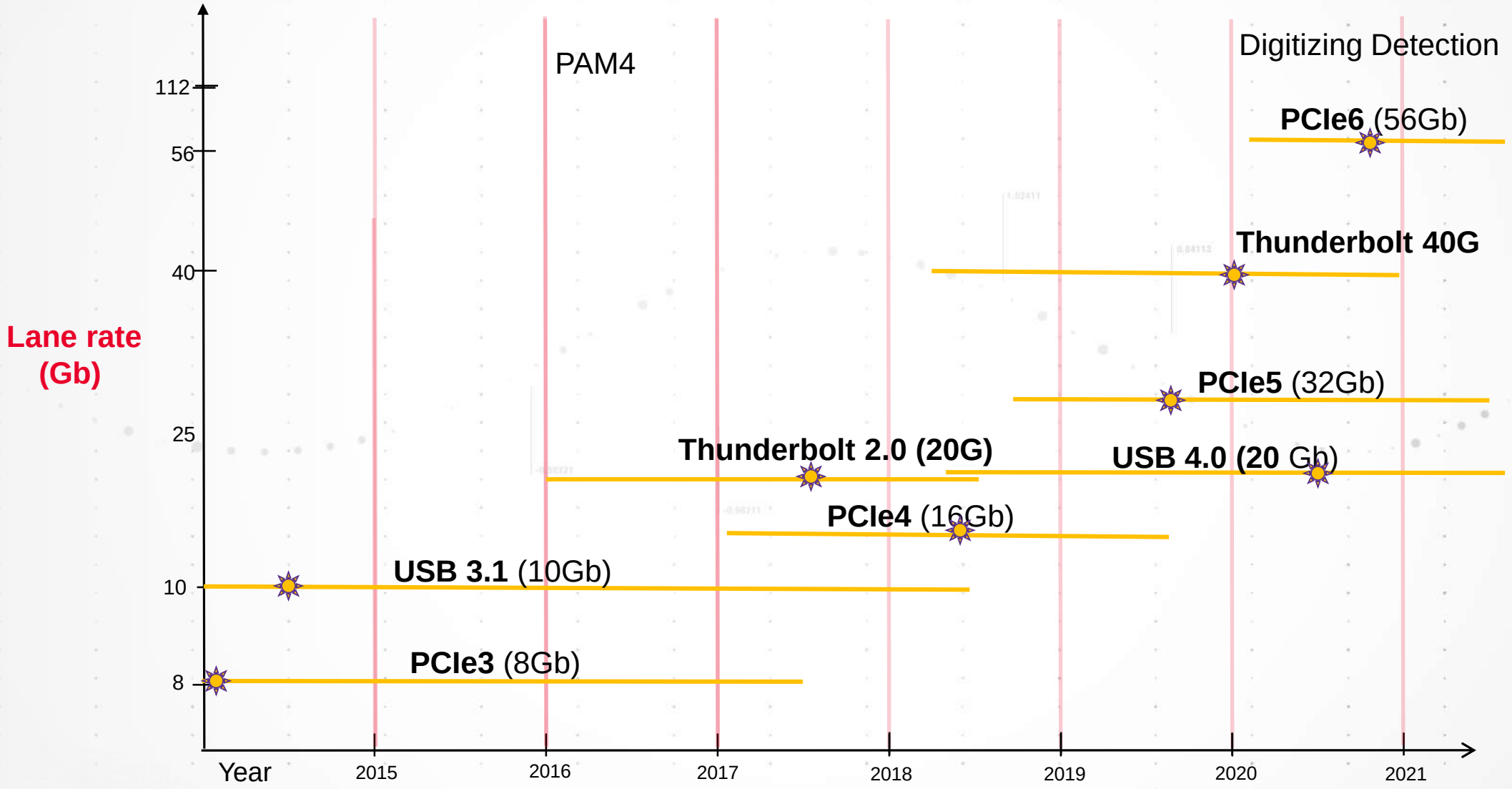
KEYSIGHT WORLD 2019

Evolution of High-Speed Server and Computing Interfaces

Senior Project Manager / Keysight Technologies

Francis Liu & Jacky Yu

High-Speed Computing Roadmap



Keysight Participation in Standards

TRUST OUR EXPERTISE

Memory Perry Keller	DisplayPort Brian Fetz	USB Jit Lim	Computer Rick Eads	Optical Comp. Greg LeCheminant	Optical WAN Stefan Loeffler	HDMI Brian Fetz	400G Steve Sekel	MIPI Roland Scherzinger
Board of Directors JEDEC Compliance Chair	Contributor Board Alumni, VESA Phy Sub Group Compliance	Contributor USB-IF Thunderbolt	Board of Directors PCI-SIG Contributor CCIX GenZ	Contributor IEEE, OIF- CEI, T11 FC	Contributor OIF, ITU, IEC	Contributor HDMI	Contributor IEEE, OIF, 64G FC	TSG Member UniPro Vice Chair MIPI Alliance Thunderbolt
DDR UFS	DP 1.3 Type C	USB 2.0, 3.0, 3.1	PCIe G3, G4, PAM4 (RT)	PAM4 (Opt), CEI 3.1	PAM4 (Opt), CEI 3.1	HDMI 2.0	IEEE802.3bs, IEEE802.3by	D/M/C–PHY, UniPro,TBT





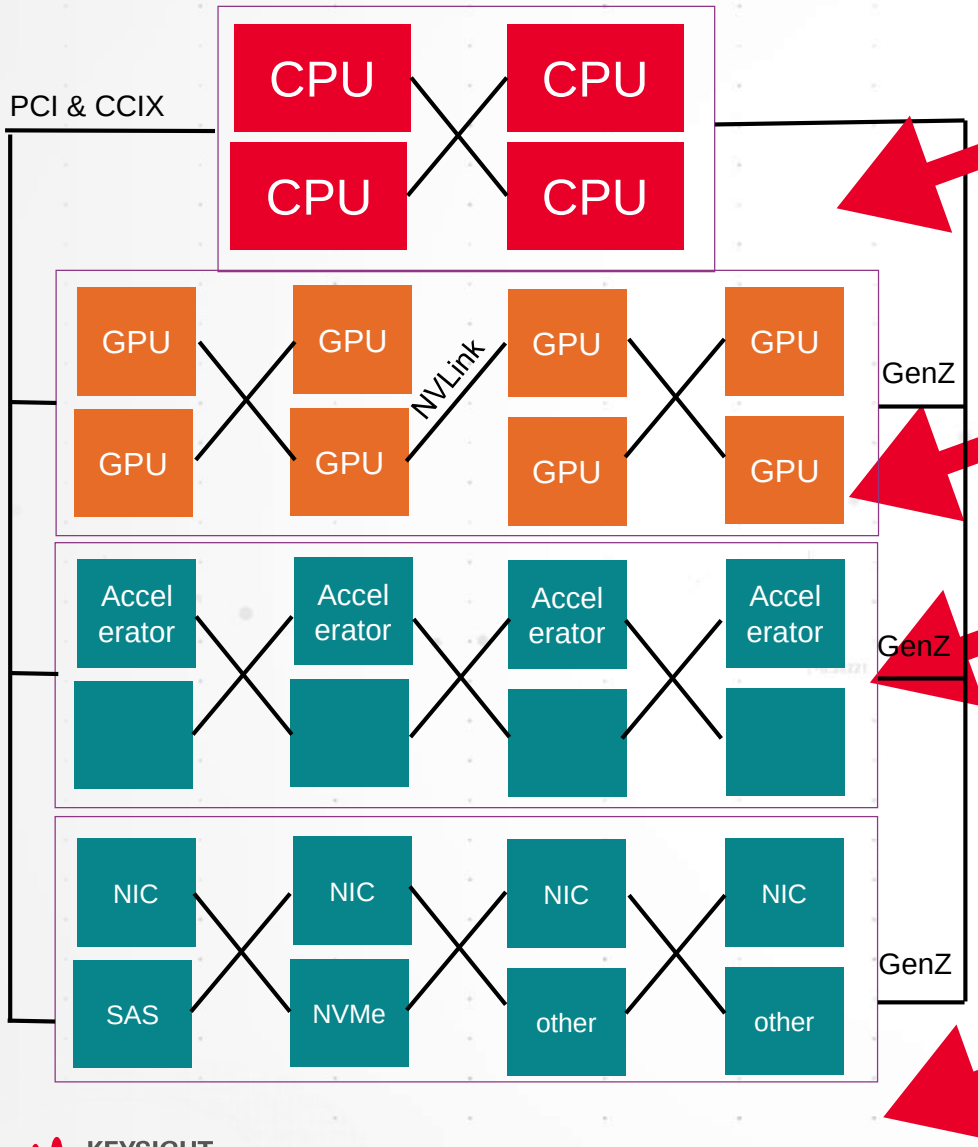




Changes in High-Speed Digital Buses

- DDR / LPDDR are moving to DDR / LPDDR5
- PCI Express (PCIe) is moving to PCIe Gen5
- Emergence of new computing standards (CCIX, GenZ, OpenCAPI, NVLink)
- Possible convergence of Type-C related buses

Server Advancements come 2019



PCI Express / CCIX offer a choice to designers for CPU to accelerator communication

