Keysight Technologies
Characterizing and Verifying Compliance of 100Gb Ethernet Components and Systems

Application Brief
Overview

The expansion in Ethernet data bandwidth from 10Gb/s through 40G to 100G (moving to 400G and 1 Tb/s in the future) is a direct result in the rapid growth of capacity in modern data centers. Their rapid growth, resulting from acceptance of streaming video media, social networking, and enterprise outsourcing of data storage from internal enterprise servers to the cloud, is expected to continue through at least two decades. Efficient use of increased data storage demands higher capacity “pipes” within the data center network infrastructure, as well as between chips, modules and cards within the servers, switches and routers themselves.

Current Ethernet deployment consists of links operating at 10 Gb/s, with complementary Fibre Channel running at 16G. The very recent release of a series of Standards supporting 100Gb Ethernet and 32G Fibre Channel enables component and system designers to implement next generation designs utilizing serial links at these speeds. Data center architects will specify equipment, copper and optical interconnects based on 100G links shortly.
Per-lane speed classes

Transmitting 100 gigabits per second through a single channel is not practical for high volume commercial deployment. Some method of multiplexing must be performed. Wavelength Division Multiplexing allows multiple serial data streams to pass through a single optical single mode fibre in long haul carrier applications, but is cost prohibitive with the tens of thousands of optical links needed in data centers. Instead, transceivers employing multiple multi-mode (MM) fibres are used. On the electrical side, these transceivers communicate with their host chips through multiple lanes using Non-Return to Zero (NRZ) signaling. More complex amplitude modulation schemes such as PAM-4, which pack multiple bit values per symbol, are being considered to enable next generation 400G and 1Tb links.

The first implementations of 40G and 100G optical links utilize per lane data rates of 10.3125 Gb/s – the exact rate of 10Gb Ethernet. Thus, 40GbE is implemented with 4 lanes of 10GbE, and 100G with 10 lanes. While matching the speed of 10GbE eliminates the need for “gearbox” ICs, practical implementation of 100G using 10 lanes is difficult from circuit board routing and pin count requirements in host chips. Clearly a faster per lane data rate was needed.

The Standards working groups came through with a per lane data rate of 25.78125 Gb/s. Not being an integer multiple of 10.3125 Gb/s, it does require gearshift ICs to interface with 10GbE, but being a perfect 2.5 multiplier, no idle fill characters need to be inserted or removed from the stream during the transition.

Figure 1. Illustration of 10 x 10G vs. 4 x 25G 100G configurations.
Designing with 25Gb lane rates

The migration from 10G to 25G per lane line rates presents significant challenges to the design engineer – beyond simply scaling down the Unit Interval (UI) time and corresponding jitter budgets by a factor of 2.5. Even with high levels of de-emphasis on the transmitter side, eyes are totally closed at the receiver inputs. Well balanced applications of both Continuous Time Linear Equalization (CTLE) and Decision Feedback Equalization (DFE) are often required on the receiver end. Retiming at every electrical input, including the Tx input driving the laser in optical transceivers is required. With total eye closure, the use of Forward Error Correction (FEC) is mandatory in order to realize links of reasonable data integrity. The raw physical link itself no longer supports a Bit Error Ratio (BER) of $10^{-12}$ or $10^{-15}$. Standards have lowered these values to $10^{-6}$ or even as low as $3.4 \times 10^{-3}$.

Many new measurement concepts and techniques are needed to characterize designs and validate compliance to Standards using 25G class line data rates. The rest of this Application Brief is dedicated to discussing these.

