Link
M8091CKPA	IEEE 802.3ck	C2C, C2M
-	OIF-CEI-112G	VSR, MR, LR

224 Gbps test solutions

Keysight Technologies test solutions portfolio addresses the entire design cycle, from simulation, design validation, conformance test, protocol test and manufacturing. State-of-the-art equipment and test solutions enable early research into emerging technologies well before measurement standards are available. Later, as standards develop, close association and participation in the standards organizations enables Keysight to provide first-to-market Test Solutions allowing timely characterization and validation of component and systems designs to run concurrently with the standards development cycle.

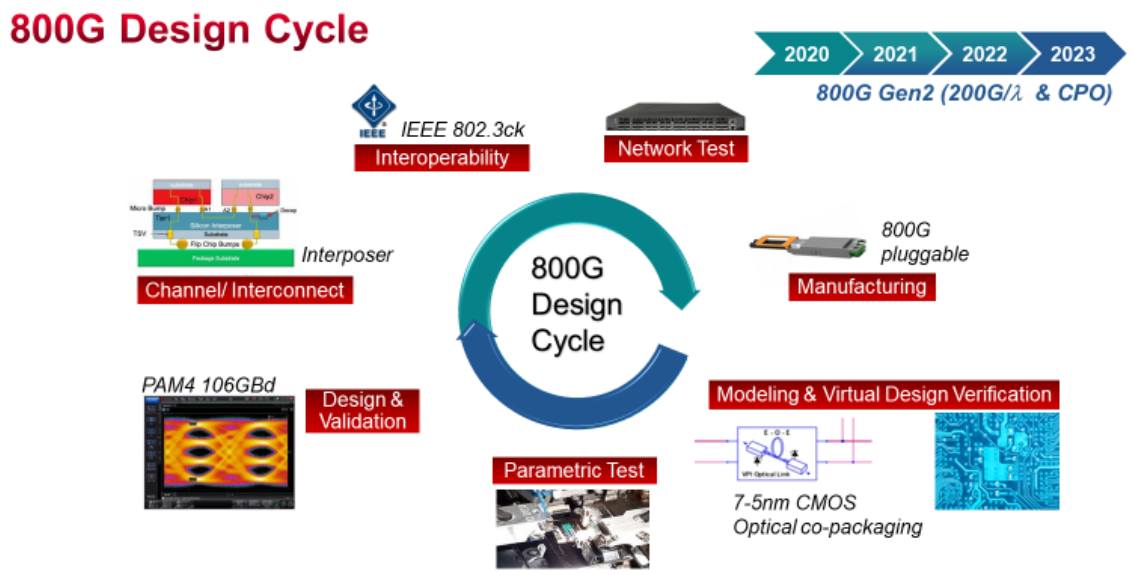


Figure 10. 800G Design Cycle

Signal generation

The path to 224 Gbps transceiver development starts with performance research at the component level such as optical modulators and detectors, driver amplifiers, switch chips etc. To test these devices requires a high-speed electrical test signal generator with flexibility to adjust amplitude (to adapt to the voltage requirements of different devices), linearity (to compensate for device gain compression), modulation format (to study the performance difference of different PAM-n modulation schemes), de-emphasis (to compensate for device packaging and transmission line bandwidth) etc.

Suitable commercial equipment that can be used in the research activities for 224 Gbs are a high-bandwidth arbitrary waveform generator (AWG) or a high-performance BERT pattern generator.

Figure 11 shows an M8199A high-bandwidth 4 channel AWG with a sampling rate per channel of 256 GSa/s and a DAC resolution of 8 bits. AWG's offer the highest flexibility in signal generation and 4 level, 8 level, 16 level or more complex signaling is easily achieved. In addition, since the arbitrary waveform generator can generate different signals by downloading different digital waveforms, it provides the ultimate capability for de-embedding channel characteristics by pre-distorting the downloaded waveforms. One limitation is memory capability and therefore AWG use is limited to short play times or short data patterns, which is appropriate for early research activities.