NVLINK replaces PCI Express for GPU to GPU communication running at >25 Gbps

PCI Express / GenZ battle for communications between memory and CPU

DDR4 is replaced by DDR5 for memory. LPDDR4 gives way to LPDDR5

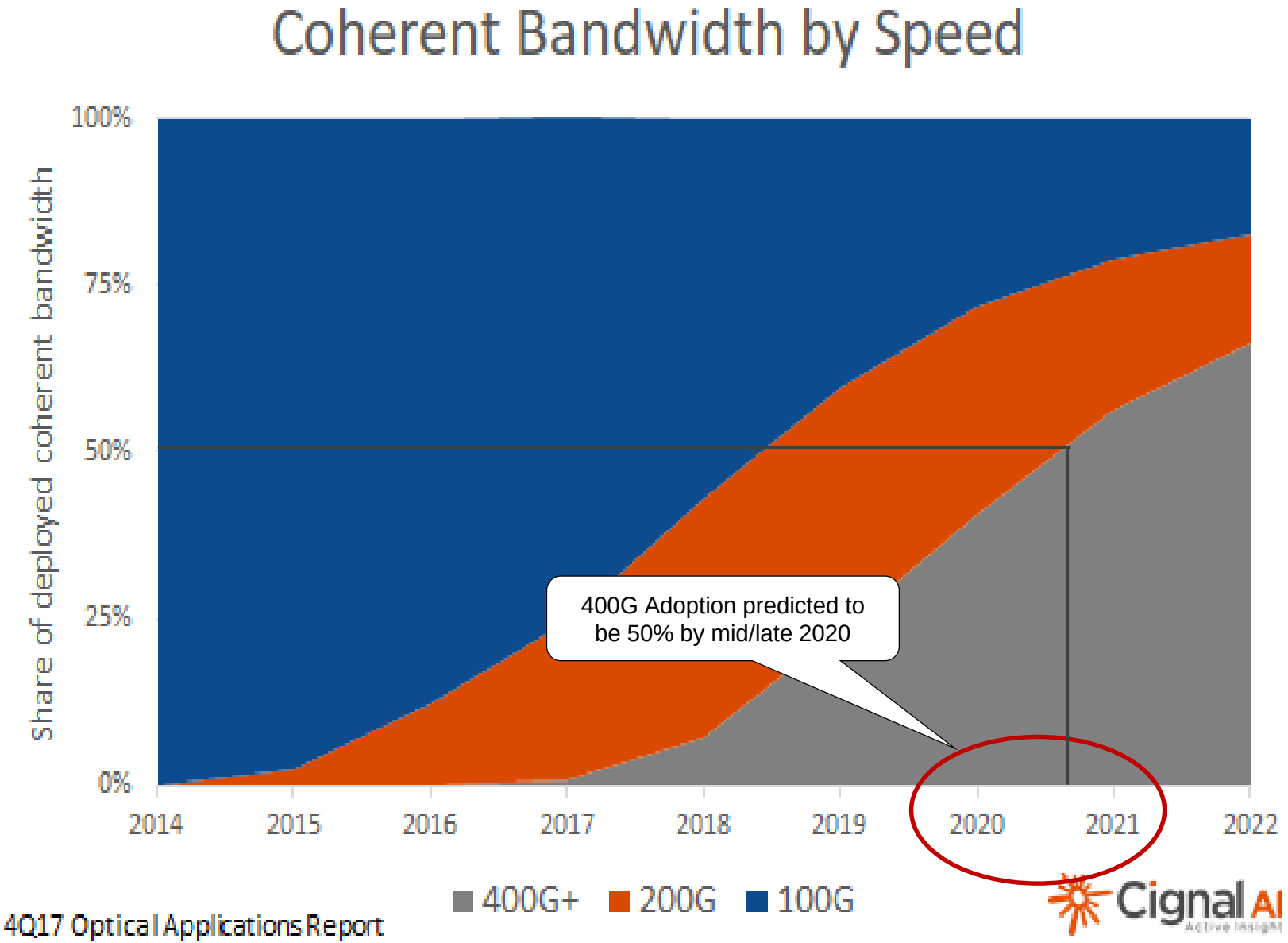
OPENCAPI and other technologies push for the fabric of the server

Intel announces new CXL open bus for coherency

PCI Express Gen5 – Moving to 32 Gbps

Drivers of PCIe 5.0 Performance

- High end networking
  - 400Gb Ethernet
  - Dual 200Gb/s InfiniBand
- Storage Networking
  - NVM Express (NVMe)
  - Big Data
- Increased IC I/O Speeds
  - Co-Processors (FPGA, GPU)
  - ASIC
  - IP
  - Artificial Intelligence Engines

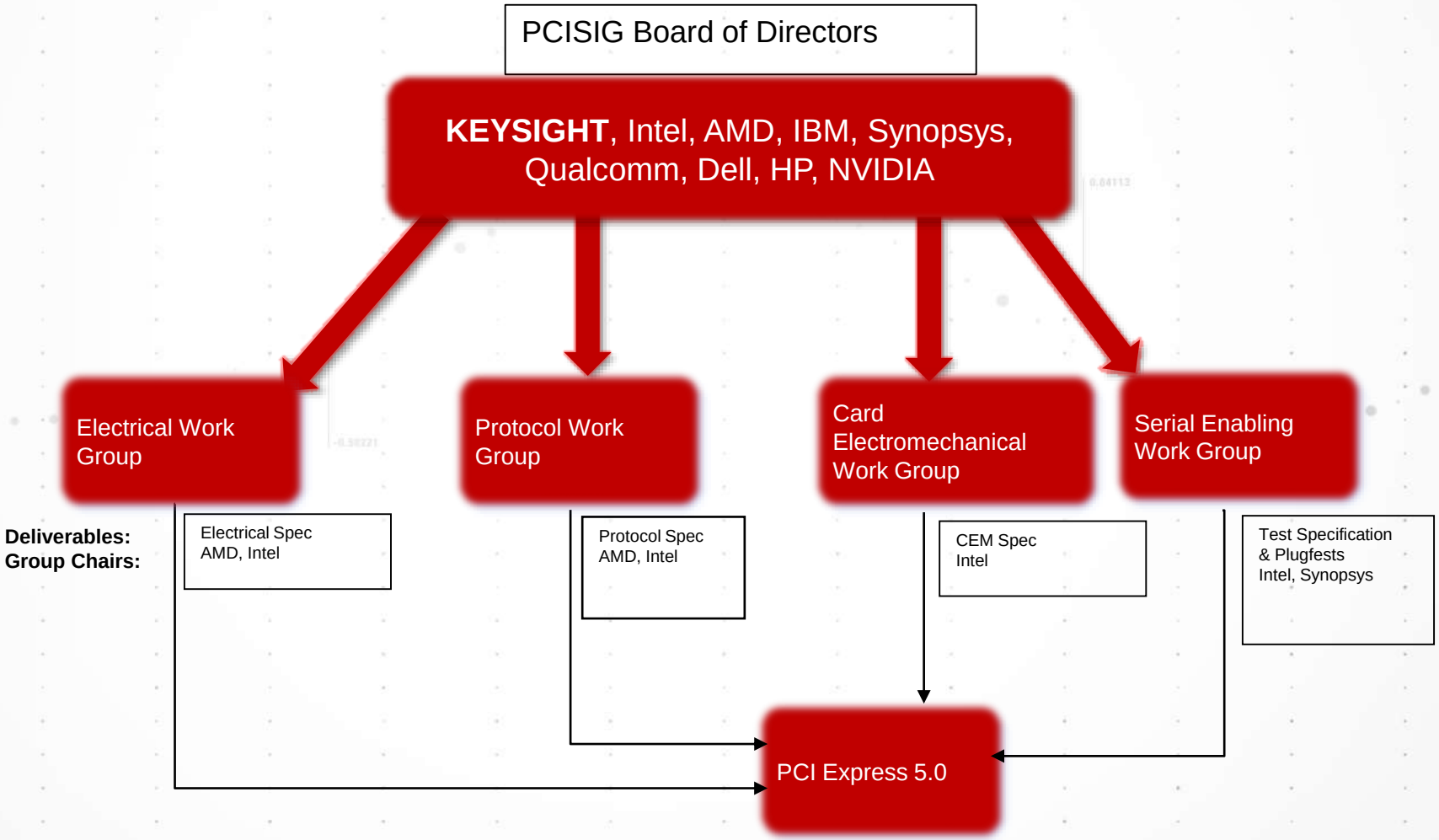


Industry Drives Higher PCIe Bandwidth Requirements

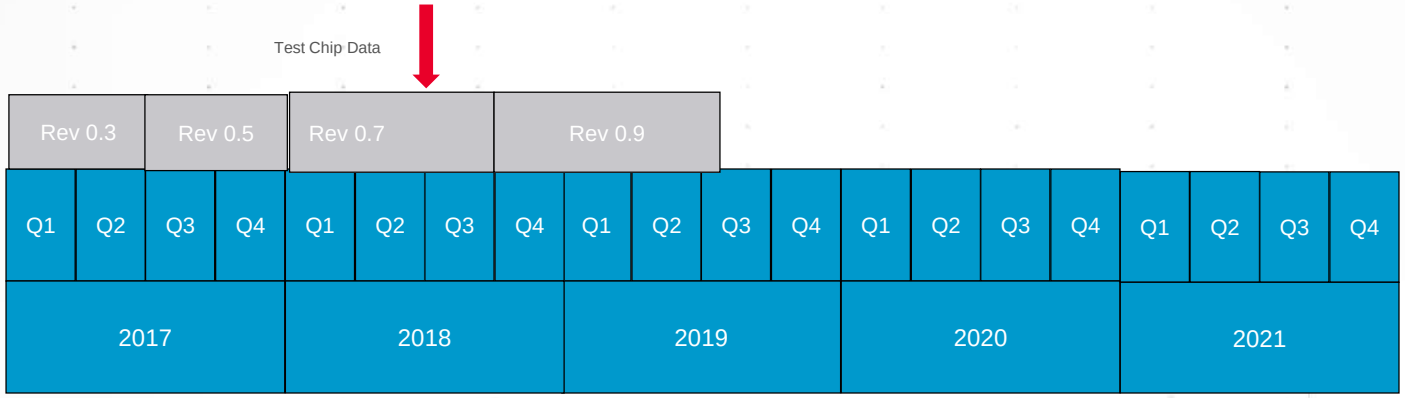
- PCIe 5.0 = 32 Gb/s
- Required for 400Gb Ethernet
  - This equates to 50GB bidirectionally
  - 16 lanes gives up to 64GB/s
  - Total full duplex BW = 128 GB/s
- CEM connector for PCIe 5 is planned to be backward compatible with earlier PCIe technologies.
- Tentative schedule for spec release in 2019

	Raw Bit Rate/Lane	Link BW	BW/Lane	Total x16 Bi-Directional Bandwidth
PCIe 1.x	2.5GT/s	2Gb/s	250 MB/s	8GB/s
PCIe 2.x	5.0GT/s	4Gb/s	500 MB/s	16GB/s
PCIe 3.x	8.0GT/s	8Gb/s	~1GB/s	~32GB/s
PCIe 4.x	16.0GT/s	16Gb/s	~2GB/s	~64GB/s
PCIe 5.x	32.0GT/s	32 Gb/s	~4GB/s	~128 GB/s

PCI Express Standards Development



PCI Express 5.0 Timeline (estimated)



- Draft 0.3 June 2017
- Draft 0.5 December 2017
- Draft 0.7 May 2017
- Draft 0.9 Stretch Goal end of 2018
- Final 1.0 “first half of 2019”

PCI Express Specifications and Scope

SELECT THE SPECIFICATIONS THAT RELATE TO YOUR NEED

Base Specification
- Contains all the system knowledge
- Can directly be applied to Chip Test

PCIe 5.0 BASE
Currently in review v0.9

Card Electromechanical (CEM) Spec
- Applies to Add-In Cards and Mother Boards
- Mitigates card manufacturer's need to study the base specification
- Increases reproducibility through PCI-SIG supplied test tools CBB and CLB (compliance base and load board)

