400GE is Revolutionary, Not Evolutionary

OVERCOME 3 KEY CHALLENGES IN TRANSCEIVER UPGRADES

KEYSIGHT TECHNOLOGIES

INTRODUCTION

Exciting New Applications Require New Data Center Speeds

Emerging technologies such as fifth-generation wireless (5G), artificial intelligence (AI), Virtual Reality (VR), Internet of Things (IoT), and autonomous vehicles will generate vast amounts of data in the network, creating new computing, storage and performance demands in the data center. Data center operators need to embrace new technologies to support the response times and high bandwidth that these technologies will require.

With the expectation of billions of devices to be connected to the internet, and the data-intensive real-time applications they will run, 100 Gigabit Ethernet (GE) speeds which are common in data centers today will not be fast enough. In addition to planning the size and location of data centers, as well as considering a shift to virtualized network architectures, data center operators are looking to evolve the speed of their networks from 100GE to 400GE.

We can help you accelerate your data center transceiver innovations.

Contents

400GE is Revolutionary, Not Evolutionary

OVERCOME 3 KEY CHALLENGES IN TRANSCEIVER UPGRADES

CHALLENGE 1

Increase Channel Capacity

CHALLENGE 1
Increase Channel Capacity

At the core of every network upgrade is the transceiver, and at the core of every transceiver is channel capacity. Traditionally, data center operators tended to upgrade their network architecture every couple of years. However, many data centers are currently at maximum capacity, and data center operators need to find a way to increase channel bandwidth to reach 400GE speeds. Reducing the power per bit is equally important. Data center operators are turning to transceiver manufacturers to move to the next-generation speed class.

THE BASICS

The move from 100GE to 400GE in the data center is revolutionary, not evolutionary. Optical transceivers will need to use advanced signal modulation and coding to reach 400GE speeds. These techniques create new test challenges for transceiver manufacturers.

Channel Capacity Theory

According to the Shannon-Hartley theorem, there is a theoretical maximum amount of error-free data over a specified channel bandwidth in the presence of noise. Either the channel bandwidth or the number of signal levels must be increased to improve the data rate or channel capacity. Non-return-to-zero (NRZ) and four-level pulse amplitude modulation (PAM4) are two modulation technologies that can enable 400GE.

NRZ is the most common signal modulation scheme for 100GE today. It is a two-state transmission system (also referred to as two-level pulse amplitude modulation or PAM2) where a positive voltage represents a logical “1”, and an equivalent (generally) negative voltage represents “0”. 100GE requires four lanes of 25 gigabits per second (Gb/s) NRZ modulated signals. Since NRZ has gradually evolved over the last 50 years, with improved speeds from 110 bits/s to 100 Gb/s, many new concepts and challenges have already been researched and addressed. By applying these same concepts, using eight lanes of 56 Gb/s NRZ signaling to move to 400GE is a logical evolution. However, as speeds of NRZ designs increase above 28 Gb/s, channel loss becomes a limiting factor.

PAM4 signals use four amplitude levels with logical bits 00, 01, 10, and 11 to represent a symbol. The number of symbols transmitted per second (baud rate) is half the number of bits sent per second. For example, a data rate of 28 gigabaud (GBaud) means there are 56 gigabits of data transmitted per second. This is double the data rate (throughput) in the same bandwidth compared to 28 GBaud NRZ, which is virtually 28 Gb/s, since one bit represents one symbol.

