

Keysight Technologies

Data Link and Transaction Layer Testing for PCI Express® and NVMe™

Application Note



Overview

PCI Express® (PCIe®) technology has been implemented broadly in systems requiring high-speed data transfer—such as video, graphics, and networking. As a serial point-to-point interconnect between two devices, PCIe buses implement a packet-based protocol for information transfer and provide scalable performance based on the number of signal lanes implemented.

Initially embraced by high-performance desktop and server systems, PCIe technology is now finding a home in embedded applications. Solid State Drives (SSDs) are a great example. Although many applications use PCIe buses, the storage industry is largely pushing the envelope of its development and technology. SSDs are quickly replacing traditional spinning Hard-Disk Drives (HDDs) due to their superior reliability and speed. It is this transition that is motivating a move from SCSI, SATA and SAS to NVMe Express™ (NVMe™) at the application layer of PCIe for storage applications (Figure 1). And with the global market for PCIe-based SSDs projected to continue growing, ensuring those units are successfully tested and meet standards is more important than ever.



The standards that comprise PCIe are managed by a non-profit organization called the Peripheral Component Interconnect Special Interest Group (PCI-SIG®). The PCI-SIG is run by a board of nine member-companies.

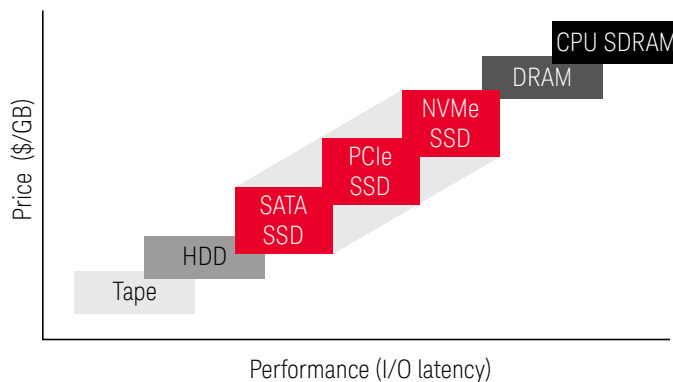


Figure 1. As storage technology has improved, the demand for higher performing interconnect has grown.

The PCIe structure includes a synchronous clock and multi-drop parallel bus with multiple embedded-clock serial links. Each serial link transfers data in one direction and can be routed as a differential trace pair independent of the other links. This architecture greatly simplifies board layout. A pair of links, one for transmission and one for receiving, called a “lane”, moves data in opposite directions between two devices. This full-duplex, multi-link transport gives a PCIe-connected SSD drive a 10x throughput advantage over standard HDD drives, as well as improved performance capabilities over other SSD interconnect alternatives (e.g., SAS and SATA).

Although PCIe is a mature technology (approaching its fourth generation, with compatible devices shipping for several years), rigorous testing for standards compliance is still required to assure interoperability, reliability, and overall improved functionality from generation to generation.

Challenge: Protocol-Level Testing

While storage has traditionally lagged behind CPU in terms of I/O capabilities, development in SSDs, PCIe interconnect, and the NVMe protocol have been closing that gap. PCIe's current (3.0) and upcoming (4.0) specifications introduce some significant changes, including higher speeds in the electrical block of the physical layer and some new low-power features that require different approaches to testing and validation.

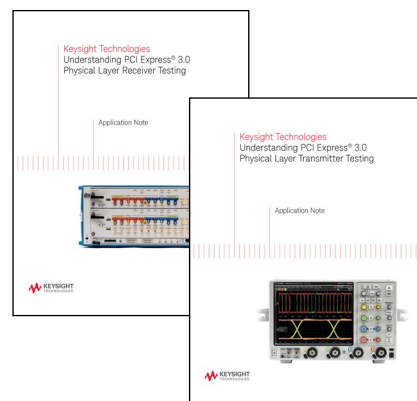
There are 3 different categories of testing required to reach compliance: physical layer, data link layer and interoperability tests. Products must successfully pass all tests at an official PCI-SIG workshop using the approved test fixtures in order to be deemed compliant and branded as an approved PCIe device. This application note will focus on protocol-level validation, specifically focusing on the data link and transaction layers.