Figure 12 shows an M8050A 120 GBd BERT pattern generator, capable of generating both real time and memory based patterns in NRZ, PAM4, PAM6, or PAM8 format from 2 to 120 GBd. Other features include 7-tap de-emphasis, ISI generation and comprehensive jitter generation impairments. BERTs are more suited to generating large patterns (e.g. PRBS31 and higher order) and to applications requiring a complex jitter stress cocktail and error analysis such as receiver testing.

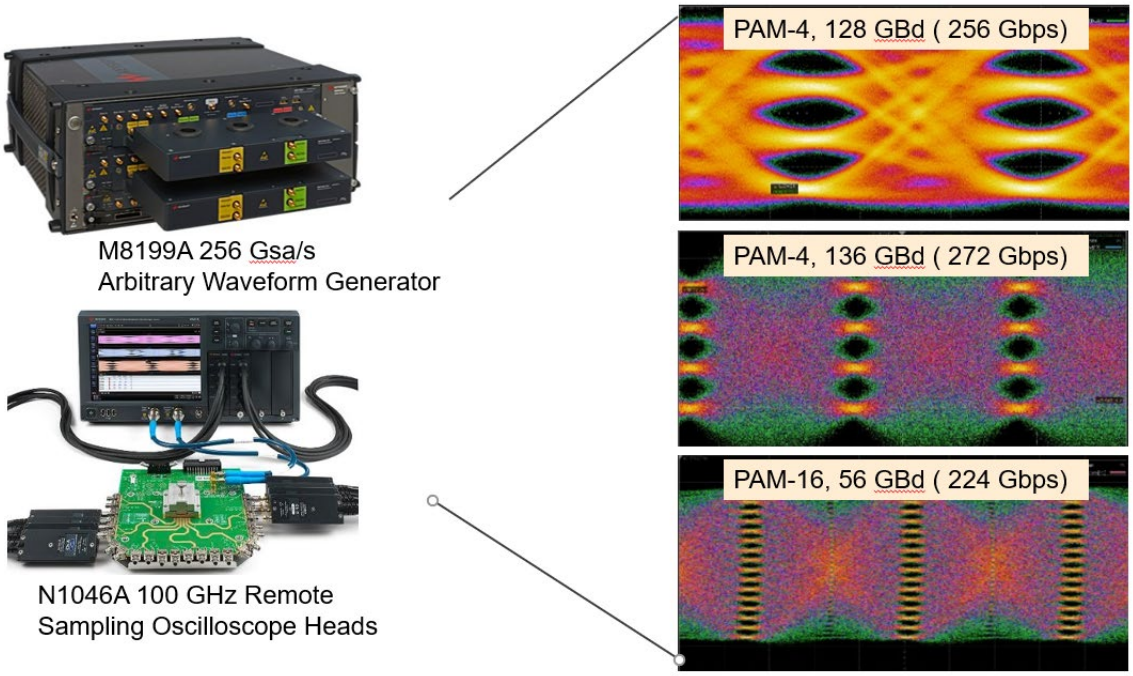


Figure 11. M8199A AWG generation of electrical signals above 224 Gbps

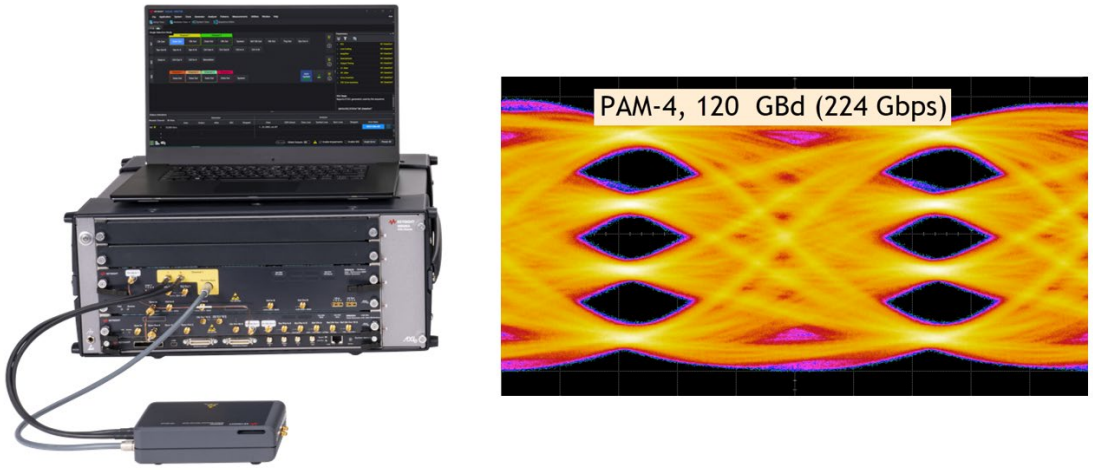


Figure 12. M8050A BERT generation of electrical signals up to 224 Gbps

Component test

In the development and evaluation of ultra-high-speed chips, interconnects and optoelectronic devices, the frequency domain parameters (bandwidth, gain, flatness, return loss etc.) are the most basic indicators to measure device performance. One proven method for optimizing interconnect design for these high-speed digital applications is to analyze scattering parameters (s-parameters). The faster the chipset and shorter the interconnect, the more the drive for 120GHz s-parameter measurements becomes.

The flagship Vector Network Analyzer from Keysight is the N5291A and consists of 4-ports for differential channel characterization up to 120 GHz. Reducing reflections can certainly be done by minimizing return loss at higher frequencies, but multi-domain analysis is now critical to developing a