Phy Test Specification
- Defines compliance tests of CEM spec in detail

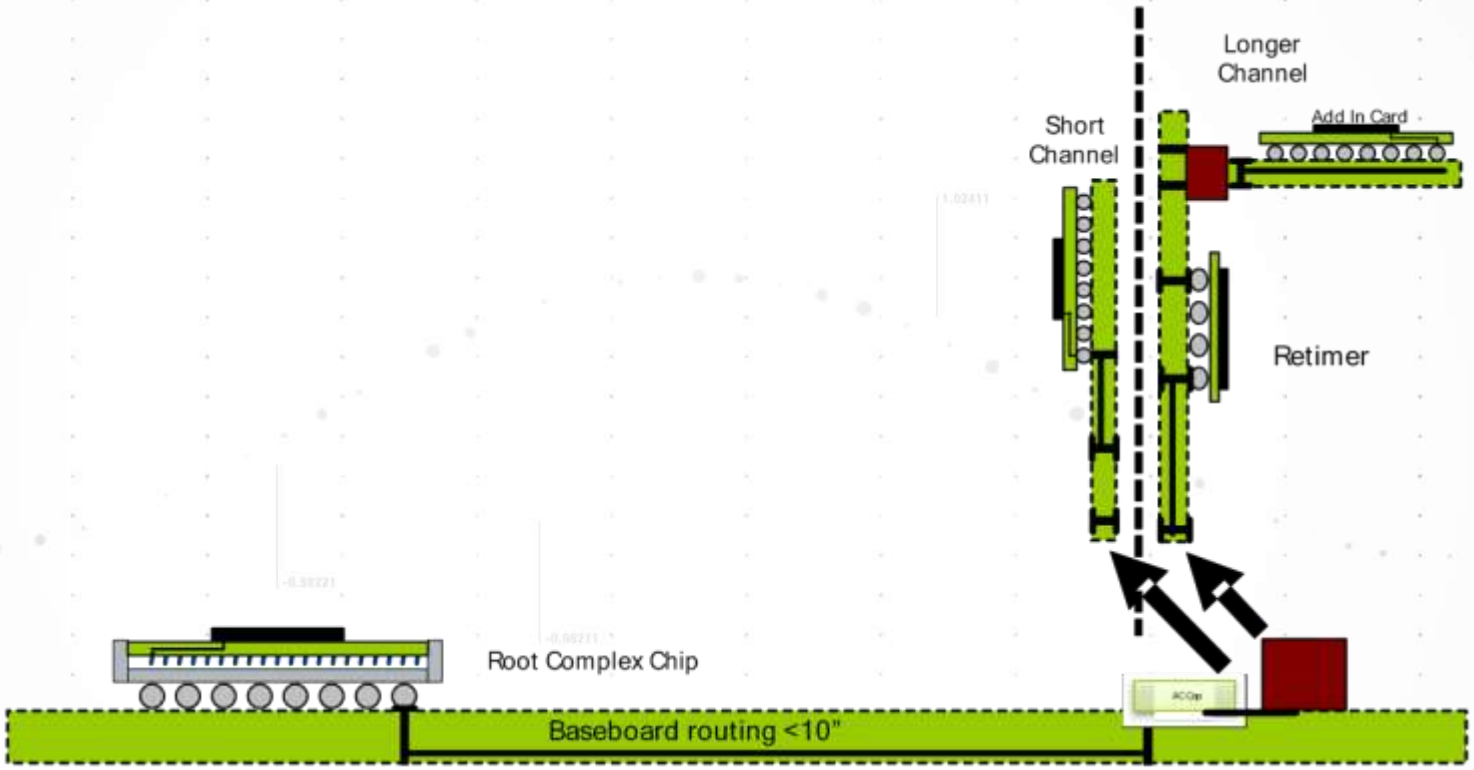
PCIe Gen5

GOALS

- BER target is 10e-12
- TX Presets P0-P10 to remain the same
- Backward compatibility with previous PCIe Gen1/2/3/4
- Same approach for TX and RX testing used for Gen4
  - Similar method for TX testing via de-embedding of breakout board traces
  - Similar method for calibrating the eye width and eye height as used with PCIe 4.0 (ISI based, fixed RJ)
- Same TX Voltage and Jitter parameters as Gen4

PCI Express 5.0 Channel Topology

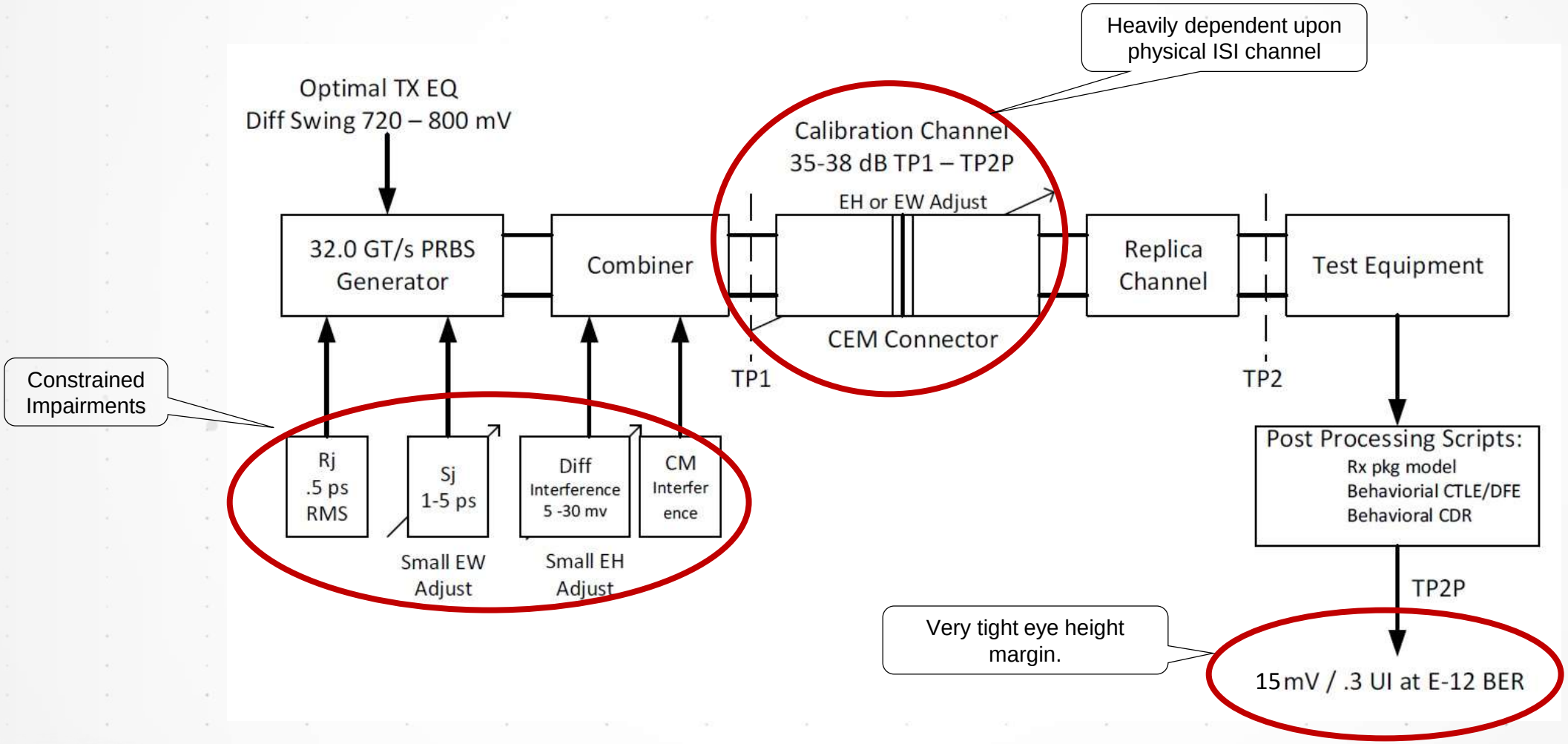
RETIMER REQUIRED WHEN LOSS EXCEEDS -36DB OR >1 CONNECTOR



- Estimated allowable loss: ~= -36dB @ 16GHz
- Root complex pkg loss allowance ~= -9dB @ 16GHz
- Add-in Card pkg loss allowance ~= -4dB @ 16 GHz
- Total AIC loss budget estimate = ~9dB @ 16GHz
- PCIe 5.0 CEM Connector loss budget ~= 1.5dB @ 16 GHz

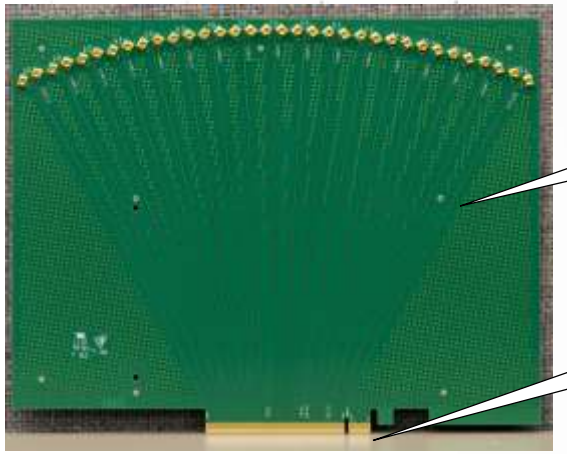
PCIe 5.0 32GT/s RX Calibration (BASE)

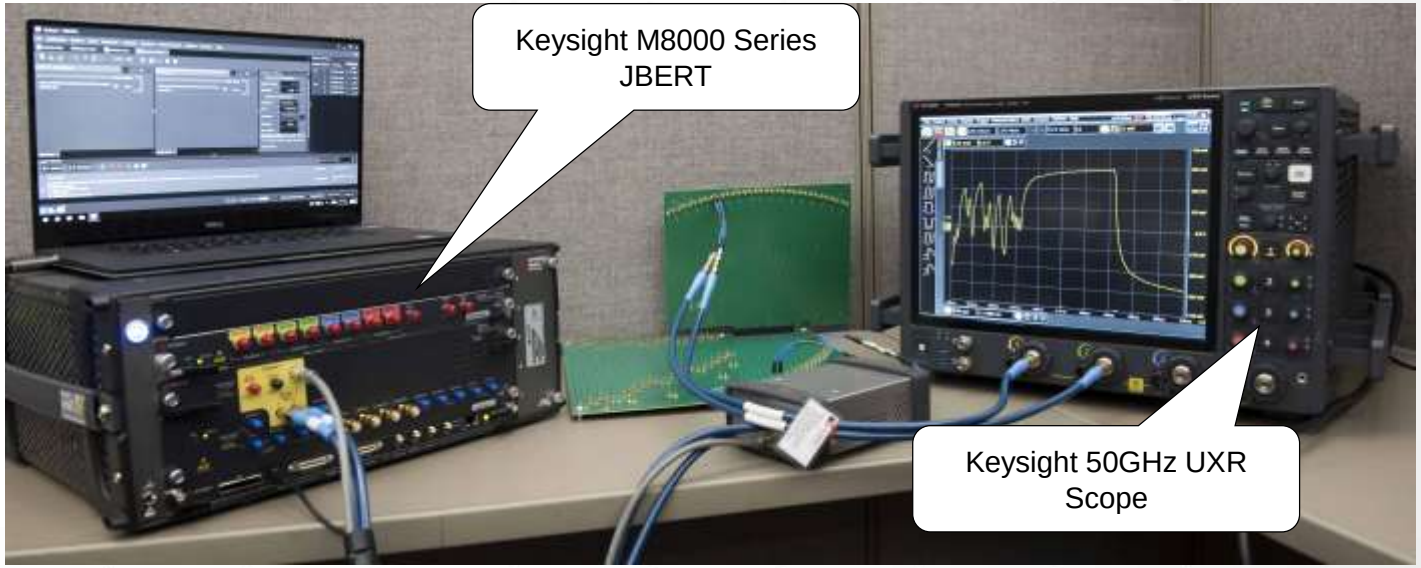
15MV EYE HEIGHT POST EQ (0.3UI EYE WIDTH) TARGET



