Electronic Design

BEN MILLER, Product Marketing Manager,
Digital Products, Keysight Technologies

Boost AI Network Reliability with End-to-End 1.6T Interconnect Testing

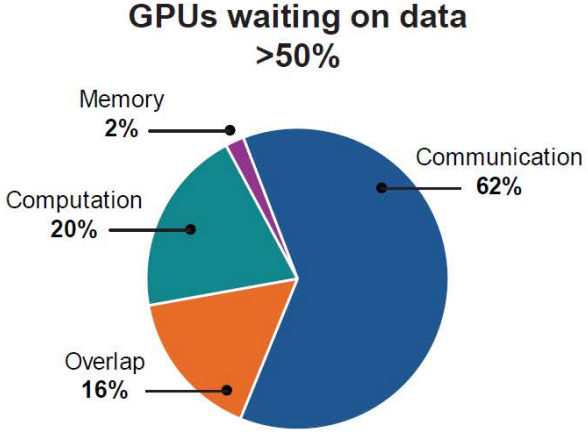
As data speeds increase to 224 Gb/s per lane and infrastructure complexity grows, even small signal-integrity issues can degrade performance.

The "Era of AI" is here, transforming how we work and live, but it's pushing data centers to their limits. Training large language models (LLMs) demands massive volumes of computing and memory, distributed across densely interconnected GPU clusters.

To keep up with model complexity and scale, hyperscale data center operators are racing to upgrade their networks to 800GE and 1.6T Ethernet. But boosting the network capacity with higher-bandwidth interconnects is only part of the story. The real challenge is how to boost interconnect reliability and efficiency to withstand the stress of continuous AI training workloads.

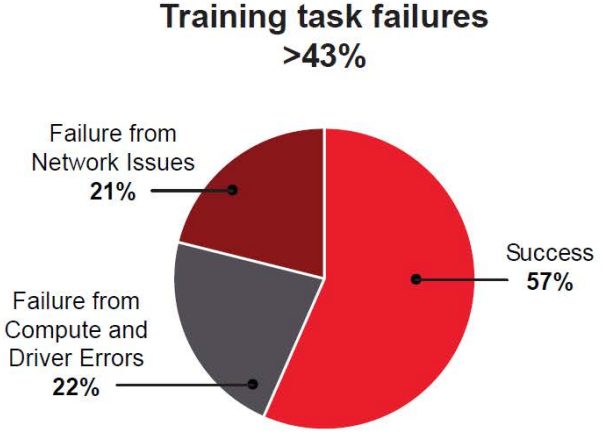
AI networks are only as fast as the weakest link between the clusters. Every transceiver, cable, and connector can affect system-wide throughput, latency, and reliability. Performance bottlenecks, interoperability gaps, and tail latency derail model training. As data center operators upgrade their networks to 1.6T networks and beyond, network architects must consider how each component performs under heavy AI workloads and real-world conditions.

Building to spec is just the beginning. Transceivers must be rigorously validated from design to manufacturing to en-



AI Computer Vision GPU Utilization

Vision transformer (ViT) example.
Source: https://github.com/facebookresearch/HolisticTraceAnalysis/



Success Rate Distribution of the Most Resource-Intensive LLM Training Tasks

Source: Unicron: Economizing Self-Healing LLM Training at Scale, Tao He[1], Xue Li[1], Zhibin Wang[1,2], Kun Qian[1], Jingbo Xu[1], Wenyuan Yu[1], Jingren Zhou[1]
[1]Alibaba Group, [2]Nanjing University

1. Training slowdowns and workload failures are caused by unoptimized networks.

2. Keysight's Infiniium UXR-B high-bandwidth oscilloscope is shown running an IEEE P802.3dj conformance test application.

sure not just interoperability, but optimal system-level performance under real-world conditions.

This article explores the challenges of enabling 1.6T networks for AI data centers and highlights the best practices for validating device performance at the physical layer and beyond. You will gain a deeper understanding of the metrics that matter, the tools you need, and the strategies to help ensure your components are stress-tested and ready for deployment at AI scale.

How Interconnects Become Data Center Bottlenecks

Training LLMs isn't just about raw compute power—it demands fast, synchronized communication across massive GPU clusters. These clusters are built with disaggregated servers connected through high-speed electrical and optical interconnects.

Training is broken down and processed in parallel across different cluster nodes, each node handling a portion of the model. All nodes must stay in lockstep to proceed efficiently *(Fig. 1)*.