well-controlled impedance environment throughout the complete channel. Therefore, using specialist software such as Physical Layer Test System (PLTS) in conjunction with the 120 GHz VNA can produce insightful information such as Time Domain Reflectometry (TDR) for small geometries, eye diagram for PAM-4 modulation schemes and automatic equalization tap selection for powerful engineering research and development.

As shown in the configuration in Figure 13, the N5291A VNA offers superior accuracy and uncertainty for interconnects employed in today's network and data centers. A typical test template will incorporate many domains of data analytics including frequency, time, eye diagram, RLCG, mode conversion as well as pre-emphasis and equalization simulation in a user-friendly format. This will enable complete channel optimization with reduced crosstalk and extended bandwidth to guarantee the best interconnect performance possible for high performance ethernet systems.

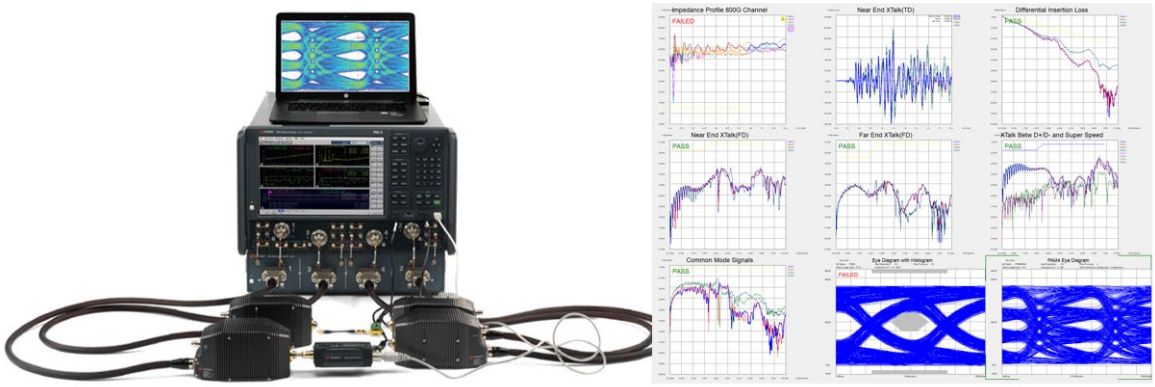


Figure 13. N5291A PNA Millimeter-wave System 900 Hz to 120 GHz, and Physical Layer Test System software

A particular challenge in the optical domain is with components such as the optical modulator and photodetector. The performance of the modulator will affect the quality of the signal converted from electrical input to laser output of the transmitter; and the performance of the photo detector will affect the quality of the receiver optical input converted back to an electrical signal.