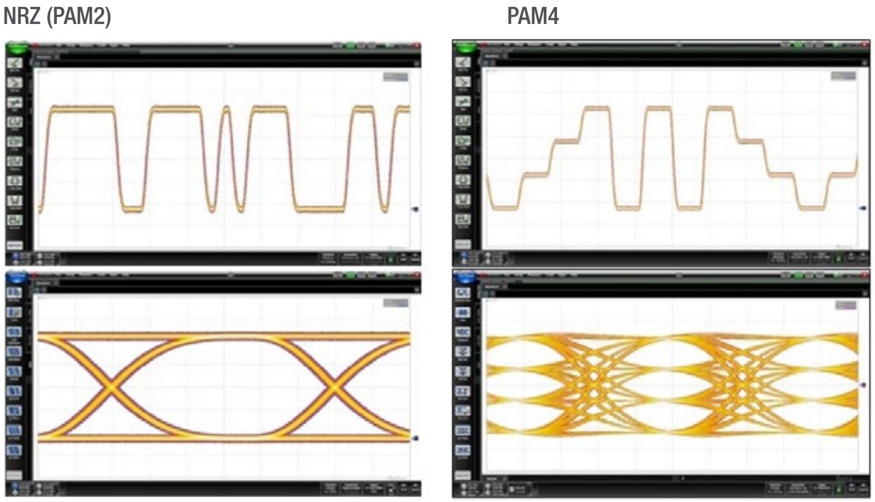


Figure 1. NRZ vs PAM4 modulated signals

THE SOLUTION: ADVANCED MODULATION AND CODING

Next-generation optical transceivers need to use revolutionary advanced modulation and coding techniques to reach 400GE. PAM4 modulation is the recommended method to move to 400GE speeds in the data center. However, PAM4 designs are far more susceptible to noise since an amplitude swing of two represents four signal levels. As a result, the signal-to-noise ratio (SNR) is lower, and analyzing noise in transceiver designs becomes a critical test factor. PAM4 will use forward error correction (FEC) to account for this.

FEC is an advanced coding technique that sends the required information to correct errors through the link along with the payload data. The decoder uses this information to recover corrupted data without the need to request the transmitter to retransmit it. Both the transmitting and receiving ends of the link must know which coding scheme is being used for the link to operate. Links employing FEC use a variety of coding systems. The more common coding schemes used in data center networking are variants of the Reed Solomon (RS) system, initially developed in the 1960s by Irving Reed and Gustav Solomon for use in satellite data links. FEC introduces new challenges in the physical layer testing of PAM4 signals. With 400GE, naturally occurring errors in the system are acceptable to a certain level and then corrected with FEC, resulting in a nearly error-free environment post-FEC.

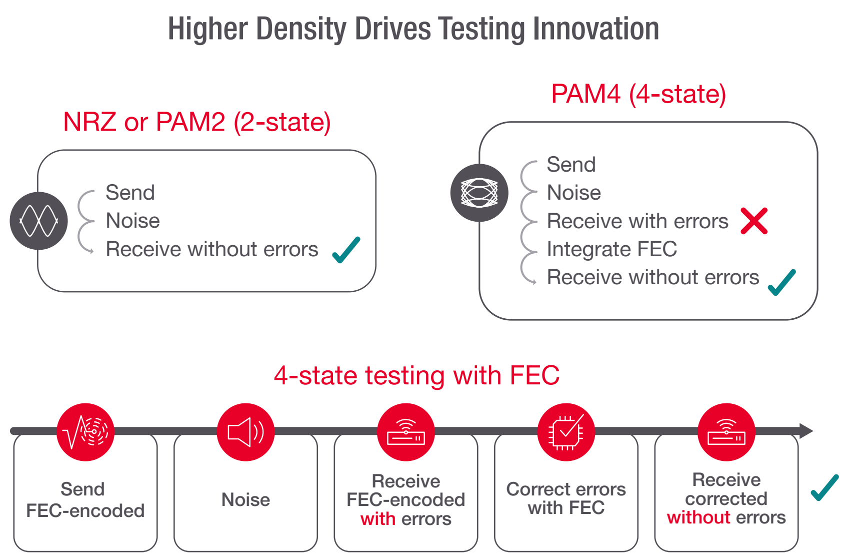


Figure 2. Testing with and without FEC

3 Main Considerations When Testing FEC Encoded Signals

1. Coding Gain

The encoding process converts the payload data to a format that allows decoding and creates the additional data required to correct errors. Code words are the result of the encoded data. Decoding is needed on the receiving end to recover the data. Coding gain is a figure of the robustness of the error correction code. A higher coding gain allows the correction of a higher number of errors. However, there are tradeoffs. RS systems using a higher coding gain need to send more overhead in the block of code words to facilitate decoding at the receiving end. Also, increasing the coding gain increases the amount of logic needed for coding and decoding, and the processing time, or latency, required to encode and decode the data. FEC with higher coding gain is required for high-speed serial data links using PAM4, which have a higher native error rate than those using NRZ line coding.