As workloads scale, so do the risks of imbalance. Parallelism creates a strong interdependence between nodes. Each node represents a potential weak link in the network. A single underperforming link, whether it's a transceiver, cable, or switch, can bottleneck the entire cluster. In unoptimized networks, GPUs are left sitting idle over half the time, waiting on their next task due to slow interconnects.

To optimize AI workload processing, data center operators need to stress-test each component and interconnect in the network. Transceiver failures are a major cause of workload failures and tail latency, and almost 50% of training tasks fail from network or compute issues. This puts the onus on the transceiver and interconnect manufacturers to design not just for spec-sheet compliance, but also for high-margin performance in the high temperature and massive workload conditions commonly present in AI data centers.

Validating Transceiver Compliance at the Physical Layer

The first step to prevent your interconnect from becoming a bottleneck in an AI data center is physical-layer performance validation. During development, every 224-Gb/s electrical and optical lane must be rigorously tested for signal integrity, interoperability, and real-world reliability under stressed conditions representative of AI training workloads.

Electrical transmitter and receiver testing

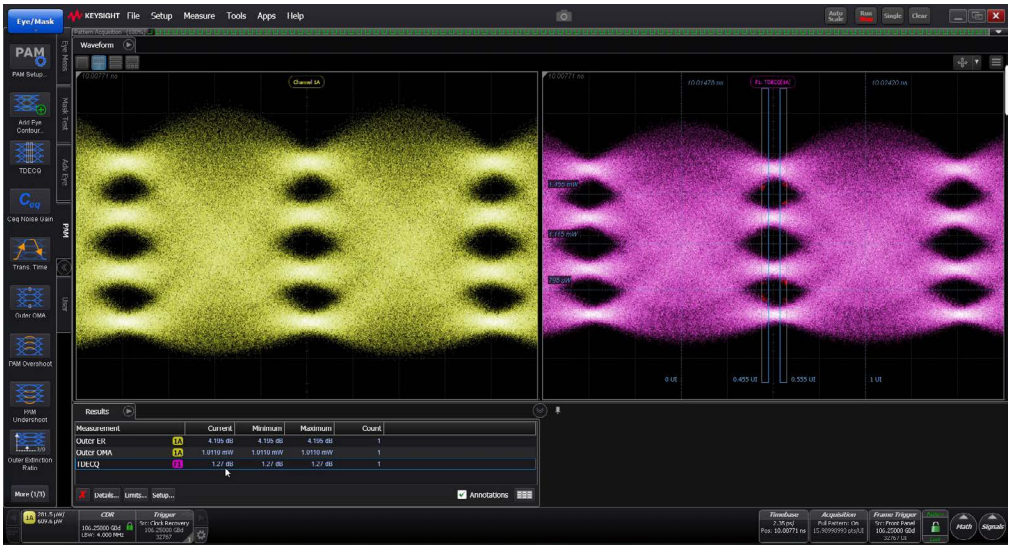
1.6T transceivers must meet increasingly tight signal integrity and noise tolerance requirements for each of the 224-Gb/s electrical and optical lanes. Standards like the IEEE P802.3dj for 1.6T Ethernet specify limits for transmitter jitter, transmitter dispersion penalties, bit error rate, and signal-to-noise-and-distortion ratio.

On top of that, real-world workloads in AI data centers push devices far beyond normal operational limits. Ensuring performance margins under increasingly strict physical-layer specifications is difficult, but important for device reliability and interoperability.

Signal integrity is paramount for electrical testing. Key transmitter measurements include jitter, signal-to-noise-and-distortion ratio, linearity, and signal-to-residual-inter-symbol-interference ratio. Characterizing and tuning equalization on the transmitter for the best, clearest transmission is key to compensating for channel loss.

Developers need a high-bandwidth oscilloscope for signal capture and analysis *(Fig. 2)*. Conformance automation software can guide users through complex validation needs and tests that determine pass/fail status for each required specification. Further signal-integrity and debugging software can help solve challenging conformance issues.

Receiver testing involves using a bit-error-rate tester (BERT) to inject stressed patterns and quantify error rates under degraded conditions. This is necessary for compli-



3. Shown is the TDECQ measurement of a 106-Gbaud PAM4 (112 Gb/s) signal.

ance testing, but it's especially vital for linear pluggable optics (LPO), a new transceiver topology that sacrifices a DSP to reduce its power consumption. This significantly reduces the performance margin for the network interface card and switch, where the host chip must accommodate for a much more distorted and noisier signal.

Picking the right BERT and oscilloscope to perform this testing is a matter of choosing a pattern generator and error detector that use the right modulation format and correct symbol rate (120-Gbaud PAM4 for 224-Gb/s signals).