The current state of the art modulators based on thin-film lithium niobate or silicon optical micro-ring modulators can achieve bandwidths exceeding 60 GHz, while photo detector bandwidths can be even higher. To achieve mature commercial use, these devices must become more consistent and stable. To perform detailed verification and evaluation of the characteristics of these devices over the entire frequency range (including the frequency roll-off characteristics outside the specified bandwidth), a lightwave component analyzer (LCA) with frequency range above 80 GHz is required.

Figure 14 shows the N4372E Lightwave Component Analyzer System with frequency range of 110 GHz. The system consists of a vector network analyzer and millimeter wave extension head to generate and receive electrical signals up to 110GHz, plus the addition of optical transmit and receive test heads and controller. Optical device test software takes care of electrical and optical calibration procedures and device testing and has the flexibility test optical-optical devices (optical fibers, optical amplifiers), electrical-optical devices (lasers, modulators), optical-electrical devices (photodetectors) and even electrical-electrical devices (amplifiers, drivers).

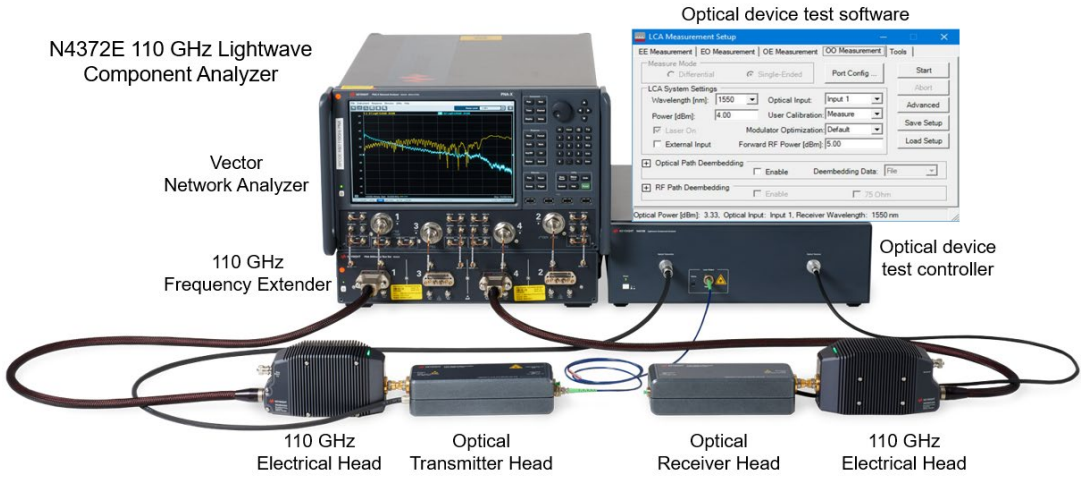


Figure 14. 110 GHz Lightwave Component Analyzer

Waveform test

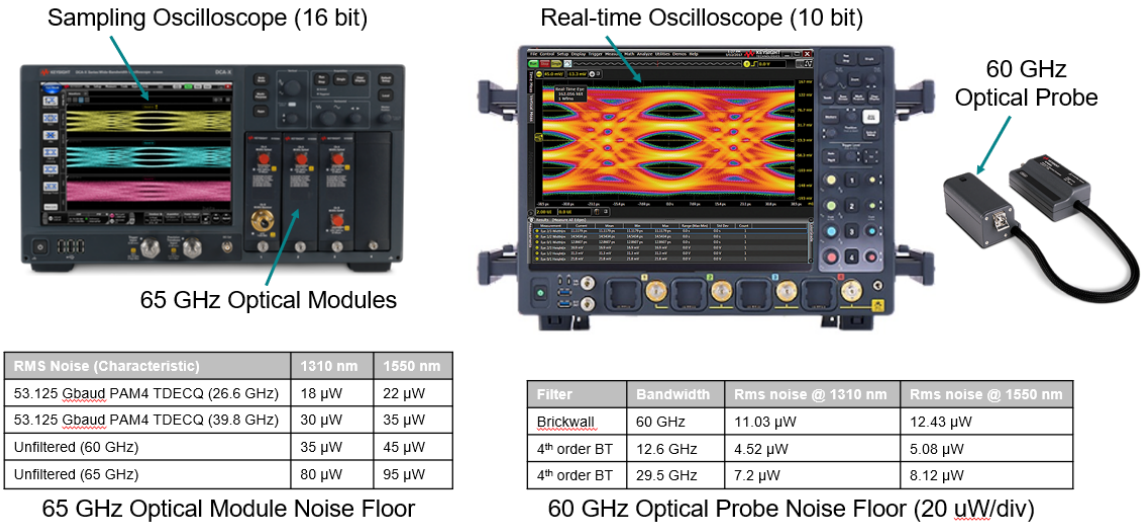
The next step in evaluation is assessing time domain performance of devices, stimulated with digital signals. These tests include eye diagram analysis, transition times, jitter, TDECQ for optical signals, VEC for electrical signals etc.

Waveform testing of high-speed signals can be accomplished using either a high-speed sampling oscilloscope or high-bandwidth real-time oscilloscope. Keysight DCA-X Sampling Oscilloscopes come with a range of both electrical modules up to 100 GHz bandwidth and optical modules to 65 GHz bandwidth. Keysight's UXR real-time oscilloscope has up to 110 GHz electrical bandwidth and the option to add an optical-electrical probe up to 60 GHz optical bandwidth.

Regardless of which oscilloscope is used, the key measurements are available on either platform, with optical measurements on the UXR leveraging the industry standard methods from the DCA-X FlexDCA application. The main differences between the platforms are in clock recovery (SW or HW), frequency response roll-off, noise floor and triggering methods.