2. Burst Errors

FEC works on the assumption that the error distribution in the link is approximately random. A large burst of errors that exceed the number of correctable errors in the frame will result in frame losses, even if the average error rate in the link is better than the specified native BER. Note that a “burst” in this context is not necessarily consecutive bits. The errors could be interspersed with correct data bits, and would still result in a frame loss if they exceed the maximum number of correctable bits for the FEC code. Error bursts are unavoidable and unpredictable. They can originate at the receiver end of the link, or anywhere within the channel. Data striping can be used to minimize the impact of them.

3. Striping

Data striping is often used to lower the incidents of frame losses in links employing multiple lanes operating at a sub-rate of the total link data rate. Striping the data rotates the individual data streams through all the available lanes in the link in a round-robin fashion. By striping, burst errors generated by pass-through re-timers within the link will have the length of an error burst effectively divided by the number of lanes used for the striping. For example, in a 100GE link using four lanes of 25.78 Gb/s NRZ data, an error burst of 100 bits generated in an optical module on a single lane would result in only 25 errors on that single lane using striping. While striping does not increase the computed coding gain, which assumes random error distribution, it effectively increases the gain when error bursts occur.

CHALLENGE 2

Ensure Quality and Interoperability

CHALLENGE 2

Ensure Quality and Interoperability

Data center operators want to make the transition to 400GE as cost-effectively as possible while ensuring interoperability of devices from different vendors. The nature of pluggable modules necessitates that any new transceiver technology must be thoroughly tested to comply with specifications to ensure seamless compatibility before insertion into the network. Optical transceiver manufacturers must test to ensure their transceivers have strict compliance with defined specifications and are interoperable with other network components and transceivers from different vendors. Network downtime due to faulty transceivers is not an option for data center operators who have guaranteed service level agreements (SLAs) with users.

THE BASICS

In simple terms, an optical communication system consists of a transmitter, an optical fiber channel, and a receiver. High-speed electrical data links then interconnect the optical modules with the switching or networking chips. The transmitter of one transceiver sends data over optical fiber to the receiver of another in a different part of the network.

3 Main Concerns of Data Center Operators

1. Cost

400GE must offer a cost advantage over four independent 100GE interfaces. There are different ways that data center operators can reach 400 Gb/s. For example, they can add more servers while using existing storage density and 100GE transceiver technology. Many data centers are in remote locations where land is inexpensive and plentiful. Expanding the data center footprint to increase capacity is a real option for increasing total storage and network throughput.

Figure 3 shows a Google data center in the Netherlands with plenty of surrounding land for expansion. But expanding the data center shell footprint has other costs, including time to deployment and long-term capital depreciation, making network upgrades a more attractive option. As a result, data center operators are relying upon transceiver manufacturers to provide a cost-effective solution to reach 400 Gb/s speeds.

Figure 3. Aerial view of Google data center near Delfzijl in the province of Groningen, Netherlands.

2. Footprint

In the past, the only option to increase bandwidth was to increase footprint—network technology beyond 100 Gb/s was not available. Today, a more compelling alternative is emerging: 400GE. Currently, many data centers are maxed out in capacity, with no room for expansion. Therefore, transceiver manufacturers must address the need for higher performance interfaces. 400GE transceivers meet that need, providing four times the speed of 100GE while maintaining approximately the same footprint. Advanced modulation and coding techniques, such PAM4, will be used to increase channel bandwidth. More complex transceiver circuitry is required to reach 400 Gb/s speeds without expanding the footprint of transceivers, significantly increasing test complexity, as well as test time.

3. Power

Huawei Technologies researcher, Anders Andrae, recently estimated that power usage by the Information and Communication Industry (ICT) could reach up to 5,860 terawatt hours (TWh) per year by 2025[1]. Data centers will account for more than 30% of that usage. Put another way, at current growth rates, 2025 data centers are on track to use 8% of the world's total power. Power is a premium resource for data center operators, and transceiver manufacturers must do their part to ensure that next-generation 400GE transceivers deliver optimal power efficiency.

[1] Andrae, Anders. (2017). Total Consumer Power Consumption Forecast: https://www.researchgate.net/publication/320225452_Total_Consumer_Power_Consumption_ Forecast

THE SOLUTION: CHARACTERIZATION AND COMPLIANCE TEST

Several standards organizations, such as the Institute of Electrical and Electronics Engineers (IEEE), International Committee for Information Technology Standards (INCITS) and the Optical Internetworking Forum (OIF), govern optical transceiver specifications and define test procedures to ensure compliance to standards and interoperability with other vendors. However, standards organizations do not address packaging and connectors as part of their standards work. A multi-source agreement (MSA) among multiple equipment manufacturers defines the form factor and electrical interface. The MSA guarantees compliant devices will function properly and allow for seamless interoperability with devices from different vendors.