Optical transmitter testing

Testing optical performance is all about making accurate transmitter dispersion and eye closure quaternary (TDECQ) measurements. TDECQ quantifies the power penalty introduced by a real transmitter compared to an ideal reference at a specific target symbol error rate (SER). It aggregates impairments such as bandwidth limitations, noise, and intersymbol interference in a single metric.

Ethernet standards rely on TDECQ as the primary metric for testing optical transceivers as a pass/fail criteria for compliance, so it's a key differentiator for transceiver reliability and interoperability.

Other important optical metrics include optical modulation amplitude (OMA—the optical power gap between optical 0 and 1 levels), extinction ratio (the ratio of power between high and low logic levels for laser performance), reference equalizer noise enhancement coefficient (Ceq), and laser relative intensity noise (RIN).

Optical measurements are typically made with a low-noise sampling oscilloscope. An ideal sampling oscilloscope for 224-Gb/s validation has extremely low intrinsic noise (<20 µW RMS) and jitter (<90 fs RMS) for the best optical accuracy and sensitivity. 112-Gbaud clock recovery is also required for compliance with 224-Gb/s optical measurements *(Fig. 3)*.

Automated test programs integrate with oscilloscopes to perform quick TDECQ for compliance validation. Sampling oscilloscopes can be used to evaluate device performance both during R&D and manufacturing, characterizing waveform quality, analyzing impairments, and allowing for fine-tuning of equalization.

The higher the optical sensitivity, the more measurement margin you can achieve, giving you a better characterization of the device's signal-integrity performance. Margin testing with stressed eye conditions over various temperature variations is key to ensuring more robust operation in real-world conditions.

Checking Signal Integrity and Cable Testing

Transceivers are only as good as the interconnects and cables that carry their signals. While most AI data centers rely on optical links today, some deploy active copper cables (ACCs) or traditional passive copper cables.

Regardless of the medium, ensuring robust signal integrity at 224 Gb/s is challenging: channel loss budgets are higher, reflections are more disruptive, and crosstalk more intense. Small discontinuities or impedance mismatches can significantly degrade performance. Characterizing and analyzing this behavior requires advanced measurement techniques.

S-parameters analysis gives a frequency-domain view of insertion loss (signal power lost along the channel), return loss (reflections due to impedance mismatches), and crosstalk (unwanted coupling between adjacent signal paths). Characterizing S-parameters for 224-Gb/s links requires a vector network analyzer (VNA) supporting 70- to 110-GHz bandwidth. Channel operating margin (COM) is a threshold metric included in recent IEEE 802.3 specs that com-

bines insertion loss, reflections, crosstalk, and noise.

Time-domain reflectometry (TDR) complements frequency-domain analysis by adding spatial resolution, revealing the location and magnitude of impedance mismatches in complex PCB traces, connectors, and cables. Many modern VNAs and high-speed oscilloscopes now include TDR modes to isolate root causes of discontinuities.

Finally, de-embedding is essential to remove the effects of test fixtures and adapters from the measurement path, ensuring that measurements reflect only the behavior of the device under test (DUT). This improves simulation correlation and measurement accuracy. A popular algorithm for accurate de-embedding of physical structures within high-speed digital channels is Automatic Fixture Removal (AFR).

4. This is a typical optical transceiver manufacturing test station that includes a sampling oscilloscope, clock recovery, and optical switch/controller running test automation software.

Scaling Transceiver Production Without Compromising Quality

Once a transceiver design is finalized, manufacturers need to ramp up volume production quickly to meet AI data center hardware demands. Slow manufacturing ramps can mean missing the market entirely, but manufacturers also can't cut corners. It's critical to maintain high yields to avoid shipping faulty units and causing problems in final deployment. The challenge for QA engineers is to ramp production quickly without compromising reliability or test accuracy.

Many test engineers build their test programs around sampling oscilloscopes, using them to "tune" their optical transceivers during manufacturing. In a fast feedback loop, the production system writes initial settings into a newly built transceiver module (such as laser bias, modulator voltage, etc.), then immediately measures a key performance metric like TDECQ.

Based on the measured result, the settings are adjusted, and the metric is measured again. This loop repeats rapidly until the transmitter is optimized and the device can pass performance thresholds.

Typically, optimization testing is performed at multiple temperature setpoints as well, ensuring that each unit will perform optimally across the device's entire operating temperature range. The sampling oscilloscope can be paired with an optical switch and test automation software to facilitate testing each lane, enabling testing multiple lanes and devices at once to reduce test time and improve hardware utilization *(Fig. 4)*.