An overwhelming concern for PCIe-based validation teams is interoperability and backward compatibility with previous PCIe generations. This requires tools to validate the parametric and protocol aspects of designs to ensure compliance, and to validate design performance. While the PCI-SIG provides the Compliance Base Board (CBB), the Compliance Load Board (CLB), and the SigTest software to facilitate electrical compliance testing at the physical level, protocol-level compliance requires a different approach.

For protocol level testing there are two different categories of tests: protocol test card (PTC) and active testing.

The PCI Express 3.0 link and transaction layer tests as defined by the PCI-SIG require the use of a PTC. This card is from an approved list of vendors and the card monitors a standard set of tests. It is considered a pass if a product successfully passes all tests on this system at a compliance workshop. All of these tests are actively executing PCIe operations to validate that the device correctly implements the PCIe protocol and is in compliance with the specification.

In addition to PTC tests, other tests are required to achieve compliance with the DUT. The focus is on monitoring the platform configuration of the DUT working under active conditions. These tests are run through a PCI-SIG approved exerciser card that steps the product through a series of conditions that it may encounter in a working environment. The exerciser puts the DUT into stressed conditions on the edge of compliance. A pass condition provides assurance that the DUT can handle those cases in a working system as well.



For a more in-depth understanding of the Tx and Rx testing processes and challenges, see 5992-0864EN *Successful PCI Express® Physical Layer Transmitter Testing* and 5992-0818EN *Understanding PCI Express® 3.0 Physical Layer Receiver Testing*.

Breaking down the layers

PCIe technology is a layered protocol. The physical layer is the physical link to transfer information. Within that is an electrical sub-block for analog signaling to send data. There's also a logical sub-block, which manages equalization training (both from the receiver and the transmitter) to set up specific parameters for achieving a clean signal operating across the link (see Figure 2).

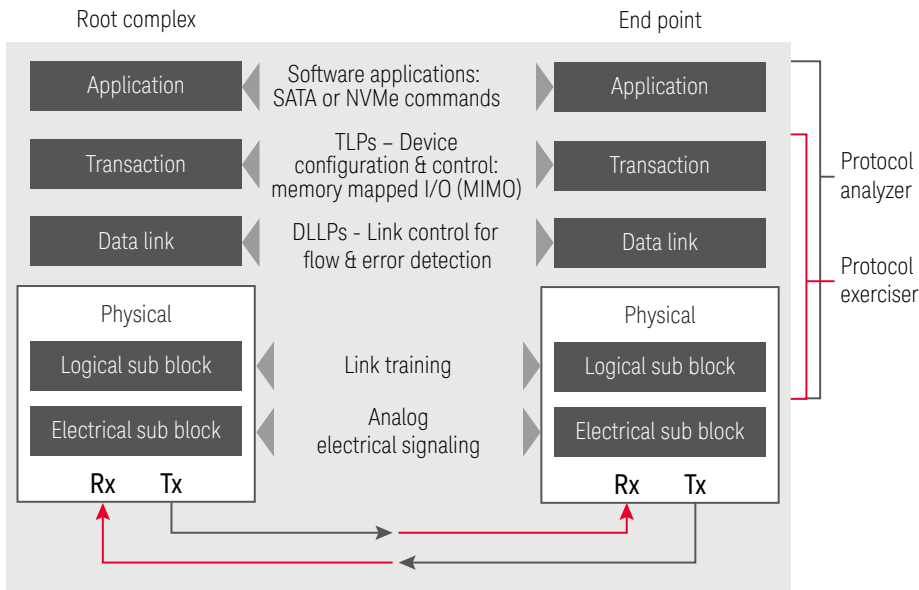


Figure 2. Different layers of the PCIe protocol require different kinds of testing.

The data link layer, above the physical layer, manages link control. All acknowledgements and retransmissions take place at the data link layer, as well as error detection and correction. The data link also contains a flow control, which ensures that the other end is ready to receive data.

The transaction layer is where data is actually moved, its memory-mapped I/O ensures that data packets are reading and writing directly to memory. The transaction layer includes configuration and control for moving data across the link. The transaction layer also includes TLPs (transaction layer protocol packets), such as memory read/write data.