PCIe 5.0 32GT/s RX Calibration (BASE) Example

TARGET TEST CHANNEL BOARD W/ NEW CEM SMT CONNECTOR





D9050PCIC New Features

MASTER YOUR BEST DESIGN

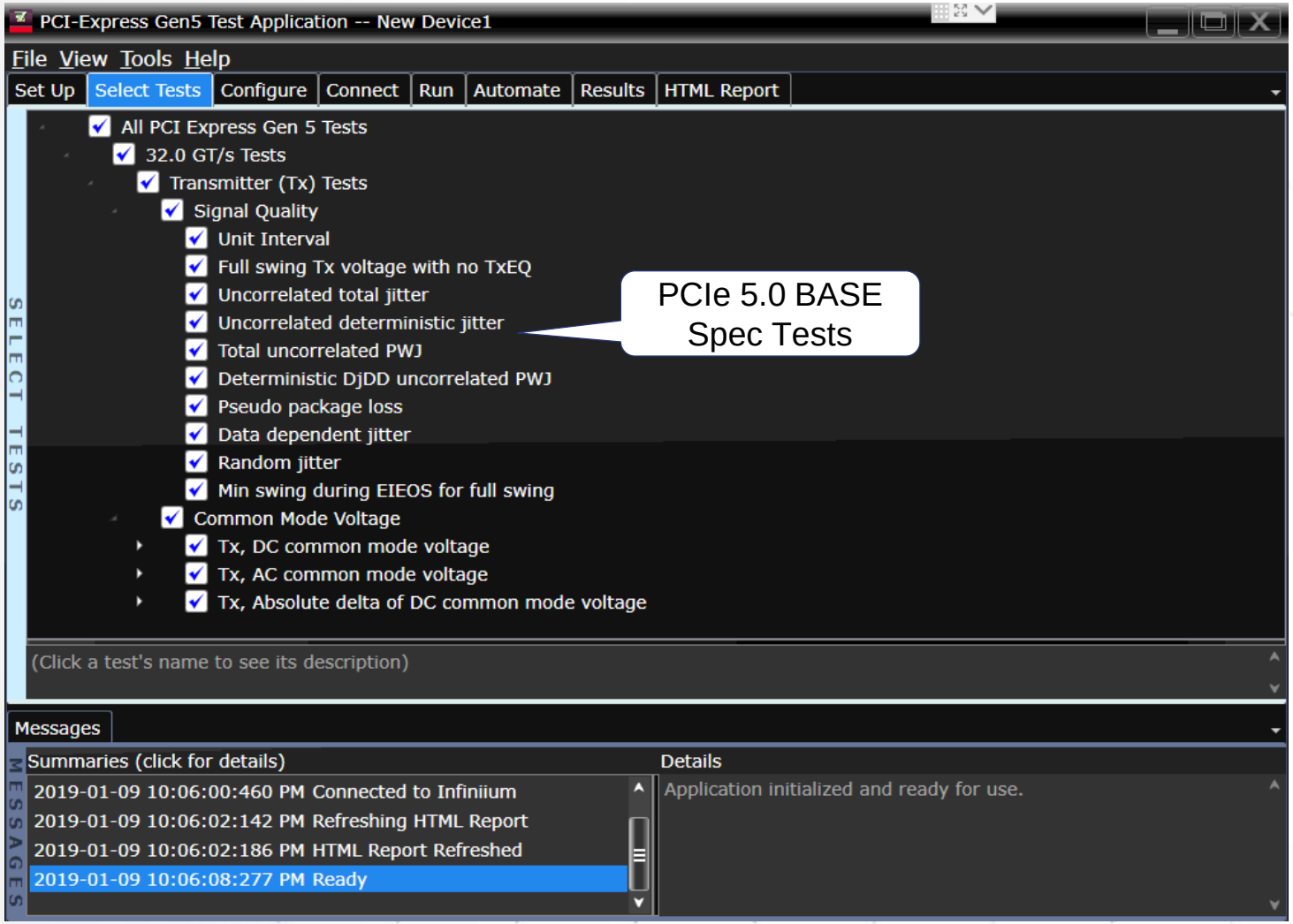
- Supports PCIe 5.0 BASE TX Testing at 32GT/s as well as 2.5G, 5G, 8G and 16GT/s (v0.9 BASE)
- Supports PCIe 5.0 Reference Clock tests (2.5G, 5G, 8G, 16G)
- Will Support CEM tests for endpoints and root complexes (2.5G, 5G, 8G, 16G, 32G).
- Integrated De-embedding of break-out channel with optional InfiniiSim
- Automated DUT control using an 81150/60A Pulse Generator ARB.
- Enhanced Switch Matrix supporting arbitrary lane mapping
- Minimum oscilloscope BW: 50GHz

clude the additional signal distortion caused by the behavioral Receiver package. If a compliance pattern waveform is used then all stresses except $V_{RX\text{-}CM\text{-}INT}$ are turned on if a step is used then all stresses are turned off. Then the resulting signal is recovered by means of Rx equalization, and a behavioral CDR function, resulting in an equivalent eye. The requirements for the waveform post processing tool used for the EH/EW calibration are described further in Section 8.4.2.1.1 Post Processing Tool Requirements .

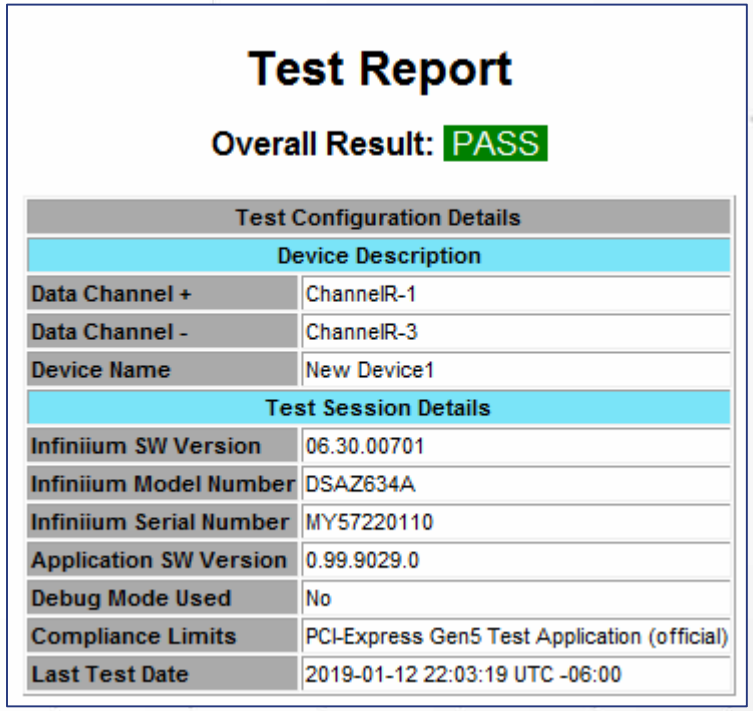
As the calibration procedure of the signal generator output contains steps where the generator is connected directly to measurement instrumentation, the transition time of the output waveform can be very fast. Therefore, it is important that the bandwidth of instrumentation used to calibrate the generator be matched appropriately to the edge rate of the generator output. This specification requires the use of a generator for 16.0 GT/s testing whose outputs have a rise time of 14 ps-19 ps (20% / 80%) which also requires a minimum oscilloscope bandwidth of 25 GHz. This oscilloscope bandwidth is also the minimum required bandwidth for transmitter measurements at 16.0 GT/s. For 32.0 GT/s testing the specification requires the use of a generator whose outputs have a rise time of 7.5 - 15.0 ps (20%/80% measured with P4) which requires a minimum oscilloscope bandwidth of 50 GHz. This oscilloscope bandwidth is also the minimum required for transmitter measurements at 32.0 GT/s. A minimum oscilloscope sampling rate that captures at least 5 samples per unit interval is required for all data rates.

Keysight D9050PCIC PCIe 5 (Gen5) TX Test Application

5.0 BASE SPEC TESTS



PCIe 5.0 BASE Spec Tests



Test Report

Overall Result: PASS

Test Configuration Details	
Device Description	
Data Channel +	ChannelR-1
Data Channel -	ChannelR-3
Device Name	New Device1
Test Session Details	
Infiniium SW Version	06.30.00701
Infiniium Model Number	DSAZ634A
Infiniium Serial Number	MY57220110
Application SW Version	0.99.9029.0
Debug Mode Used	No
Compliance Limits	PCI-Express Gen5 Test Application (official)
Last Test Date	2019-01-12 22:03:19 UTC -06:00

PCIe 32G/16G/8G RX Test Setup

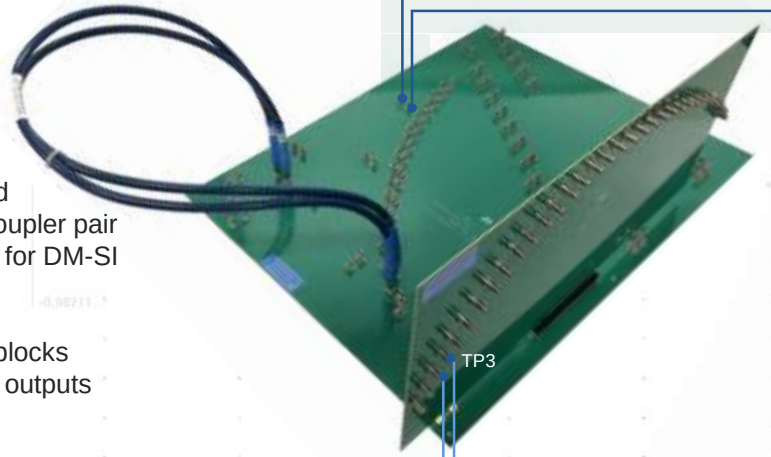
SETUP USING M8000 SERIES 64G HIGH-PERFORMANCE BERT

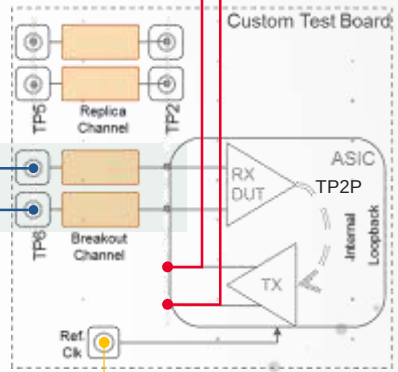
New Interference Source module M8054A for DM-SI & CM-SI

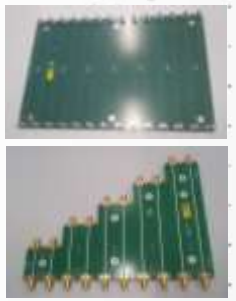
M8046A with integrated CDR option 0A4 / 0A5 and equalization with new option 0S2 SKP OS Filtering PCIe 8G/16G/32G and CCIX 20G/25G and option 0S1 Interactive Link Training PCIe 8G/16G/32G

New matched broadband coupler pair M8045A-803 for DM-SI & CM-SI

N9398F DC blocks used coupler outputs







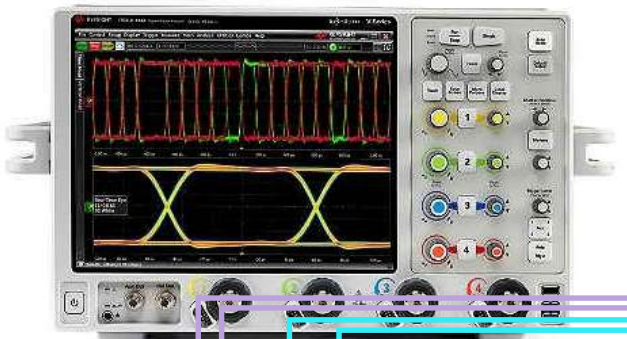
Substitute PCIe 5 Base Channel Boards by new ISI traces M8049A-003 & M8049A-001

M8049A-001 offers traces from approximately 1.5dB to 3.5dB in 0.5dB steps.