Going Beyond Physical-Layer Testing

Meeting physical-layer compliance is necessary, but it's only the beginning for ensuring transceiver performance in AI data centers. Standards like IEEE 802.3dj define specifications at 224 Gb/s per lane, but many issues can crop up in real-world conditions that aren't detectable from physical validation. As links approach and expand beyond 1.6T speeds, you need to take testing to the next level by stress testing beyond the physical layer.

"Layer 1.5": Forward Error Correction

At these speeds, forward error correction (FEC) is essential to keep the aggregate bit error rate (BER) at an acceptable level across all lanes for reliable data transmission. While receiver compliance tests focus on pre-FEC BER, a compliant receiver still needs to perform at an acceptable BER level for FEC to be effective. Post-FEC analysis is about determining the error distribution where FEC becomes unreliable. Even with an average BER within spec, burst errors can exceed FEC limits and lead to unrecoverable frame loss.

Because AI data centers demand near-zero data loss, FEC performance is crucial to analyze. That's why analyzing post-FEC performance, specifically error distribution and frame loss rate (FLR), is just as important as traditional receiver compliance metrics. Interconnect test platforms can stress devices under full 1.6T loads using real packet flows and impairments to simulate system-level noise and congestion.

By validating end-to-end link behavior, including FEC tail analysis to examine how a receiver handles consecutive errors, developers can assess whether a transceiver is truly ready for AI deployment *(Fig. 5)*.

	A	B	C
1	Name	10.109.3.230/1	10.109.3.230/2
62	PCS Local Faults	0	0
63	PCS Illegal Ordered Set	0	1
64	PCS Illegal Idle	0	0
65	PCS Illegal SOF	0	0
66	PCS Out Of Order SOF	0	0
67	PCS Out Of Order EOF	0	0
68	PCS Out Of Order Data	0	0
69	PCS Out Of Order Ordered Set	0	0
70	Transmit Neighbor Solicitations	0	0
71	Transmit Neighbor Advertisements	0	0
72	Receive Neighbor Solicitations	0	0
73	Receive Neighbor Advertisements	0	0
74	FEC Total Bit Errors	74	7,315,343,017,487
75	FEC Max Symbol Errors	1	>15
76	FEC Corrected Codewords	74	6,318,610,141,288
77	FEC Total Codewords	28,217,238,144	133,027,704,084,304
78	FEC Frame Loss Ratio	0.00e+00	1.46e-12
79	pre FEC Bit Error Ratio	4.82e-13	1.01e-05
80	FEC Codewords with 0 errors	28,217,238,070	126,709,093,942,822
81	FEC Codewords with 1 error	74	5,449,731,753,954
82	FEC Codewords with 2 errors	0	757,697,968,918
83	FEC Codewords with 3 errors	0	96,679,636,348
84	FEC Codewords with 4 errors	0	12,609,889,284
85	FEC Codewords with 5 errors	0	1,644,585,451
86	FEC Codewords with 6 errors	0	214,484,273
87	FEC Codewords with 7 errors	0	27,920,406
88	FEC Codewords with 8 errors	0	3,542,046
89	FEC Codewords with 9 errors	0	358,678
90	FEC Codewords with 10 errors	0	1,926
91	FEC Codewords with 11 errors	0	2
92	FEC Codewords with 12 errors	0	0
93	FEC Codewords with 13 errors	0	0
94	FEC Codewords with 14 errors	0	0
95	FEC Codewords with 15 errors	0	2
96	FEC Uncorrectable Codewords	0	194
97	FEC Transcoding Uncorrectable Events	0	97
98	L1 Bits Sent	0	50,374,532,755,858,080
99	L1 Bits Sent Rate	0	0
100	L1 Bits Received	0	50,372,278,653,946,752
101	L1 Bits Received Rate	0	0
102	L1 Line Rate Transmit (%)	0.0000	0.0000
103	L1 Line Rate Receive (%)	0.0000	0.0000
104	Transmit Arp Gratuitous	0	0
105	Transmit Arp Reverse	0	0
106	Transceiver Temperature (C)	47 C	47 C
107	Transceiver Interrupt Asserted	No	Yes
108	Encoding	PAM4 106G	PAM4 106G
109	PGID Overflow	0	0
110	Active FEC mode	KP4-FEC	KP4-FEC
111	Transceiver Voltage (V)	3.233 V	0.000 V

5. Optimized (left) and unoptimized (right) interconnect FEC margin performance. The decreasing number of FEC codewords as the number of errors increase is called the "FEC tail."

Layers 2-3: Network Performance Testing

For AI data center transceivers, testing can't stop at Layer 1. It's critical to extend validation to full protocol stack performance. Developers should validate Layer 2/3 behavior under realistic conditions to uncover issues related to MAC addressing, routing, IP packet handling, and transport efficiency, ensuring that transceivers can support parallel data transfers of AI training workloads.