### Relevant Standards enabling 100G class data rates

<table>
<thead>
<tr>
<th>Application</th>
<th>Standard</th>
<th>Standards Body</th>
<th>Data Rate</th>
<th>Link Sub-Clause</th>
<th>Electrical PHY Sub-Clause</th>
<th>Per lane Signaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Networking (Extensions for optical networking and backplanes @ 40G &amp; 100G)</td>
<td>802.3ba</td>
<td>IEEE</td>
<td>40G, 100G</td>
<td>40G (4 x 10G): XLAUI, 100G (10 x 10G): CAUI-10, 100G (4 x 25G): CAUI-4</td>
<td>10.3125 Gb/s NRZ, 10.3125 Gb/s NRZ, 25.78125 Gb/s NRZ</td>
<td></td>
</tr>
<tr>
<td>Copper cable &lt;10 m</td>
<td>802.3bj, clause 92</td>
<td>IEEE</td>
<td>100G</td>
<td>100GBASE-CR4</td>
<td>25.78125 Gb/s NRZ</td>
<td></td>
</tr>
<tr>
<td>Backplane &lt;1 m</td>
<td>802.3bj, clause 93</td>
<td>IEEE</td>
<td>100G</td>
<td>100GBASE-KR4, 100GBASE-KP4</td>
<td>13.59375 Gb/s NRZ, PAM-4</td>
<td></td>
</tr>
<tr>
<td>Optical networking</td>
<td>802.3bm, clause 94</td>
<td>IEEE</td>
<td>40G, 100G</td>
<td>40GBASE-CR4, 40GBASE-SR4, 40GBASE-LR4, 40GBASE-ER4, 100GBASE-CR10, 100GBASE-SR10, 100GBASE-LR4, 100GBASE-ER4</td>
<td>10.3125 Gb/s NRZ, 10.3125 Gb/s NRZ, 25.78125 Gb/s NRZ</td>
<td></td>
</tr>
<tr>
<td>Optical networking &lt;100 m (MM) or &lt;10 km (SM)</td>
<td>Fibre Channel PI-6</td>
<td>T11-INCITS</td>
<td>32G</td>
<td>28.05 Gb/s NRZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical Board to Board &lt;686 mm + 2 connectors</td>
<td>CEI-25G-LR</td>
<td>OIF</td>
<td>19.9 - 25.8 Gsymbols/s</td>
<td>19.9 - 25.8 Gb/s NRZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical Board to Daughter Board &lt;500 mm + 1 connector</td>
<td>CEI-28G-MR</td>
<td>OIF</td>
<td>19.9 - 28.1 Gsymbols/s</td>
<td>19.9 - 28.1 Gb/s NRZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical Board to Daughter Board &lt;300 mm + 1 connector</td>
<td>CEI-28G-SR</td>
<td>OIF</td>
<td>19.6 - 28.1 Gsymbols/s</td>
<td>19.6 - 28.1 Gb/s NRZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical Chip to Module 10 dB loss (100 mm) + 1 connector</td>
<td>CEI-28G-VSR</td>
<td>OIF</td>
<td>19.6 - 28.1 Gsymbols/s</td>
<td>19.6 - 28.1 Gb/s NRZ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What needs to be tested?

As customary with high speed serial links, measurements of the transmitter (Tx) output waveform and receiver (Rx) susceptibility to signal impairment on its input are performed for both general characterization and Standards compliance. The jump from 10G to 25G has changed areas of focus in some of these.

Two instruments are generally required for these tests – a high speed sampling oscilloscope and a bit error ratio tester (BERT). The oscilloscope is used for measuring the transmitter output parameters, as well as calibrating the levels of “stress” impairments used for testing receiver inputs. The BERT is used for generating some special patterns to drive the transmitter for Tx output measurements, and verifying receiver operation during Rx stressed input tests. When testing inputs with a BERT, the receive signal must be retransmitted to the BERT error detector.

Figure 2. Illustration of generic scope Tx measurement and BERT Rx measurement.
When characterizing optical transceivers, both optical and electrical signals need to be generated and measured. Until recently, all data center applications used multi-mode (MM) devices exclusively. However, to achieve link reaches greater than 100 m, the latest 32G Fibre Channel and a new 100G-PSM4 MSA (Parallel Single Mode 4 lane, Multi Source Agreement) project, single mode variants have been included.