Storage software operations originate in the application layer. This is where protocols like NVMe exist as an application for storage management. Packaged within the TLPs are the NVMe application commands, which focus on storing and retrieving the data across the link.

Protocol-level validation challenges

Once users achieve a data valid state at the physical layer level, they can begin to test the higher layers at the protocol level with a protocol analyzer and exerciser. The analyzer connects to the link between the two devices, in an active state to observe and evaluate different data at each level. The exerciser is connected as an end device to emulate stressed conditions and record how the DUT responds.

Validation of the PCIe data link layer is performed by the specification tests that check for data link layer protocol packets (DLLPs) being transferred, as well as acknowledgements (Ack), negative acknowledgements (Nak), retransmissions, and flow control.

Validation teams need robust systems that can recover from all errors, including intermittent failures. Ideally one could inject errors into real-life traffic, see how the design responds, and make improvements accordingly. This approach, however, is challenging given the difficulty of emulating a real-world setup and the ever-present need to reduce validation time.

1 Functionality and speed-to-market

The primary challenge of protocol validation is to test system functionality with speed and accuracy so that the product can go to market. Protocol errors must be detected, analyzed, and corrected in an efficient manner.

Debugging PCIe protocol means capturing at-speed traffic including power management transitions. Protocol debug tools need to lock onto traffic quickly, then trigger on a unique protocol sequence. Debugging lower level problems, such as power management, requires exceptionally fast lock times. Once traffic is captured, viewing the data at different levels of abstraction makes it possible to narrow in on the problem.

2 Non-intrusive probing

Another key challenge that occurs within PCIe protocol debug has is gaining access to the signals in a non-invasive manner. The probe connected to the PCIe bus must not disturb or change the signals in a manner that impacts bus values or operation.

Where and how signals are probed depends upon each unique system design. Mid-bus probes provide access to traffic, but must not impact signal quality. Slot interposers need to pass signals passively over a long trace and not change the signals. Validation teams require flexible access with multiple probing options.

3 Rapidly increasing protocol-layer capabilities

With the arrival of PCIe 4.0, increasing speeds to 16 GT/s will continue some of the same trends seen in PCIe 3.0 at 8 GT/s, requiring even greater signal quality tuning and equalization capabilities to achieve validation.

The NVMe protocol presents many new test challenges. NVMe creates a new high-performance scalable host controller interface for PCIe-connected devices that maps memory from the host machine for moving data to and from the storage device, with low latency and end-to-end data protection. The interface will provide optimized command submission and completion paths and supports parallel operation.

What is NVMe?

The NVM Express (NVMe) 1.2 specification defines an optimized register interface, command set and feature set for PCIe-based SSDs. The goal of NVMe is to unlock the potential of PCIe SSDs now and in the future, and standardize the PCIe SSD interface.

NVMe significantly improves both random and sequential performance by reducing latency, enabling high levels of parallelism, and streamlining the command set while providing support for security, end-to-end data protection, and other client and enterprise features users need. NVM Express provides a standards-based approach enabling broad ecosystem adoption and PCIe SSD interoperability.

NVMe benefits:

- Standardized high-performance controller interface for SSDs
- Developed by an open industry consortium
- Architected for non-volatile memory
- Uses a queuing model to send commands and data to the drive
- Support for up to 64k queues, and up to 64k commands per queue
- High-performance, low-latency protocol
- End-to-end data protection

The University of New Hampshire interoperability lab (UNH-IOL) has defined an NVMe conformance specification test suite that will continue to expand. The UNH-IOL offers conformance and interoperability testing for NVMe hosts and devices, helping products to qualify for the NVMe Integrators List. UNH-IOL also uses their NVMe interoperability test bed to enable heterogeneous testing across various OS, drivers, and hardware platforms to prove the viability and robustness of NVMe.

For more information visit,
<http://nvmexpress.org/>

Protocol Validation Test Setup

In order to fully test the data link, validation is required to ensure protocol messages are correctly sent and received. As a result, both a protocol analyzer and exerciser are required.

The ultimate goal of validation with the analyzer is testing the link and data transmission to assure device functionality and interoperability. To achieve this, the protocol analyzer needs to provide visibility into what happens at each unique protocol layer. When troubleshooting, users must track specific errors to the appropriate layer for correction.