There are different sets of optical and electrical tests are needed for transmitters and receivers, and the tests must also include channel quality impairments. As systems become faster and more complex, it is also necessary to test receivers using non-ideal signals with different types of impairments. This is commonly referred to as a stressed signal. Since performance levels usually vary significantly from one generation to the next, standards and components often evolve together.

TRANSMITTER & RECEIVER TESTING

As data rates increase, so do the challenges of device design, validation, and test. Errors are introduced into the signal in many ways—from the transmitter, through the channel, and to the receiver. The optical and electrical links are not expected to have raw, error-free performance. Due to the reduced SNR, PAM4 links are expected to have errors due to noise and require forward error correction (FEC). Standards organizations are actively discussing several FEC considerations for 400GE testing. In addition to noise susceptibility, another increasingly challenging issue is testing probe or fixture interference. Test fixtures and cables have not traditionally been a substantial source of interference. However, at higher data rates, signal degradation occurs even in high-quality cables, and noise and interference become significant factors. Device test interference can arise from the simple insertion of measurement instruments. In this case, it is vital to de-embed the noise introduced by the test equipment from the signal. De-embedding is a post-measurement process that mathematically removes the noise characteristics of test fixtures from the overall measurement. Test equipment with built-in de-embedding will simplify transceiver characterization considerably.

CHALLENGE 3

Test Faster, Reduce Cost

CHALLENGE 3

Test Faster, Reduce Cost

Massive growth in data center traffic is driving the need for bandwidth upgrades in the data center. Data center operators require cost-effective next-generation optical transceivers to support their migration from 100GE to 400GE. Large hyperscale data centers have more than 50,000 optical fibers in them. With a transceiver at each end of the fiber, these data centers house upwards of 100,000 transceivers. Keeping the cost of the optical transceivers low is a high priority for data center operators. To be competitive, transceiver manufacturers must find ways to drive down production costs.

THE BASICS

Like most new technologies, the price of new optical transceivers tends to drop sharply after introduction to the market. Development costs amortize over time as volume ramps. Next-generation transceiver technology, such as 400GE, will likely reach mature pricing within a year of introduction. At maturity, the cost of transceivers is directly proportional to the complexity of the design and the number of optical components.

THE SOLUTION: TRANSCEIVER TEST EFFICIENCY

The data center transceiver market is extremely cost sensitive and competitive. Test time contributes significantly to overall transceiver cost. More efficient testing of the broad range of transceiver data rates accelerates innovation and lowers cost.

At each phase of the product development lifecycle, there are test solutions that can be used to maximize test efficiency. These tools can shorten design cycles, dramatically improve productivity, ensure quality, and significantly reduce costs. The process begins in the research and development phase with the design of transceiver components, continues to validation test of the transceivers, and then through manufacturing test.