Optical signals are handled with appropriate optical to electrical (O/E) and electrical to optical (E/O) conversion. As these reside within the measurement signal path, their accuracy impacts measurement accuracy and repeatability. For transmitter measurements, optical sampling heads are generally used. The bandwidth required for 25.8 and 28G line rates can be a challenge for optical heads which accept multi-mode fibre, such as the Keysight 86105D option 281. Optical sampling heads and precision photo diode detectors used for 10G will need to be upgraded to handle the new speed class. At the higher data rates, it is preferable to use integrated optical sampling heads, which include the photo detector and electrical sampling circuits within the same plug-in, rather than a discrete O/E photo detector cabled to an electrical sampling plug-in. Even with the highest quality cables, the latter configuration is prone to reflections from small mismatches at the connector interfaces, which distort eye diagrams and measurement repeatability between instruments.

For testing optical receiver inputs in transceiver modules, a laser reference transmitter can be used. However, since wavelength multiplexing is not used in data center components, the laser from a second DUT can be used in conjunction with an MM optical attenuator with output power measuring capability.

![Figure 3. Illustration of test setup for optical transceiver (simplified – O/E & E/O with instruments only).](image-url)
Compliance test boards and de-embedding

When measuring an IC or optical transceiver module, a break out fixture is required to bring the signals from the device leads or pads to standard connectors used on test instruments. At the frequencies present in 100G devices, even the shortest practical lengths of circuit board trace on these boards have some signal loss. In order to assure repeatable measurements, the Standards working groups have defined Compliance Test Boards with tightly controlled trace loss and reflections for the purpose of device interfacing. Both Host Compliance Boards (HCB) and Module Compliance Boards (MCB) are specified. Because devices may have different physical pin layouts, these boards must be designed and fabricated by the users. As they are an important part of the measurement chain, high accuracy is essential. Be sure to verify the S-parameters of your compliance boards when they arrive from fabrication before performing device characterization or compliance measurements.

An important side note – the SMA or 3.5 mm connector which has traditionally been used in break out and compliance test boards for the last several years has resonant nodes within the frequency spectrum required for 25G and 28G devices. Thus, a higher performance connector such as a 2.4 mm, or 1.85 mm (Type ‘V’) will be required.

Measurements are made from and test signals are calibrated at specified test points on the compliance test boards. These are generally the point of cable connection to and from the instruments. These points can be thought of as a reference plane. The degrading effects of cable loss between the instruments and the connection points of the test boards must be compensated for using suitable de-embedding techniques.

Overview of testing methods - outputs

Device outputs are measured directly with a sampling oscilloscope. All optical output and some electrical measurements require a “Reference Receiver”. A reference receiver is simply a calibrated oscilloscope having tightly specified frequency roll off characteristics. As the Unit Interval (UI) and corresponding link jitter budgets are smaller at 25 and 28 Gb/s, the sampling scope must be equipped with a precision time base in order to perform accurate jitter measurements.

The standard amplitude and timing measurements of eye height and width are similar to those performed in lower speed Standards. Crosstalk coupling is a concern as it is more difficult to manage in higher data rate systems. Both the input receiver susceptibility tests and the transmitter output measurements now require crosstalk to be present on the back channel (receive path next to the transmitter) or alternate transmit output lanes in a four lane device. To assure the phase alignment of the transmitter is worst case, the clocking of the crosstalk signal must be asynchronous to the output signal being measured.

Output jitter must be measured as well, however some new specific measurements from the decomposition are required. As most devices supporting 100GbE or 32GFC use half rate clocks, the possibility of Even/Odd jitter results from non-symmetrical clock waveforms. Even/Odd jitter is seen when two eyes formed for consecutive bits or “double tracking” at the transitions in a single eye when averaging is applied to remove random jitter components. (Note it will not be seen in a single eye if the sampling oscilloscope is triggered with a sub rate clock that is a power of 2).
J2 and J9 jitter measurements are also required in some of the 100GbE Output compliance tests.

Due to the use of FEC, the BER of uncorrected links allowed to be much higher than in the lower speed class Standards such as 10GbE. Hence, measurements of total jitter do not need to be made at the deep BER levels as in the past. “Bathtub” plots are not required for either Standards compliance or device characterization. In the past, a BERT was necessary to obtain the required sample depth for the total jitter measurement. With 100GbE and 32GFC, a sampling scope provides the required sample depth for accurate measurements in a short amount of time.