M8049A-003 offers traces from 10dB to 22dB in 2dB steps, IL measured @ 16GHz

SAS-4 RX Test Setup – Stressed RX Jitter Tolerance

SETUP USING M8040A + M8054A INTERFERENCE SOURCE



New Interference Source module M8054A for DM-RI & CM-RI
N9398F DC blocks used at interference source outputs

Matched coupler pair M8045A-803 for CM-RI
N9398F DC block

DM-RI

CM-RI

ISI trace

Simulate SAS-4 test channel by new ISI traces M8049A-001, -002 and -003

M8049A-001 offers traces from approximately 1.5dB to 3.5dB in 0.5dB steps.

M8049A-002 offers traces from approximately 3dB to 9dB in 1dB steps.

M8049A-003 offers traces from approximately 10dB to 22dB in 2dB steps.

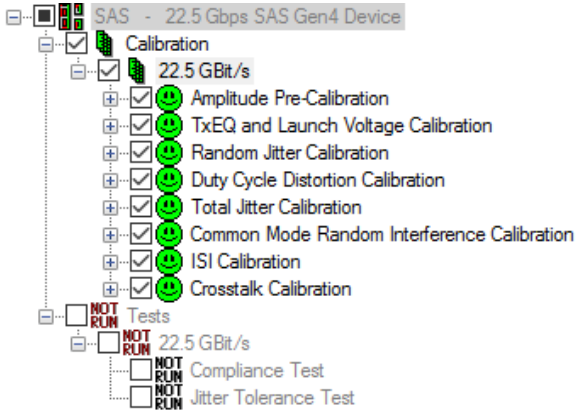
IL @ 16GHz

Matched directional coupler pair M8045A-802 for DM-RI for ICN

To support SAS-4 RX testing:

- M8040A system with 1 pattern generators and one error detector
  - All jitter sources
  - Multi-tap de-emphasis is built-in
  - CDR as well as CTLE are built-in
- AWG used as common mode random interference source
- AWG used as differential mode random interference source
- SAS RX Calibration and Test SW – N5991SA4A (June/July 2019)

2nd data lane sends 1010 as clk reference for M8040A



PCIe Tx EQ De-Emphasis and Rx Link Equalization

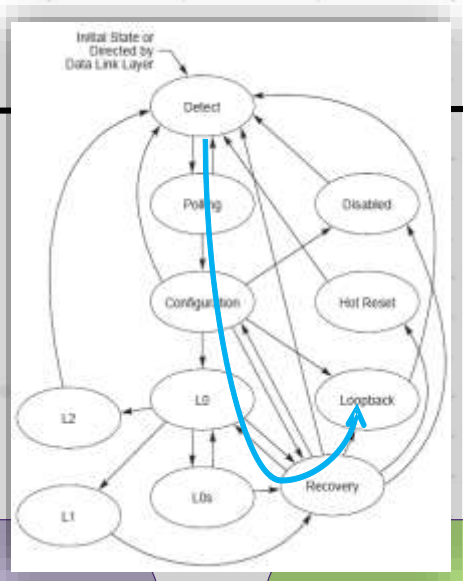
- Rx LinkEQ Test
- No Jitter Tol. Test

Request TxEQ setting

Algorithm for determining equalization and de-emphasis

Means for measuring signal quality

Control



De-Emphasis Controller

CDR

Equalization

2 Tap DFE

1 order CTLE

Receiver

Transmitter

3-Tap De-Emphasis

- Initial Tx EQ Test
- Tx LinkEQ Test
- Preset Test

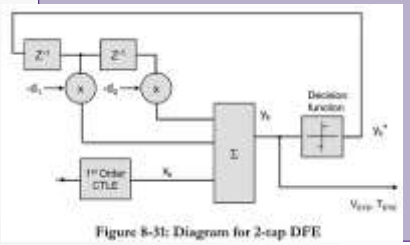


Figure 8-31: Diagram for 2-tap DFE

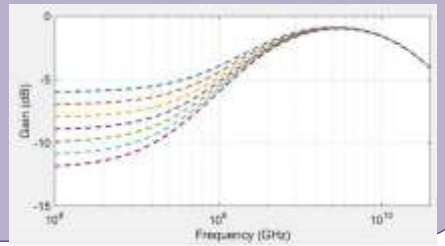
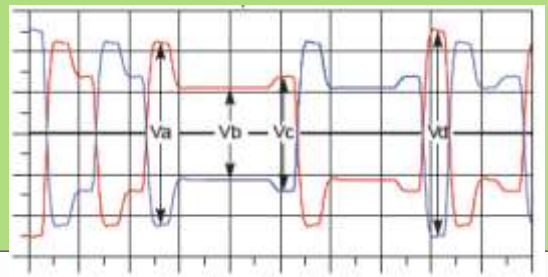


Figure 8-29: Loss Curves for 16.0 GT/s Behavioral CTLE



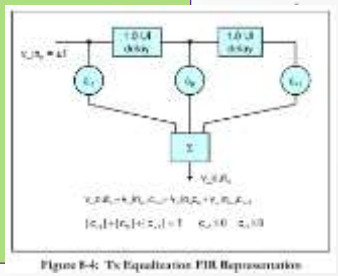
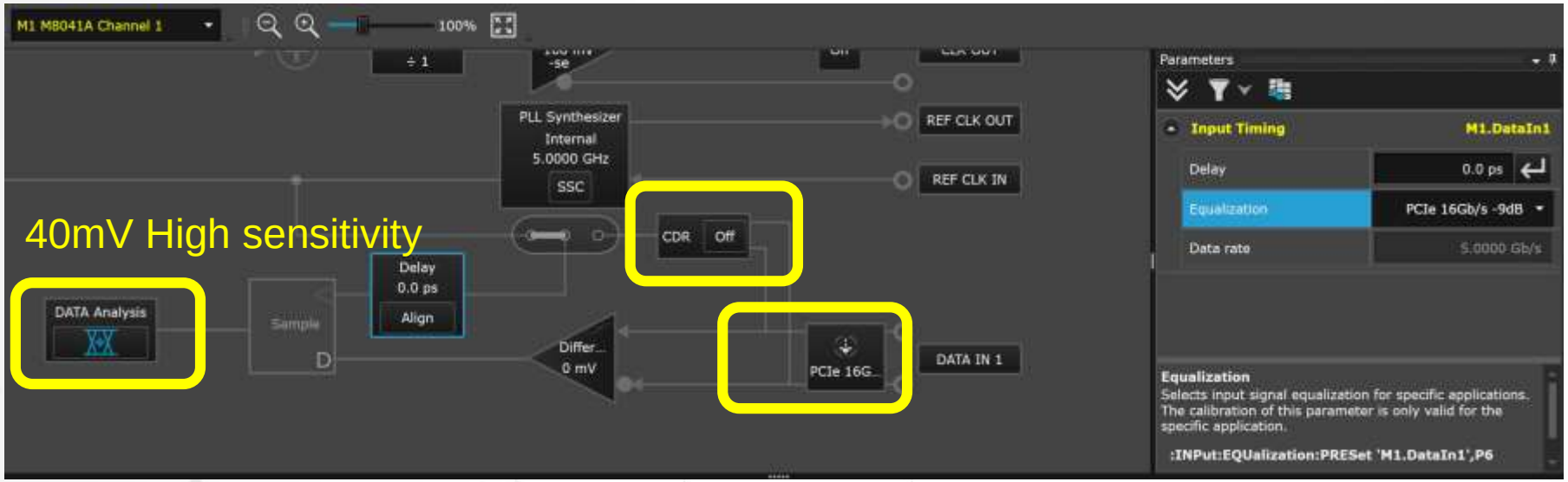


Figure 8-4: Tx Equalization FIR Representation

Key of Successful PCIe Gen4 Rx testing: CTLE & CDR

CHALLENGE IN CEM LONG TRACE

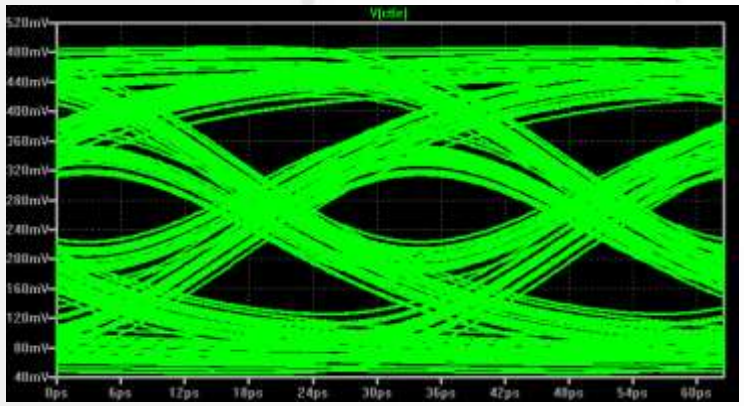


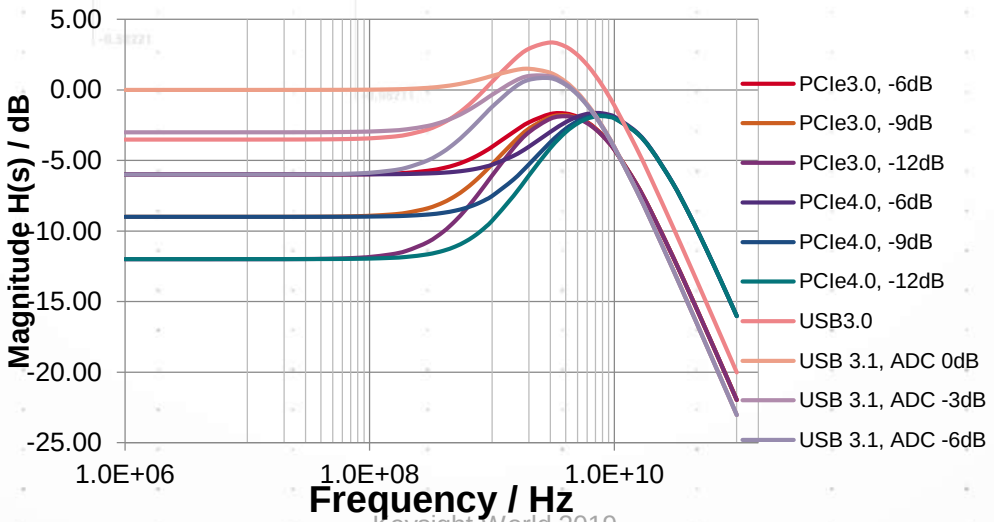
40mV High sensitivity

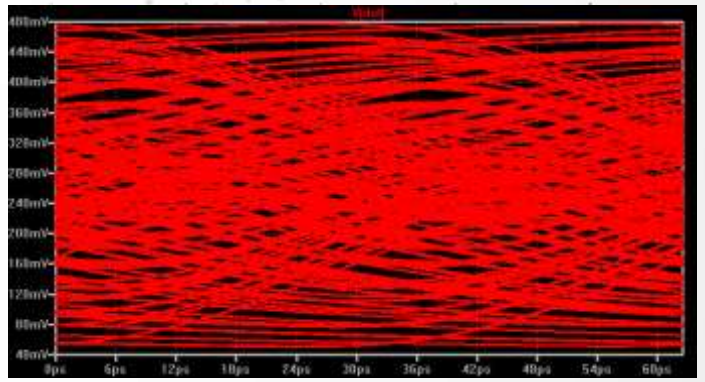
The Key with Error Free:

- No additional loss with integrated CDR

Tx signal after > 10 inches







DDR5 - Bring on Receiver Testing for DRAM

The New Age - DDR Signaling Evolution

DDR6?