To validate the device, the protocol exerciser needs the ability to emulate either a root complex or an end point in the same card. The exerciser acts as an ideal link partner by sending appropriate I/O traffic to stimulate the device under test. The device under test can be exercised under various conditions and scenarios without influencing its performance parameters.

Throughout these functions and tests users must address the complexities of how and where signal access happens. Signal access happens with consistent representation for accurate data recovery. The probing solution needs to be not only high-performance, but also versatile to accommodate different types of system designs. Probing needs to remain non-intrusive, so that access won't impact signal quality or protocol operation.

Each of the below functions needs to be tested with both an exerciser and an analyzer in order to ensure the data link and transaction layers are functional and compliant:

- Equalization (EQ) is the process of training the two terminations to reliably transfer each bit. Each individual lane must be trained. Within a x4 link, each link can have different tuning parameters. Training is dynamic and requires that each lane's pre-cursor, cursor and post-cursor values are set correctly. Signal quality is critical for successful operation. As speeds increase, additional EQ functions are required.
- Link Training Signaling State Machine (LTSSM) analysis observes and controls the different possible states of the link. State transitions handle link up, recovery, and power management. The goal of LTSSM is to reach a state called L0—where the link is active and data can be transferred. Other states within LTSSM, such as Configuration, Training, and Error Recovery Routines, provide link control and recovery as well as power management capabilities.
- Packet capture looks at the individual PCIe packets to decode the responses and device configuration and enumeration. The process of reading and writing from memory is based on addresses contained in the packets in order to transfer data across the link.
- Performance analysis and flow control affects response times and overall throughput. Flow control, clearly identifies and tracks credit starvation issues to see if the receiving device and buffers are ready to accept the data being transferred.

Key protocol testing for NVMe

- Complex protocol analysis: queue create/delete, command processing, reservations, PRP (physical region page), SGL (scatter gather lists), and throughput
- Emulation: Creation of multiple test queues to ensure controller capabilities

After all tests are met to establish a stable communication channel, the next step is to validate the application layer (e.g., NVMe) over the PCIe bus.

In terms of NVMe protocol testing, users will require a tool to observe how the different parts interact. The data link layer, acknowledgements, flow control, multiple queues, and multiple commands need to be coordinated. There may be thousands of packets to analyze in a single NVMe read. While all the specific tests within the NVMe specification have not yet been defined, a protocol analyzer will continue to be an invaluable tool for users to validate designs.

Another important aspect of NVMe testing is emulation. Users need the capability to create multiple queues and commands in operation to ensure a controller can manage the different types of configurations. The companion tool for this sort of emulation process is called an NVMe exerciser.

Keysight Solution (U4301B PCIe Analyzer)

Keysight has developed tools in the physical, data link and protocol layers to meet the challenges posed by the PCIe standard as shown in the figure below. The solutions are designed to reduce development costs, provide accurate test results and simplify the measurement processes for test engineers.

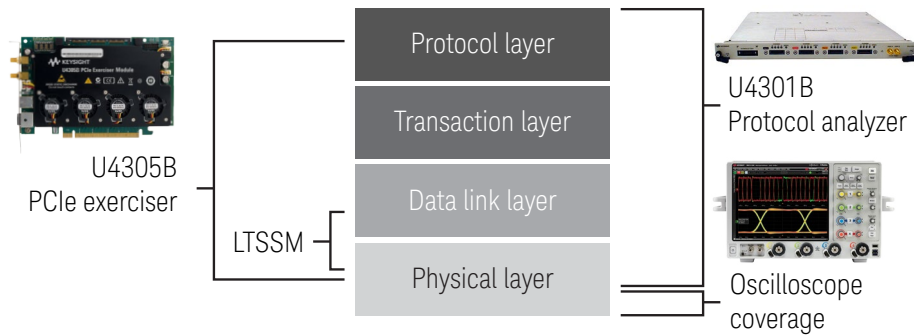


Figure 3. Application coverage of the Keysight U4305B PCIe 3.0 exerciser.

Keysight Technologies' high speed U4301B PCI Express® 3.0 analyzer module is a protocol analyzer supporting all PCI Express® applications from PCIe 1.0 through 3.0. The U4301B analyzer captures and decodes PCIe data and displays it in a packet viewer window.