The signal to noise ratio of the transmitter output signal is an important parameter in systems which employ Decision Feedback Equalization (DFE) in the receivers, and is now included in the list of output measurements for 100GbE and 32GFC Standards. The individual Standard specifies the exact method of measurement. Generally a noise measurement is made in the rails near the center of the eye, and compared with a standard histogram based eye height measurement.

### Input measurements

As with the lower speed class devices, the receiver behind a device’s input needs to be tested for its ability to handle impaired (“stressed”) input signals. The composition of stress types and relative amplitudes have changed considerably with the higher data rates, as a result of different topologies used in the devices themselves.

Three types of stressed input tests are performed independently. They are optical unstressed input sensitivity test, a PLL tracking test, and the Jitter or Interference tests which use higher frequency stress sources.
For optical inputs, a sensitivity test is performed. The data center Standards do not require the complicated stressed optical input tests used in the longer reached Ethernet Standards. Generally the laser from a second DUT can be used for a source, assuming it has sufficient power stability. The setup is shown below:

The Standards require a specified test pattern, provided by the BERT. Optical transceiver modules do not have internal error counting capability, so a conventional BERT electrical error analyzer is used to determine the power threshold where the device begins to experience errors.

The PLL tracking or Jitter tracking test verifies the ability of the phase locked loop (PLL) in the device clock recovery to track higher amplitude low frequency jitter components. This test has been isolated from the higher frequency stress test to avoid stressing devices with high frequency sinusoidal jitter.

The PLL Tracking test uses only Sinusoidal Jitter (SJ) as a stress component. The test set-up is the same as used with the higher frequency test, so Intersymbol Interference (ISI) is also present to exercise the CTLE in the receiver.
Depending on the device and the Standard, the input receiver is verified by using loop back into the input of a traditional BERT error analyzer, or with internal error counters in the device itself. For Standards compliance test, discrete tests are performed at either two or four spot SJ frequencies or amplitudes. Engineers characterizing device operating margins may elect to use a conventional Jitter Tolerance test (J-TOL) using swept amplitudes and stepped SJ frequencies.

The Receiver Stressed Input Test (RSIT) or high frequency jitter test uses higher frequency jitter components and/or interference (amplitude noise). In many aspects, this test has become more complicated for 100GbE and 32GFC.

100GbE and 32GFC links use both Tx pre-emphasis and Rx equalization, implemented with CTLE and often adding DFE. With the exception of backplanes which have variable channel loss (covered in 802.3bj described below), the Standards call out 9 preset values of equalization boost in the receiver. In use, the design engineer will program the initiation software with the appropriate equalization value determined by the design channel loss. The Standards require that the RSIT be performed with the input equalization set to the optimum value for maximum eye opening, and also one equalizer boost setting above and one setting value below the optimum.
To determine the optimum equalization setting, a scope set up as a reference receiver and programmable software CTLE measures the signal at the end of the test setup which drives the DUT input during the test. The scope equalizer boost setting is adjusted to get the widest eye opening. This value is noted and programmed into the input receiver under test.

A typical stress mix may include Random Jitter (RJ), Intersymbol Interference (ISI), Uncorrelated Bounded High Probability Jitter (UBHPJ or BUJ), and Random Interference (RI). The ISI is provided by a Frequency Dependent Attenuator, which is generally implemented as a circuit board trace. The loss characteristics are specified as S-parameters, which include both the circuit board loss, the insertion loss of the summation network used for the RI injection, and any cabling interconnecting these. It is imperative to carefully calibrate the entire set up with a Vector Network Analyzer to assure it is within the tight tolerance band allowed by the Standard.

Random Interference is injected after the ISI channel. This is a differential mode signal, but due to the wide band of spectral content required, it is generally injected only in one leg of the differential pair. The newer Standards do specify the frequency content – typically being spectrally flat from data rate divided by 5 to data rate divided by 2. A dedicated noise source is required to fulfill this requirement.