"Light Speed"
- Complex modulation?

IQ Threshold

DDR5

"Hyper Speed" - No visible eye
- Pre/De-Emphasis, Impulse response

Impulse Response Threshold

DDR4

"Serial Speed" - Clock often embedded
- Eye Diagrams/Masks
- Error rates, Jitter/Noise

Jitter & Noise Threshold

DDR3

DDR2

DDR

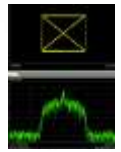
"High Speed" - Transmission lines
- Timing
- Impedance / load curves

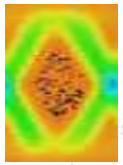
Signal Integrity Threshold

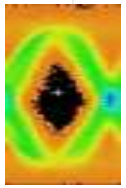
Static RAM

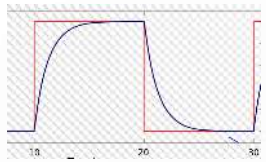
"Low Speed" - Auto place and route
- Fanout
- Capacitance

Signal

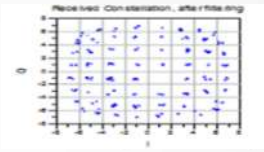


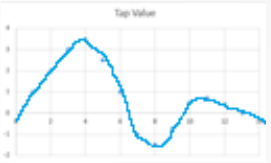


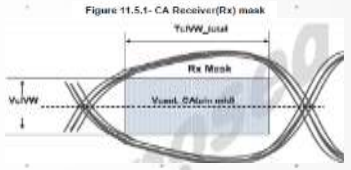


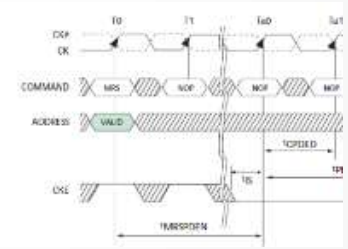


Specification





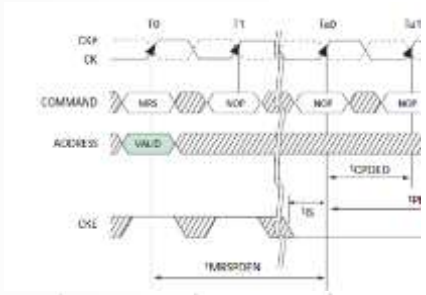


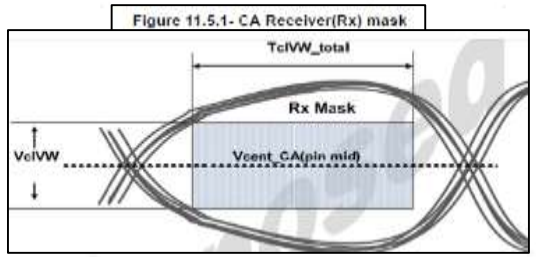


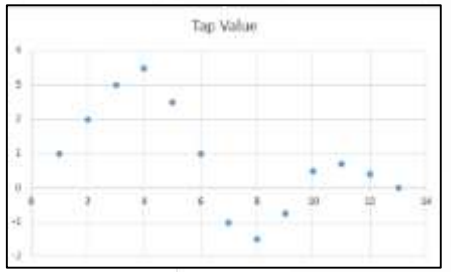
Symbol	Parameter
$V_{OH}(DC)$	DC output high measurement level (for IV curve linearity)
$V_{OM}(DC)$	DC output mid measurement level (for IV curve linearity)
$V_{OL}(DC)$	DC output low measurement level (for IV curve linearity)
$V_{OH}(AC)$	AC output high measurement level (for output SR)
$V_{OL}(AC)$	AC output low measurement level (for output SR)