Figure 15 shows two typical high-bandwidth optical signal measurement solutions. Both have adequate bandwidth for 224 Gbps research. If the 224 Gbps standards adopt PAM4 modulation (112GBd), and if as per previous PAM4 optical signal standards, the required reference receiver bandwidth is to be 1/2 of the signal baud rate (56 GHz in this case) then the instruments' 60/65GHz optical bandwidth can cover this need.

The DCA-X sampling oscilloscope's high bandwidth, low noise, and very mature optical path calibration technology has established it as the industry standard for optical signal testing. Traditionally, real-time oscilloscopes have been used in the electrical domain only, however with the development of real-time oscilloscopes with up to 110 GHz bandwidth, 10-bit ADC sampling, and optical probes up to 60GHz, optical testing is now possible. The frequency response characteristics of real-time oscilloscopes are typically flat with a brick wall roll-off response. High-speed digital signal testing usually requires the oscilloscope to emulate a 4th order Bessel-Thomson filter for device measurements, which generally means side-by-side a higher bandwidth real-time oscilloscope vs sampling scope is required in order to create the filter response.



RMS Noise (Characteristic)	1310 nm	1550 nm
53.125 Gbaud PAM4 TDECQ (26.6 GHz)	18 µW	22 µW
53.125 Gbaud PAM4 TDECQ (39.8 GHz)	30 µW	35 µW
Unfiltered (60 GHz)	35 µW	45 µW
Unfiltered (65 GHz)	80 µW	95 µW

65 GHz Optical Module Noise Floor

Filter	Bandwidth	Rms noise @ 1310 nm	Rms noise @ 1550 nm
Brickwall	60 GHz	11.03 µW	12.43 µW
4th order BT	12.6 GHz	4.52 µW	5.08 µW
4th order BT	29.5 GHz	7.2 µW	8.12 µW

60 GHz Optical Probe Noise Floor (20 uW/div)

Figure 15. Optical signal measurement scheme with bandwidth above 60GHz

Another consideration is noise. Sampling oscilloscopes have an excellent and constant noise floor performance, whereas the noise floor of the real-time oscilloscope depends on the input voltage range. In some situations, with low amplitude signals the UXR series real-time oscilloscope may have a lower noise floor than the DCA-X sampling oscilloscope as shown in the tables in Figure 13.