3 Main Development Phases To Maximize Transceiver Test Efficiency

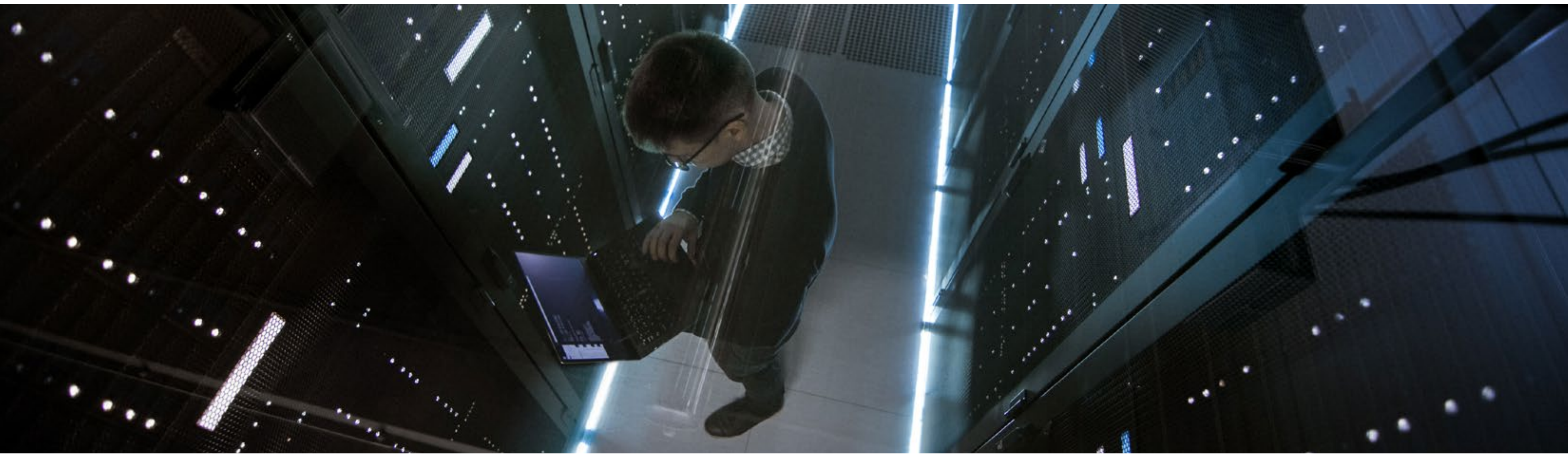
1. Design and Simulation

In the research and development (R&D) phase, starting with powerful design simulation software is the first step to ensure test efficiency and lower the cost of test. Today's design and simulation software enables transceiver designers to optimize their designs, ensure performance and robustness, and avoid costly additional board design cycles. They can identify the most sensitive design components early on and decide how to set specifications to improve manufacturing yield. Once optimized, designers can test design performance using post-processing data analysis functionality without rerunning simulations.

2. Validation Test

As new 400GE transceiver designs transition from simulation to first prototype hardware, engineers face the challenging task of developing a thorough, yet efficient, test plan. The operating margins in 400GE optical links are the tightest of any generation thus far. 400GE designs create additional test challenges, as small measurement errors can quickly consume the entire operating margin. Fortunately, the methods that describe how to characterize 400GE designs are becoming more stable, and engineers can review and follow the guidelines outlined in the standards when developing their test plan to characterize their transceivers.