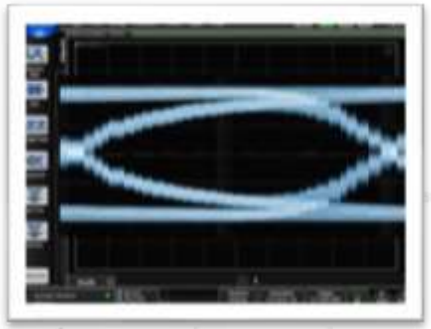
Just When You Thought it Was Safe…

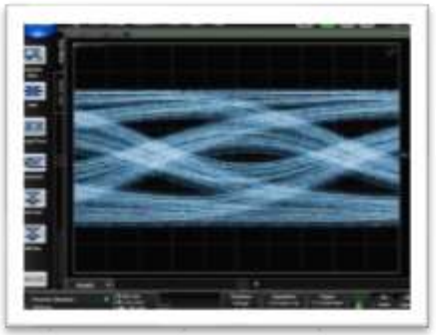
RULES ARE CHANGING WITH EVERY GENERATION- DDR5 IS COMING

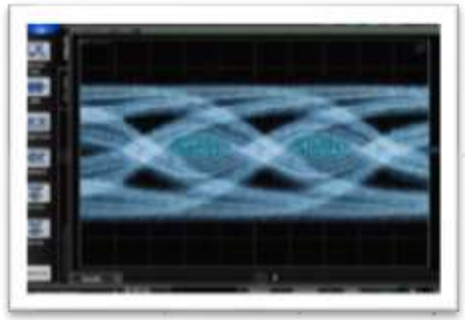












DDR1/2/3

High-Speed Digital

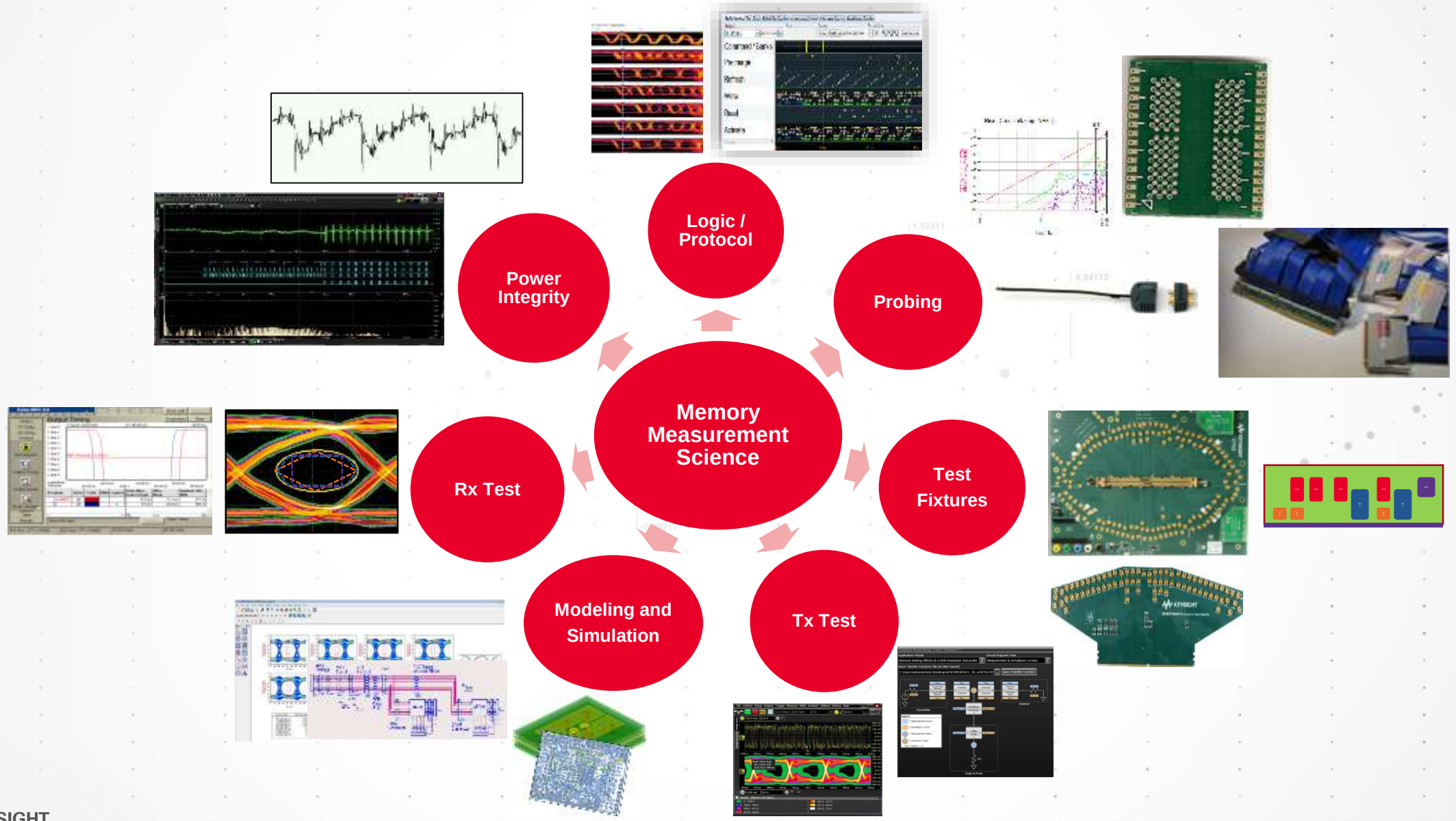
DDR4

Serial Speed

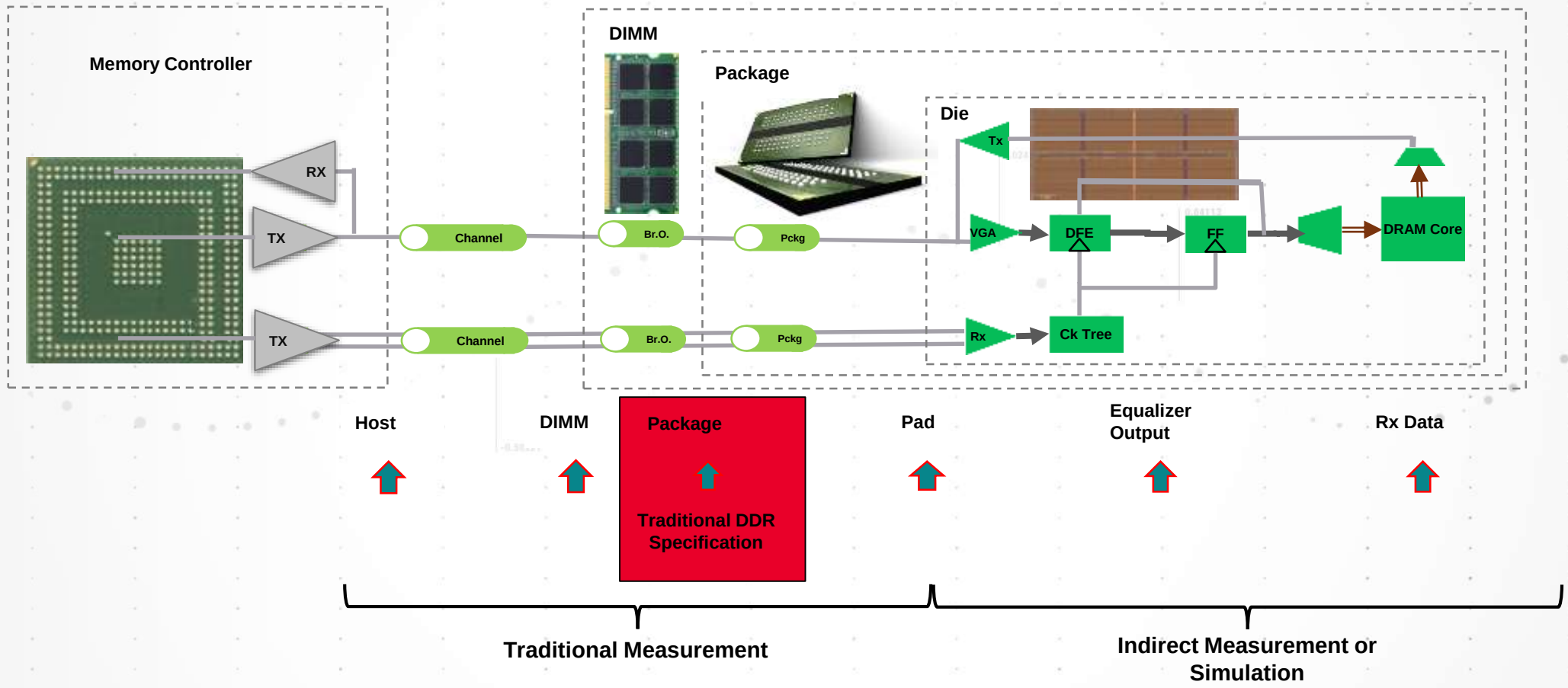
DDR5

Hyper Speed

Measurement Science Domains



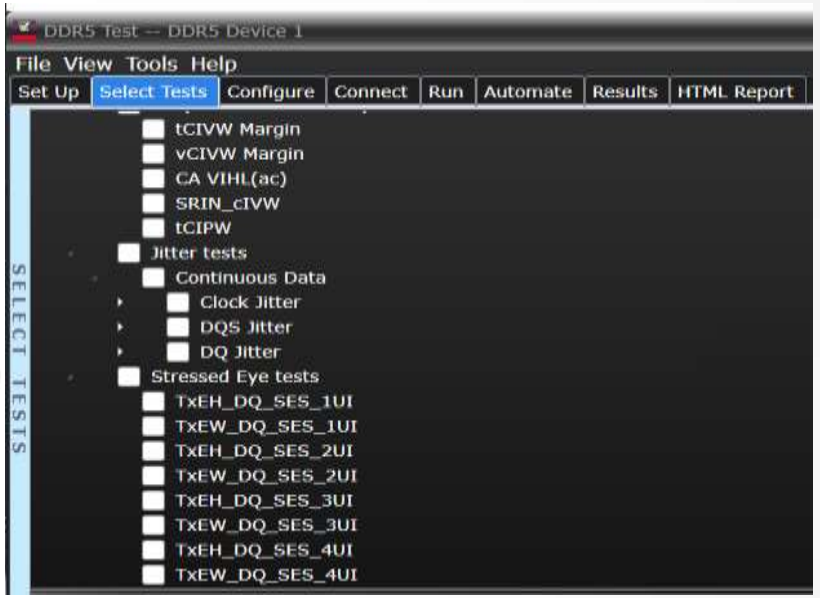
DDR5 Specification and Measurement

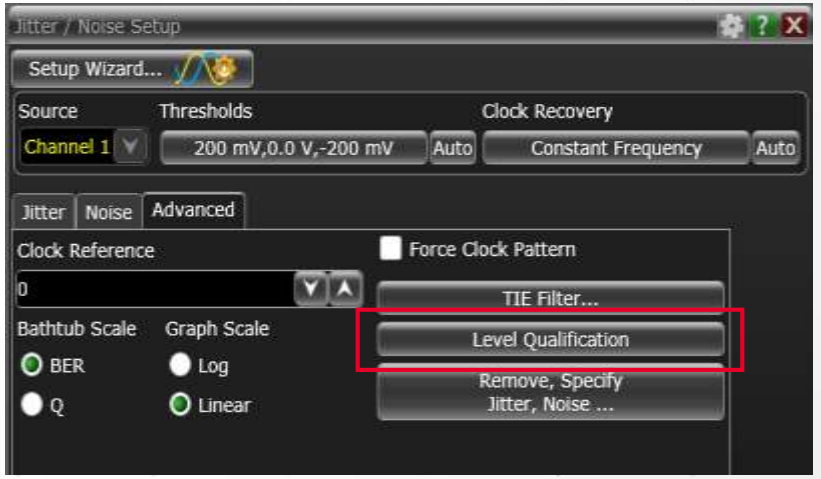


New Tx Measurement Science

- **Characterization of CLK, DQS, DQ**
  - Mismatch between DQS and DQ (tDQS2DQ)
  - Rj/Dj jitter separation to allow designer to identify design issues
    - Rj – source of noise
    - Dj – source of crosstalk and duty cycle distortion

- **DQ stressed eye test**
  - Eye height and eye width measurement
  - DFE for higher data rates
  - Account for instrument noise





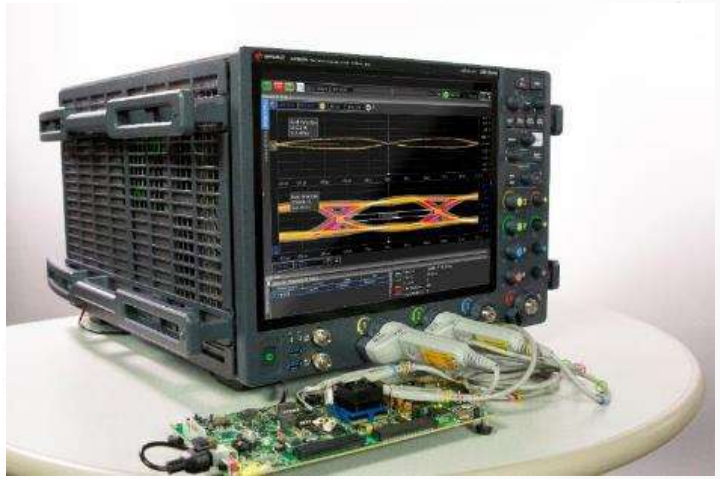
New Read/Write Separation Method

Command protocol decode with mixed signal oscilloscopes

- Use command truth table to decode read and write commands
- Pros: More robust
- Cons: Requires additional signals. More signals -> More loading

Read or write only control

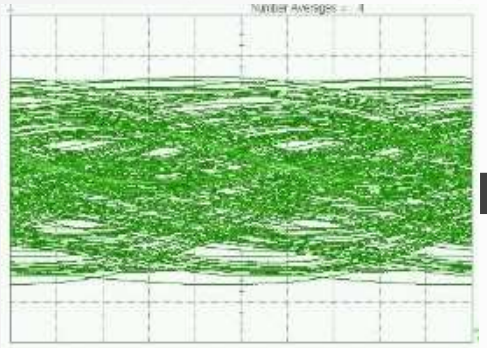
- System designer has control over data transition type through SOC, FPGA.
- Pros: Eliminates the need for read/write separation
- Cons: Not an option for system integrators



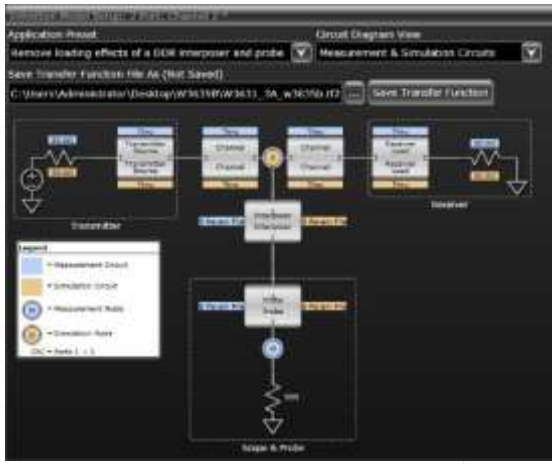
De-embedding and Equalization

VIRTUAL PROBING METHOD AT RX EQ

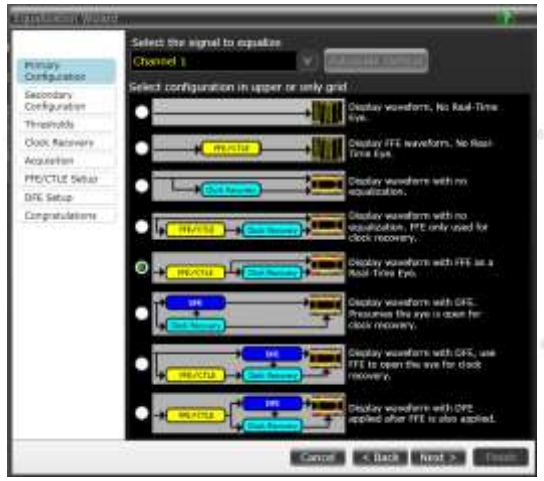
DRAM Ball



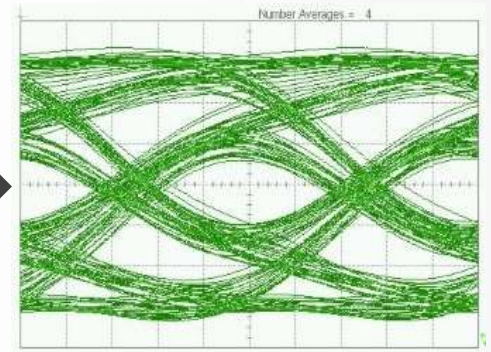
Package Model

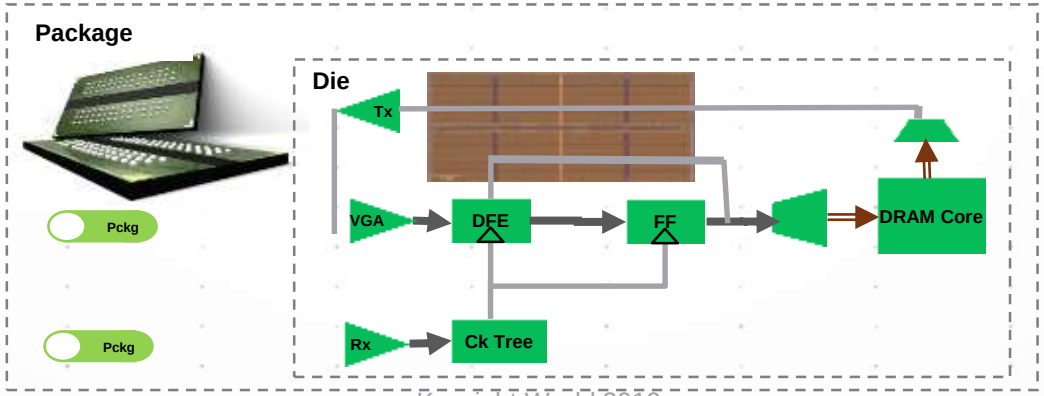


Rx Model



Scope display





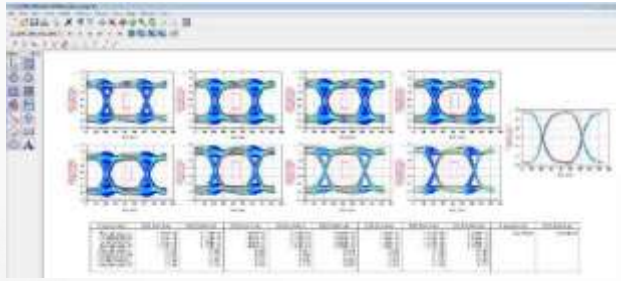
Probing Technology

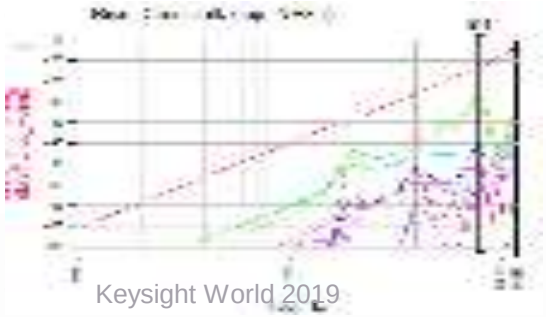
Higher speeds require smaller, higher density probing with lower loading