Channel operating margin based receiver stressed input tests

The inclusion of Decision Feedback Equalization (DFE) is now necessary in many receivers used in 100GbE applications. In all cases, a Continuous Time Linear Equalizer (CTLE) must be used in front of a receiver to open the eye enough for the clock recovery to function. The balance of equalization performed by the CTLE versus the DFE requires careful design consideration. The high frequency gain provided by the CTLE does not discriminate between data and noise. Applying too much gain in the CTLE results in the receiver being overly susceptible to channel noise. In many cases, the susceptibility to amplitude noise (vertical eye closure) dominates over the effects of timing jitter (horizontal eye closure.) Many of the 100G class Standards have incorporated an interference stressed input test.

The interference stressed input test verifies the receiver operates at the specified BER limit or better when being driven from a test transmitter which has the specified worst case signal to noise ratio. The signal to noise ratio is expressed as Channel Operating Margin (COM), in units of dB.

The basis of the test methodology is to measure the signal amplitude and native noise on the test pattern generator output, and pad it up as needed by adding random interference to the COM limit as seen at the interface test point on the compliance test board. The calibration process requires careful measurements to avoid under or over stressing the device under test.
The forward and return losses of both the test channel and RI injection path are both measured. Using an equation provided in the Standard, these determine the actual COM seen at the compliance board test point by the device under test. Factoring the loss of the injection channel, solving the equation determines the amplitude of the noise which must be applied at the input to pad the COM to the limit specified in the Standard. Band limited random interference with flat frequency content from data rate/5 to data rate /2 and rms amplitude computed by the user is added to the input of the injection channel. Once the COM is calibrated, the prescribed test pattern is applied and BER is measured at the receiver output.

Depending on the Standard, the COM based RSIT may also include timing jitter stress or just apply the RI as the only stress component. When timing jitter is also used, it is generally a mix of RJ and BUJ. Higher frequency SJ is not used in the stress mix, as DFE systems are very susceptible to high levels of periodic jitter which is not correlated to the data pattern.
Interactive link training in 802.3bj

IEEE 802.3bj applies to backplanes and copper cables. As the channel losses vary with the particular slot the card is positioned in, or the length of the cable, it is not practical to predetermine the proper equalization setting for the link during design. The link equalization is optimized during an interactive training process which takes place during power on initialization. The training process and the control protocol used for interactive communication between the transmitter and receiver are specified in the Standard.

Validating compliance for input susceptibility to stress first requires that the receiver complete a proper link training cycle with the test channel and transmitter. This would be impractical using a BERT or other instrument for pattern generation, as it would be required to act on commands from the test receiver and generate the proper response protocol in order to perform the link training. The Standard accommodates this by allowing the use of a second DUT as the test transmitter.

Two tests are performed to verify input compliance – a PLL tracking test and a COM based RSIT. The latter only uses RI as a stress component, which is injected at the far end (near the compliance test board input) of the ISI channel. The test method first verifies the output of the second DUT meets the compliance criteria for Tx outputs. Once validated to be compliant, the second DUT is configured to be the pattern generator. The PLL tracking test uses four static SJ frequencies and amplitudes. As most DUTs will retlime their clock inputs with a PLL, the low frequency of the SJ will be tracked and pass to the output of the DUT serving as the pattern generator. A standard communications grade synthesizer with FM capability can be used as a clock generator for the transmitter DUT. Calibration of the SJ amplitude is performed by direct measurement on the Tx output while the synthesizer FM modulation depth is adjusted. The modulation is turned off and the link is allowed to train. The modulation is then turned back on and transmitter is set to produce the prescribed test pattern of FEC encoded scrambled idle. The receiver behind the input under test has built in error detection and counting capability for the prescribed scrambled idle pattern, and measures its own BER. The value is read from the DUT and checked to be less than the limit specified by the Standard.

For the COM based RSIT, the process is repeated using the RSIT test set-up. A clean clock drives the transmitter. The calibration and test are performed using the COM test method described in the previous section.
Note on aggressor signals

With faster edge speeds, crosstalk susceptibility can be more prevalent at higher data rates. Most of the standards require aggressor signals be applied to the back channel and alternate forward lanes. To emulate real data, a long PRBS pattern such as PRBS-31 is usually required. The amplitude and transition time of the aggressor signals are set to the highest voltage and fastest transition time the Standard allows for the transmitter output. It is usually calibrated at the input connector interface point of the compliance test board.