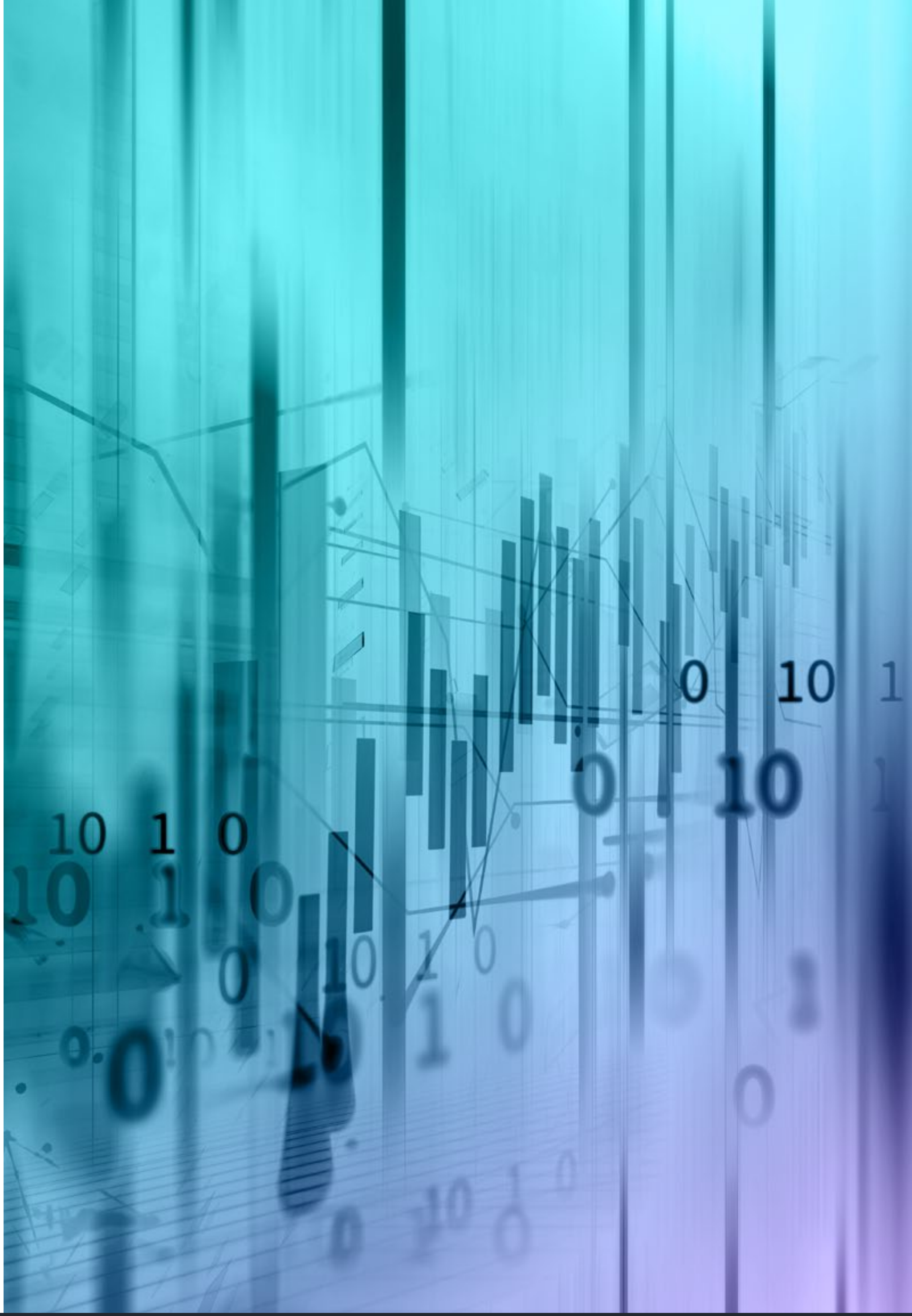
Test automation software can reduce test time down from hours to minutes. It is essential to choose automated compliance test software that is verified to test to the exact specifications of each technology standard. Test automation software guides the test engineer through setup, calibration, and compliance measurements, and allows them to quickly run through test cases without being an expert on test procedures—saving hours of test time. By using compliance test applications, transceiver manufacturers can ensure that a test that passes in their lab will pass in their customers' labs as well. More importantly, ensuring transceivers are compliant to standards will minimize the risk of interoperability issues with network switches and routers once millions of them are installed in data centers around the world.



3. Manufacturing Test

Research and development of 400GE transceivers is well underway. Engineers are still struggling with how to test PAM4 modules, and 400GE standards continue to evolve. Once 400GE transceivers reach the manufacturing phase, any issues found will mean a costly rework of designs. Designing for manufacturing is ideal.

Several tools can help transceiver manufacturers create manufacturing-friendly designs. For example, by using a common software platform with built-in data analytics from initial design concept through manufacturing and deployment, it is possible to accelerate the overall product development workflow and minimize errors found during manufacturing. Costs increase significantly as issues are uncovered later in the development process. It is ideal to discover all problems early in the design phase.



SUMMARY

Get Ready for 400GE

100GE is widely deployed in data centers around the world today, but 400GE links will be the next step to increase network bandwidth for 5G-capable data centers. Data center operators can ensure the seamless transition from 100GE to 400GE by introducing next-generation transceivers that increase channel capacity, guarantee quality and interoperability, and reduce test time and cost. Addressing these challenges will enable 400GE in data centers to soon become a reality.

With 400GE transceiver technology squarely on the horizon, data center operators are also looking for new ways to design and operate their networks to withstand the kind of traffic that billions of devices will generate. Many are shifting to virtualized networks using software-defined networking (SDN) and network functions virtualization (NFV). SDN is a network architecture that enables software programmable network control of a virtualized network infrastructure. Network functions virtualization is an architectural concept that automates entire classes of network node functions into building blocks that can be connected to create communication services. Once data center operators shift to a virtualized network, they need to make sure that data flows through it as expected. Complete network test of Layers 2–7, including SDN/NFV validation and traffic loading, becomes the next hurdle to overcome.

FOR MORE INFORMATION

To find out how Keysight's solutions can help you address your 400GE data center implementation challenges, check the following links:

- To accelerate your data center infrastructure innovations, check out our **Data Center Infrastructure Solutions**
- To understand the three key challenges of testing 400GE transceivers, check out our **Data Center Transceiver Test Solutions**