Bit error ratio test

The remaining step in high-speed data link testing is a bit error ratio (BER) or symbol error ratio (SER) measurement to validate the quality of the link. There are essentially three available methods for measuring BER of high-speed digital components and links:

The traditional method is to use a bit error rate tester (BERT) such as the Keysight M8040A 64 GBd or M8050A 120 GBd High Performance BERT. The BERT can generate the test signal and send to the device, and then the signal is looped back through the transmit path to the BERT analyzer for BER measurement and error analysis. This method is good if the loop back path does not introduce more errors into the signal. With PAM-n signaling and higher baud rates, no transmission is error free, and the loop back path starts to impact the BER measurement. For 112 Gbps this approach can still be used in some cases.

Alternatively, the BER may be read directly from the device under test as in many cases the device has an internal error counter to facilitate BER testing. This is becoming increasingly common, but in the early stages of new chip development, the internal error counter functionality may not be available.

The third method is to capture the digital signal with a high-speed real-time oscilloscope and determine BER through software decoding and compare with the expected pattern. Limitations of this approach are available capture memory and time and trying to achieve BER of e.g. 1E-12 is simply not practical. However, with the introduction of PAM4 signaling and FEC, the raw BER targets of 112 Gbps links are in the range of 2.4E-4 (optical) to 1E-5 or 1E-6 (electrical) making the real-time oscilloscope BER measurement viable. Target uncorrected BER for 224 Gbps is expected to be even higher e.g. 1E-3 for

an optical link. An additional feature of the real-time oscilloscope as an error detector is the ability to capture and analyze the bit stream waveform, including FEC signaling, and correlate bit errors positions in the signal waveform.

An example of this method is the Keysight Infiniium UXR-series oscilloscope, available with up to 110 GHz bandwidth, configured as a real-time error detector (RED) with the M8070B BERT control software. The oscilloscope captures and converts the incoming signal into digital form and these bits are sent over USB/LAN to a controlling PC running M8070B, the BERT control software. The M8070B then compares this data against the expected pattern to generate a BER measurement. All necessary setup and control of the real-time oscilloscope is also handled by the M8070B software.

The pattern generator which can either be BERT PG (e.g. M8050A) or AWG (e.g. M8199A) adds flexibility in terms of transmitter equalization (de-emphasis capabilities, raised cosine filtering, error insertion etc.) and the real-time oscilloscope offers high sensitivity, flexible clock recovery and enhanced automated receiver equalization capabilities such as FFE, DFE, CTLE. This solution can handle multiple modulation formats such as NRZ, PAM4, PAM6, PAM8 and at data rates up to 240 Gbps (120 GBd).

Figure 16 shows an example of PAM4 bit error rate test at 224 Gbps with the M8050A 120 GBd BERT pattern generator and UXR real-time oscilloscope as the error detection engine. The table illustrates a few BER measurements and confidence levels for different target BER limits; in these examples 5-tap FFE was applied at the analyzer before BER was measured. In the example of 1E-5 target BER, the measurement of 5.2E+5 bits was realized in about 1.2 seconds, and the confidence level of the result exceeds 99%. As the complexity of the equalization increases, the test time required to achieve the corresponding BER confidence level will also increase. For practical purposes the real-time error detector test solution based on the UXR oscilloscope can support bit error rate measurements down to a bit error ratio of 1E-6.

Target BER	Measurement Time (seconds)	Number of bits received	BER Confidence Level
1E-4	0.39	2.62E5	99.9999 %
1E-5	1.19	5.24E5	99.46 %
1E-6	7.53	3.14E6	95.68 %
1E-7	75.03	3.01E7	95.08 %

Figure 16. PAM4 bit error rate test with real-time oscilloscope

Test solutions summary

Figure 17 summarizes the 800G Ethernet module or chip interfaces and test equipment solutions. Keysight Technologies has 112 Gbps compliance test solutions available now, based on the DCA-X sampling scope, UXR real-time scope and M8040A bit error rate tester for both electrical and optical interfaces. The OIF-CEI and IEEE 802.3 standards for 112 Gbps are still in development and the Keysight compliance test solutions track that progress to provide timely test solutions throughout the standards development cycle.