Most of the Standards also stipulate that the aggressor signals be clocked at the same rate, but asynchronous with the data clock of the receiver being tested. For a meaningful test, it is imperative that asynchronous clocks be used for the aggressors. While all four lanes in a 100GbE transmitter will likely use the same clock, there is no assurance that the lane to lane skew in the test setup replicates the real life environment. The receiver under test is only susceptible to crosstalk induced errors when the coupled crosstalk edge is near the detector sample time, usually in the center of the eye. When the aggressor is clocked independently, there will be an inherent small difference in the clock frequencies, usually only a few parts per million. This difference will create a continuous slow phase slip between the signals, assuring that any crosstalk coupled from the aggressor transitions will at some point cross through the sample time of the receiver being tested.
Aggressor channel transition induces a voltage waveform on the victim channel.

Amplitude of induced crosstalk affected by:
- Aggressor source impedance
- Victim load impedance
- Physical separation
- Aggressor amplitude
- Aggressor rise time

The induced voltage closes the victim channel eye, but maybe not at the central sampling point.

Multiple aggressors further impair the victim channel.

Each aggressor impacts the victim channel differently.

But since all the aggressors are clocked with the victim transmit clock, the impairment does not occur during the receiver sample time window.

No crosstalk susceptibility is found.

Asynchronously clocking the aggressors guarantees the worst-case impairment of the victim channel because the effect is spread across the eye.

Multiple aggressors should each have a clock offset different than the other aggressors.

Frequency offsets should be low to allow enough “dwell time” in the sampling interval of the receiver to catch a possible victim pattern susceptibility as well.

Figure 11. Crosstalk impairments.
BERT selection for 100G input testing

Keysight offers two series of BERT products which cover 100GbE speeds. The J-BERT M8020A with the M8041A, M8061A and N4877A are a high performance, full featured system for use up to 32 Gb/s. The system features industry leading intrinsic random jitter under 200 fs rms, integrated pattern generator de-emphasis and error analyzer equalization, all of the calibrated timing jitter impairments required for Standards compliance testing as well as compliant clock recovery.

Figure 12. Illustration of M8020A configured for generic 25G input test.
The N4960A is a fully integrated BERT system for use up to 32 Gb/s which is a cost effective system ideal for general design characterization. With fast rise time pattern generator output, it is also ideal for aggressor generation when testing crosstalk susceptibility. The small pattern generator and error detector heads can be remotely located close to the DUT, minimizing the degradation caused by interconnecting cables. Pattern generator selections include integrated de-emphasis and high amplitude outputs, capable of directly driving VCSELs and laser modulators.

Detailed information of features, specifications and how these BERTs can help you speed up your 100GbE testing can be found in the product data sheets and at www.keysight.com/find/M8020A and www.keysight.com/find/N4960A.

References:
- IEEE 802.3ba
- IEEE 802.3bj
- IEEE 802.3bm
- Fibre Channel PI-6 and MSQS-2
Evolving Since 1939

Our unique combination of hardware, software, services, and people can help you reach your next breakthrough. We are unlocking the future of technology. From Hewlett-Packard to Agilent to Keysight.

myKeysight

www.keysight.com/find/mykeysight
A personalized view into the information most relevant to you.

http://www.keysight.com/find/emt_product_registration
Register your products to get up-to-date product information and find warranty information.

Keysight Services

www.keysight.com/find/service
Keysight Services can help from acquisition to renewal across your instrument's lifecycle. Our comprehensive service offerings—one-stop calibration, repair, asset management, technology refresh, consulting, training and more—helps you improve product quality and lower costs.

Keysight Assurance Plans

www.keysight.com/find/AssurancePlans
Up to ten years of protection and no budgetary surprises to ensure your instruments are operating to specification, so you can rely on accurate measurements.

Keysight Channel Partners

www.keysight.com/find/channelpartners
Get the best of both worlds: Keysight’s measurement expertise and product breadth, combined with channel partner convenience.

www.keysight.com/find/M8020A