Keysight Product Info Highlights: U4301B PCIe Analyzer



- Equalization testing: Accurate signal integrity testing
- LTSSM analysis: High-level analysis of all link training processes for channel synchronization
- Transaction decode: Packet-level decode of NVMe and transactional analysis
- Performance analysis: Unmatched capability to compute 50+ statistics with graphical analysis
- Flow control analysis: Clearly identifies and tracks credit starvation issues

The U4301B analyzer is a blade that is installed in an AXIe two-slot M9502A or five-slot M9505A. It offers designers insight to the equalization process and all of the state transitions, with multiple analysis views that can be customized to meet engineering requirements. The analyzer's LTSSM overview pinpoints specific training sequence issues through easy-to-interpret analysis results.

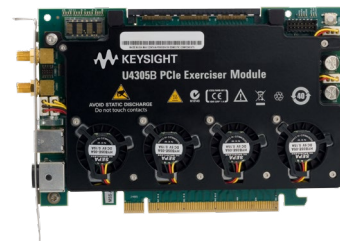
Keysight's transactional decoder includes a viewer that allows users to select NVMe queues and performance information from the analyzer's NVMe transaction overview pane. This organizes the transactions by direction and by queue to follow the data flow across the interface, with one-click control. Individual physical region page (PRP) lists contain all of the key information of the NVMe queues. Quickly review and validate the data flows over the PCIe connections.

The performance analysis package includes the real data throughput calculations, with response-time measurement of the PCIe data flow. Measure and understand throughput performance, PCIe response times, and other operational measurements that provide the insight needed to optimize device performance.

PCIe stimulus and response testing is accomplished with the addition of the U4305B PCIe 3.0 exerciser.

The U4305B offers a broad range of test tools for validation of PCIe 1.0, 2.0 and 3.0 operation for all lane widths up to x16. The U4305B PCIe 3.0 exerciser with pre-defined LTSSM test cases can help validate the complex and hard-to-test state transitions of a DUT's LTSSM. Using the U4305B PCIe 3.0 exerciser, validate your device with the ability to emulate either a root complex or add-in card in the same card. The emulated personality is easily changed with a simple software switch. The LTSSM test cases come with detailed reporting to clearly define the source of the found failures for fast debug.

Keysight Product Info Highlights: U4305B PCIe Exerciser



- Flexible hardware support: Link width x1 – x16, emulate root complex (RC) or add-in-card (AIC)
- LTSSM software capabilities: Pre-defined LTSSM tests to help validate the complex and hard-to-test state transitions of the DUT's LTSSM
- NVMe Root Complex (RC) emulation: Add NVMe commands (available as predefined templates) to submission queues and submit these to the NVMe DUT.

Solution Benefits

Keysight’s U4301B PCI Express® 3.0 analyzer addresses the challenges of protocol-level validation; functionality, non-intrusive probing and increased requirements of NVMe testing.

1 Decoding and analyzing PCIe traffic

Data is displayed from the 8 GB capture buffers using several different views that mark exactly the data desired. The packets window decodes each captured packet and enables quick and easy finding, filtering, and decoding of specific PCIe information. This easy-to-use view lets users define the parameters of interest and display the details of each packet.

Additional displays enable a fast understanding of the data in the buffer. Errors are highlighted in red and tool tips provide details of the error detected (see Figure 4).