Research on 224G optical ports and electrical ports is at an early stage and has yet to achieve consensus within the industry regarding the technical implementation details. Keysight Technologies is the leading supplier of test equipment used in 224G research with the M8050A 120 GBd BERT and M8199A high-speed arbitrary waveform generator for generating 224 Gbps electrical PAM-n signals (as well as other modulation schemes such as DMT and partial-response PAM4), DCA-X and UXR oscilloscopes with up to 110 GHz electrical / 65 GHz optical bandwidth to measure 224 Gbps transmitter waveform quality and parameters.

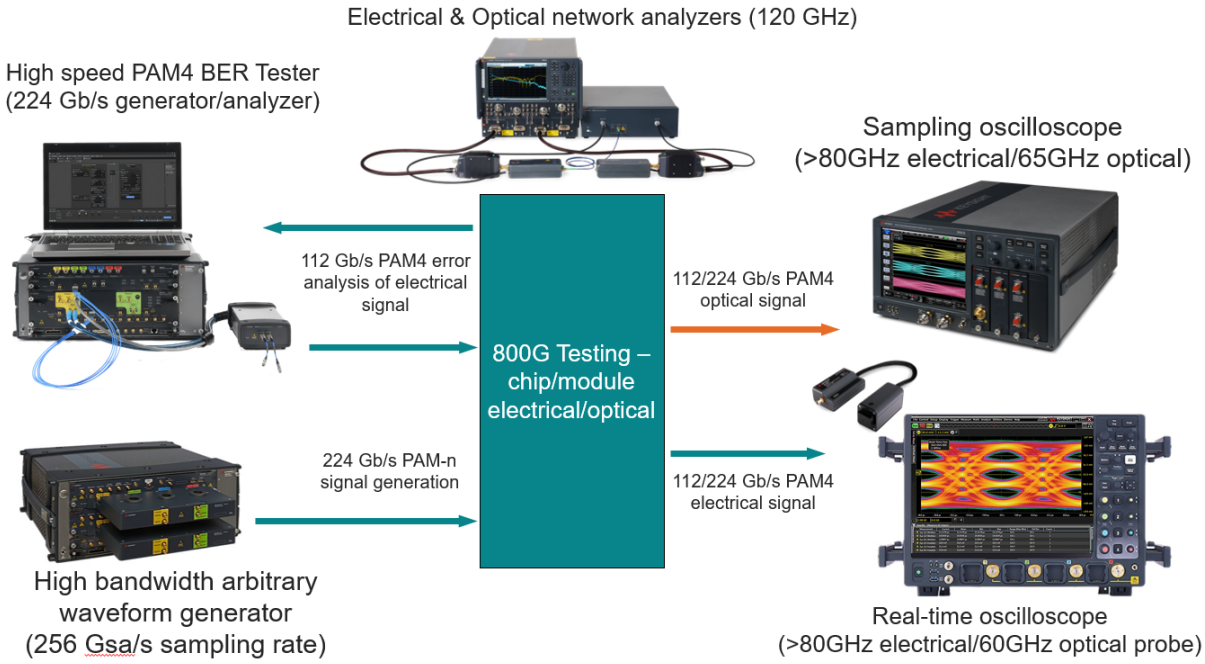


Figure 17. 800G equipment/optical module/chip test plan

For performance evaluation of key components such as integrated circuits, interconnects, optical modulators, and detectors used for 224 Gbps, the electrical PNA and lightwave component analyzers with frequency ranges up to 120 GHz can be used, and for bit error ratio testing on electrical or optical ports (mainly for key optoelectronic devices in TOSA or ROSA), a real-time oscilloscope based error analyzer solution is available.

For more information at www.keysight.com/find/800G