Reduced loading
DIMM interposer
for logic analysis

Via back side
probing (DDR4)

BGA interposer





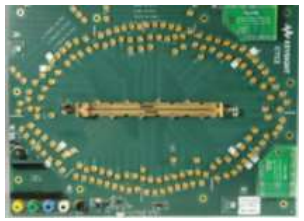
Detailed simulation modeling
correlated to physical prototypes

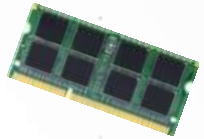
DDR/PCI Express Comparison

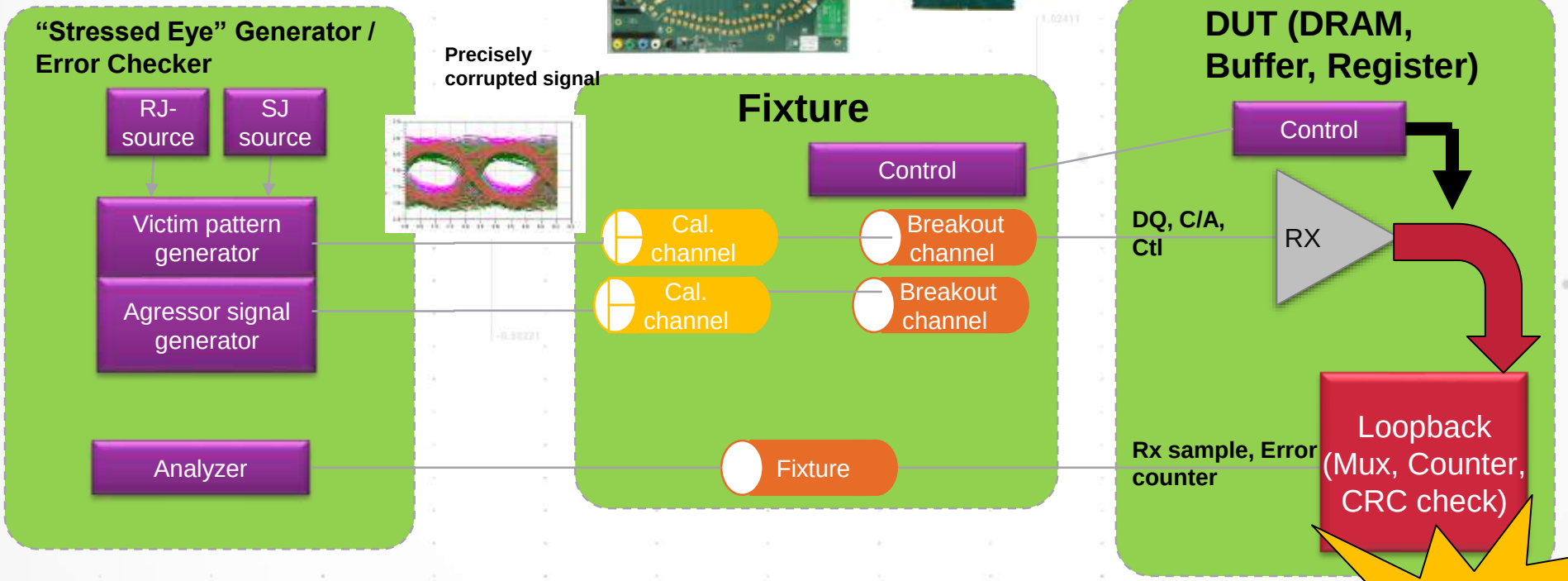
- **Full rate clock**
  - No Rx PLL needed
- **Wide, single ended**
  - Xtalk dominated vs. loss dominated
  - Increased emphasis on multi-channel Rx test impairments
- **Bidirectional**
  - Extra parasitic loading when Rx and Tx share pins
  - There is no clean way to terminate the bus
  - Reflections due to ODT changes, adjacent rd/wr bursts
- **Bursty**
  - ISI impact differs on preamble, first bits, rest of burst
- **Rx must work using “slow” DRAM process**
  - Full rate loopback impossible
  - “Dumber” Rx increases test complexity

Receiver Testing Begins in DRAM

RX TESTING AND EQUALIZATION





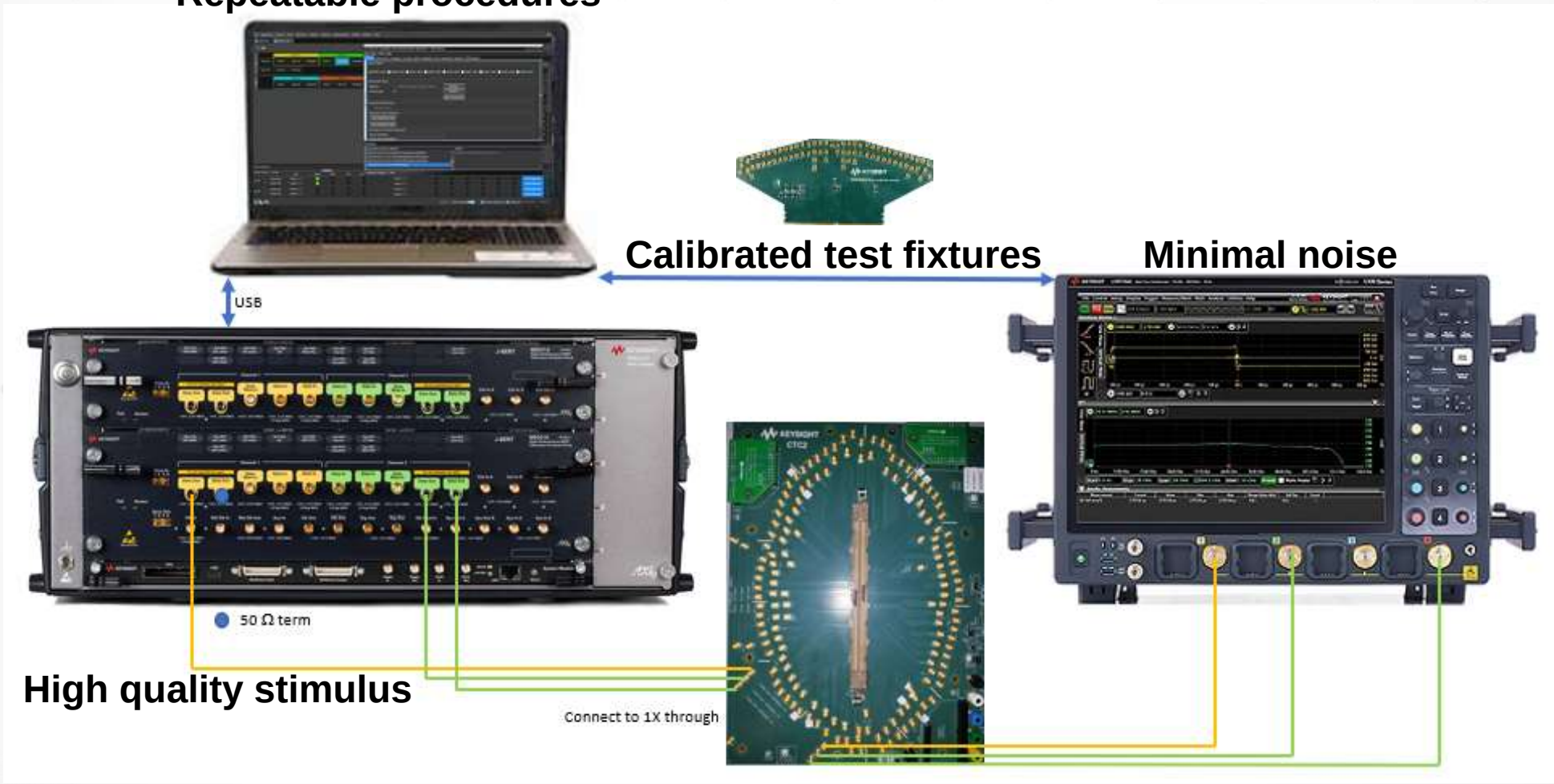


- Challenge the DRAM with a "stressed" signal / known pattern
- Compare expected vs. actual
- Repeat 3e10 to 3e16 times (depending on BER goal)

New to DRAM

Receiving Testing Needs

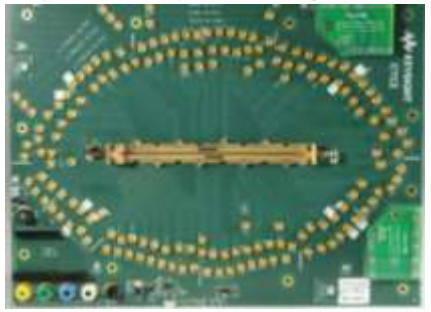
CREATE A WELL CONTROLLED TEST SETUP



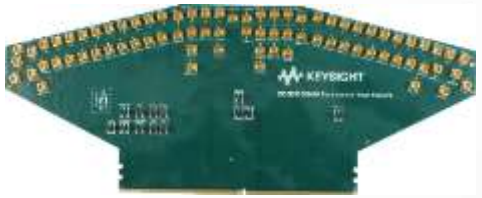
Device and System Testing Fixture

AUTOMATED DEVICE CHARACTERIZATION AND GOLDEN CHANNEL

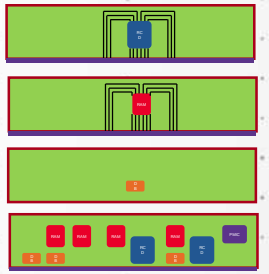
- **Channel Test Card "CTC2"**
  Socket for plugging in DIMM test card for calibration, DIMM module to be tested or device test board. SMP connectors to provide access to CA/CTRL, strobe, and data
- **Golden Channel Modeling Board (or ISI Board)**
  Contains different length standard compliant channels
- **Device Test Card**
  Plugs into CTC and can be used by DRAM/memory/register manufacturers to test their chips
- **Fully Passive DIMM Test Card**
  Plugs into CTC. Signal breakout board for clock, CA, CTRL, strobe and data. Used for signal inspection and Rx stress signal calibration. Signals are routed to SMP connectors.



CTC2



DIMM Test Card



Potential Device Test Cards


KEYSIGHT
WORLD 2019