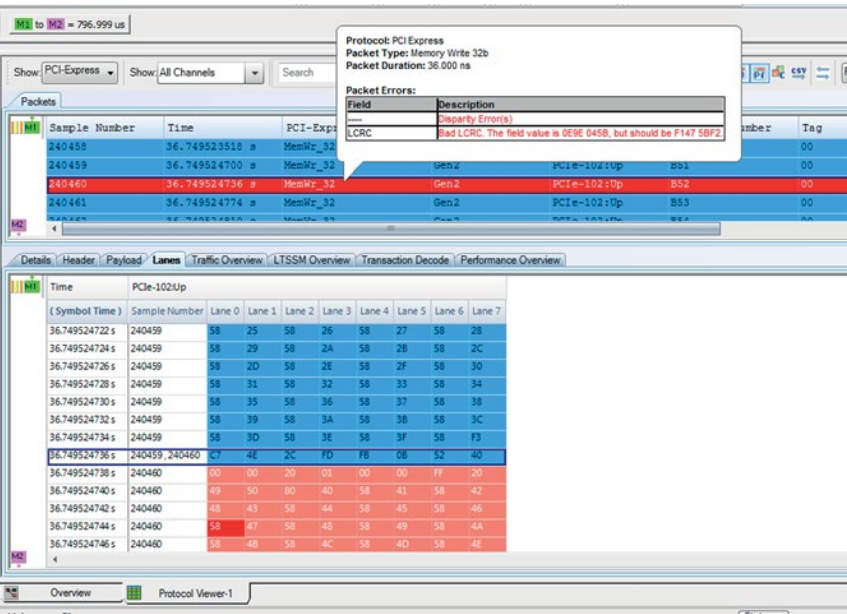


Figure 4: A memory read packet has a bad link cyclical redundancy check and the lane view shows which byte and lane are corrupted.

Keysight's U4305B PCI Express® 3.0 exerciser

The Keysight U4305B Exerciser can be configured to provide sub-protocol layer test and debug for legacy and next generation PCIe devices. The U4305B Exerciser for PCIe is an advanced traffic generator that can be used to send and respond to TLP, DLLP, and physical layer packets to stimulate PCIe devices and systems. The U4305B PCIe exerciser supports a rich programming interface that allows automation of all of the exerciser capabilities and can therefore accommodate specific DUT case requirements. U4305B is a standard height, half-length PC form factor card as described in the PCI Express specification.

Keysight gets involved, you benefit

Keysight's solutions for digital applications are driven and supported by Keysight experts that are involved in the various international standard committees. We call it the Keysight Digital Test Standards Program. Our experts are active in the Joint Electronic Devices Engineering Council (JEDEC), PCI Special Interest Group (PCI-SIG®), Video Electronics Standards Association (VESA), Serial ATA International Organization (SATA-IO), USB-Implementers Forum (USB-IF), Mobile Industry Processor Interface (MIPI) Alliance, and many others. Our involvement in these standards groups and their related workshops, plugfests, and seminars enables Keysight to bring the right solutions to the market when our customers need them.

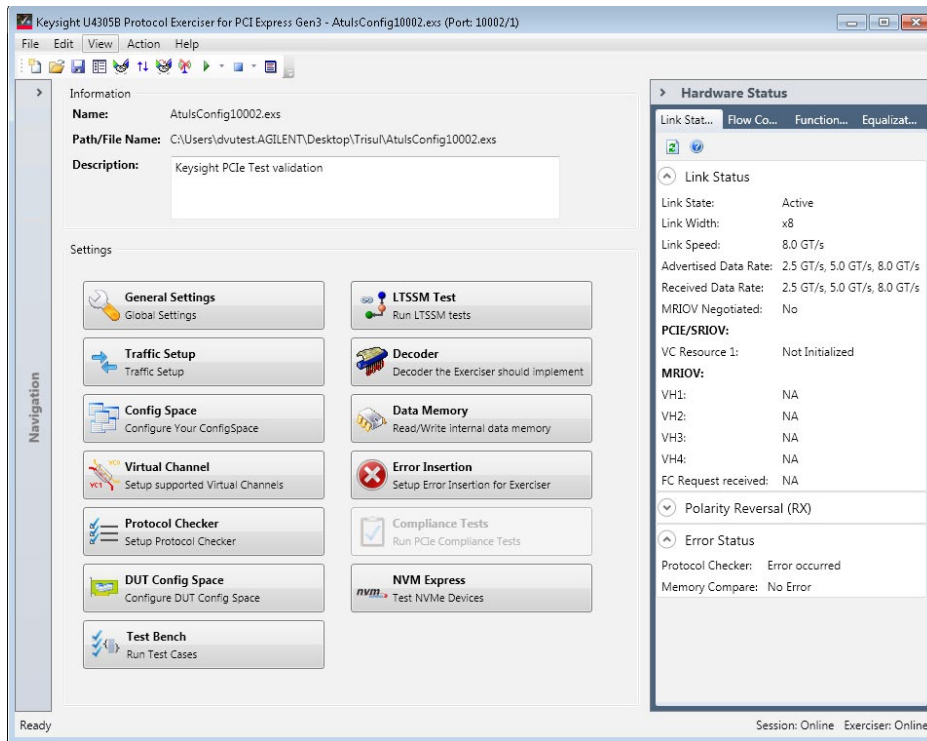


Figure 5. The U4305B can emulate host or add-in cards and tests all PCIe 3.0 link speeds and widths. Controls for traffic generation and error detection or insertion are simple to access and configure.