Emulating real Ethernet/IP traffic patterns at full line rate can expose issues in routing, flow control, latency, link stability, and congestion that aren't visible through physical waveform analysis alone *(Fig. 6)*.

Combining physical-layer validation, FEC-aware receiver testing, and full protocol stack (Layers 1-3) emulation, transceivers developers can have a more complete picture of transceiver performance. This holistic approach may seem like a lot, but it's critical to ensure the reliability, throughput, and efficiency of AI data center interconnects as networks upgrade to 1.6T Ethernet and beyond.

What's Next? Early Pathfinding and R&D

As the industry prepared to deploy 1.6T Ethernet over 224-Gb/s lanes, some developers are already looking toward the future. The next step, the technological jump that will push data center speeds to 3.2T Ethernet, is 448-Gb/s lanes.

	A	B	C
1	Name	10.109.3.230/3	10.109.3.230/4
2	Link State	Link Up	Link Up
3	Line Speed	800GE	800GE
4	Frames Sent	15,114,349,228	15,114,349,227
5	Frames Sent Rate	0	0
6	Valid Frames Received	15,114,349,227	15,114,349,228
7	Valid Frames Received Rate	0	0
8	Bytes Sent	967,318,350,592	967,318,350,528
9	Bytes Sent Rate	0	0
10	Bytes Received	967,318,350,528	967,318,350,592
11	Bytes Received Rate	0	0
12	Fragments	0	0
13	Undersize	0	0
14	Oversize and Good CRCs	0	0
15	CRC Errors	0	0
16	Vlan Tagged Frames	0	0
17	Flow Control Frames Received	0	0
18	Oversize and CRC Errors	0	0
19	User Defined Stat 1	0	0
20	User Defined Stat 2	0	0
21	Capture Trigger (UDS 3)	15,114,349,227	15,114,349,228
22	Capture Filter (UDS 4)	15,114,349,227	15,114,349,228
23	User Defined Stat 5	0	0
24	User Defined Stat 6	0	0
25	ProtocolServer Transmit	0	0
26	ProtocolServer Receive	0	0
27	Transmit Arp Reply	0	0
28	Transmit Arp Request	0	0
29	Transmit Ping Reply	0	0
30	Transmit Ping Request	0	0
31	Receive Arp Reply	0	0
32	Receive Arp Request	0	0
33	Receive Ping Reply	0	0
34	Receive Ping Request	0	0
35	Transmit Duration(Cleared on Start Tx)	0 :7 :6.158924188	0 :7 :9.282154263
36	Pause End Frames	0	0
37	Pause Overwrite	0	0
38	Scheduled Frames Sent	15,114,349,228	15,114,349,227
39	Asynchronous Frames Sent	0	0
40	Port CPU Frames Sent	0	0
41	Link Fault State	No Fault	No Fault
42	Local Faults	0	0
43	Remote Faults	0	0
44	Scheduled Transmit Duration	0 : 0: 0.0	0 : 0: 0.0
45	Bytes Sent / Transmit Duration	2,269,853,558	2,253,339,303
46	Bits Sent	7,738,546,804,736	7,738,546,804,224
47	Bits Sent Rate	0	0
48	Bits Received	7,738,546,804,224	7,738,546,804,736

6. 800GE link testing statistics showing zero packet loss; packets/bytes sent and receive match.

Three primary signaling options are under evaluation: 224-Gbaud PAM4, 174-Gbaud PAM6, and 150-Gbaud PAM8. Each presents tradeoffs in complexity, bandwidth, and noise tolerance. Early research shows that any of these potential methods could be valid, and some developers are already working to find the best solution using high-speed arbitrary waveform generators and sampling oscilloscopes.

At the same time, emerging transceiver topologies like LPO, enabled by new photonic ICs, may affect these choices as performance demands and priorities change and reshape the next generation of data standards.

Interconnects are No Longer Just Passive Links that Require Advanced Testing

Regardless of the direction taken by the next generation of data standards, one thing is clear: Interconnects are no longer just passive links in the system—they're critical performance enablers. In the new AI data center architecture, every component, transceiver, and interconnect is a potential weak link. Developers and architects must shift their mindset beyond interoperability compliance to performance optimization at every level.

It's a large task to take on, but with advanced end-to-end test, automation, and emulation tools from forward-looking companies like Keysight, every step in the design and validation cycle is achievable, and every improvement can have an outsized impact on efficient, reliable AI data center operations.