2 Non-intrusive probing



Figure 6: The U4322A mid-bus probe supports x1 to x16 unidirectional or x1 to x8 bidirectional.

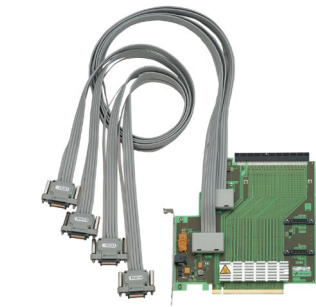


Figure 7: The U4321A slot interposer will support x1 to x16 unidirectional or bidirectional.



Figure 8: The U4324A flying lead solder down probe offers support for x1 and x2 bidirectional capability on a single probe; other standard lane width configuration support is x4, x8, & x16.

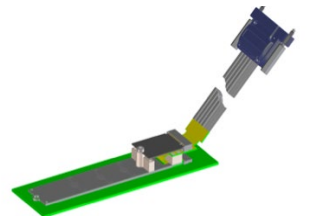


Figure 9: The U4328A M.2 interposer supports testing of M/B-M PCIe SSDs.

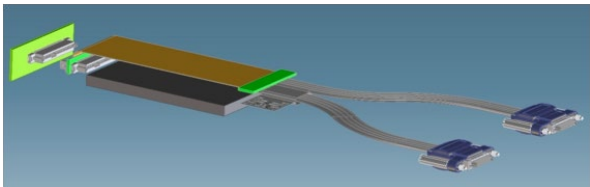


Figure 10: The U4330A SFF-8639 interposer can be used with SSDs with either single- or dual-link support.

3 NVMe testing: transaction decode

The transaction decode tab in the protocol viewer window allows users to compute and view transactions decoded from the captured PCIe traffic. The decoding and display of transactions is done according to the relevant storage protocol specifications (such as NVMe, AHCI, PQI) to help easily correlate the decoded data to the protocol specifications and evaluate the DUT's compliance to specifications. Transaction decoding of NVMe data includes analysis of the trace to identify all of the relevant transaction types to present a transaction overview table that enables quick and easy navigation of key packet types.

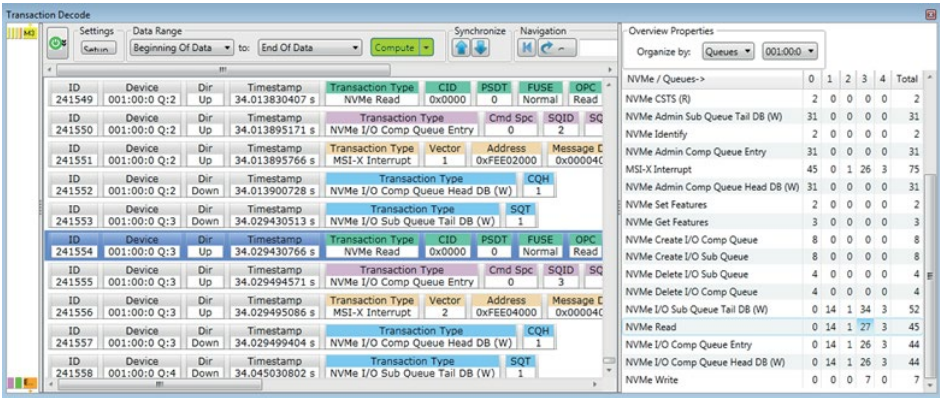


Figure 11: NVMe transaction information provides easy to understand displays of NVMe payloads and PRP messages. Transaction overview of all transactions enables instant access to all messages in the capture trace.

Summary

The ultimate goals for validating a device to PCI-SIG standards are interoperability, reliability, backwards compatibility to previous PCIe standards, and speed to market. PCIe technology continues to provide fast and reliable point-to-point connectivity. PCIe-based SSDs offer potential for a 10-20x improvement in performance over comparable SAS and SATA technologies. The collaborative development of the NVMe protocol specification (which includes Keysight's participation) will help to close the I/O gap between systems and storage.

Protocol-level testing requires tools designed with flexibility not just for the current generation's challenges, but ideally with built-in capabilities that anticipate the evolution of these layers in future versions of specifications.

Whether it's PCIe or NVMe-based, Keysight will continue to help validation teams overcome the challenges of validation with the tools and support they need for fast and accurate testing.

Download your next insight

Keysight software is downloadable expertise. From first simulation through first customer shipment, we deliver the tools your team needs to accelerate from data to information to actionable insight.

- Electronic design automation (EDA) software
- Application software
- Programming environments
- Utility software



Learn more at

www.keysight.com/find/software

Start with a 30-day free trial.

www.keysight.com/find/free_trials

myKeysight**myKeysight****www.keysight.com/find/mykeysight**

A personalized view into the information most relevant to you.

**www.axiestandard.org**

AdvancedTCA® Extensions for Instrumentation and Test (AXIe) is an open standard that extends the AdvancedTCA for general purpose and semiconductor test. Keysight is a founding member of the AXIe consortium. ATCA®, AdvancedTCA®, and the ATCA logo are registered US trademarks of the PCI Industrial Computer Manufacturers Group.

**www.lxistandard.org**

LAN eXtensions for Instruments puts the power of Ethernet and the Web inside your test systems. Keysight is a founding member of the LXI consortium.

**www.pxisa.org**

PCI eXtensions for Instrumentation (PXI) modular instrumentation delivers a rugged, PC-based high-performance measurement and automation system.

**Three-Year Warranty****www.keysight.com/find/ThreeYearWarranty**

Keysight's commitment to superior product quality and lower total cost of ownership. The only test and measurement company with three-year warranty standard on all instruments, worldwide.

**Keysight Assurance Plans****www.keysight.com/find/AssurancePlans**

Up to five years of protection and no budgetary surprises to ensure your instruments are operating to specification so you can rely on accurate measurements.

**www.keysight.com/go/quality**

Keysight Technologies, Inc.
DEKRA Certified ISO 9001:2008
Quality Management System

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/PCle**www.keysight.com/find/hsd-insight****Related Keysight Products**

Keysight U4301B PCI Express® 3.0 Analyzer Module
<http://www.keysight.com/find/U4301b>

Keysight U4301B PCI Express Protocol Analyzer
www.keysight.com/find/u4305b

For more information on Keysight

Technologies' products, applications or services, please contact your local Keysight office. The complete list is available at: www.keysight.com/find/contactus

Americas

Canada	(877) 894 4414
Brazil	55 11 3351 7010
Mexico	001 800 254 2440
United States	(800) 829 4444

Asia Pacific

Australia	1 800 629 485
China	800 810 0189
Hong Kong	800 938 693
India	0124 229 2010
Japan	0120 (421) 345
Korea	080 769 0800
Malaysia	1 800 888 848
Singapore	1 800 375 8100
Taiwan	0800 047 866
Other AP Countries	(65) 6375 8100

Europe & Middle East

Austria	0800 001122
Belgium	0800 58580
Finland	0800 523252
France	0805 980333
Germany	0800 6270999
Ireland	1800 832700
Israel	1 809 343051
Italy	800 599100
Luxembourg	+32 800 58580
Netherlands	0800 0233200
Russia	8800 5009286
Spain	800 000154
Sweden	0200 882255
Switzerland	0800 805353
	Opt. 1 (DE)
	Opt. 2 (FR)
	Opt. 3 (IT)
United Kingdom	0800 0260637

For other unlisted countries:

www.keysight.com/find/contactus

(BP-02-06-15)

PCI-SIG®, PCIe® and the PCI Express® are registered U.S. trademarks and/or service marks of PCI-SIG

The NVM Express™ design mark and NVMe™ word mark are trademarks of NVM Express, Inc.