Tektronix Innovation Forum

Enabling Innovation in the Digital Age

High performance storage systems technology update and introduction to 12G physical layer validation

Presenter: John Calvin

Tektronix®

Agenda: High performance storage systems technology update and introduction to 12G physical layer validation

- Introduction
  - Industry Timeline
  - SATA 3.1 Spec Revisions
    - UTD 1.4.2 overview
    - ECN50: Asymmetric Amplitude
    - ECN51: Framed Composite waveform
    - SATA/PCIE: SATAExpress Convergence
  - Measurement Considerations at ~12G
    - 12+ G Design Problem
    - Crosstalk
    - BUJ
    - WDP
    - Signal Amplitude
    - Instrumentation

REFERENCES

(1) Richard Mellitz: Intel: T10/ 11-275r0 SAS-3 12Gbs Transmitter Device Test Proposal

(2) Doron Lapidot: Tyco: T10/10-219r0 SAS 3.0 B-t-B Connector & Cable assembly Channel Performance @ 12Gbps "Modeling, Measurements & Simulations for BER compliance with multi Aggressor System Interconnect"

(3) Mickey Felton: EMC: T10/11-239r0 Channel compliance points and lengths

(4) Kevin Witt: Maxim: T10/11-221r4 SAS-3 Electrical Spec (Draft)

(5) Mathieu Gagnon : PMC-SIERRA: T10/11-008r3 SAS-PHY: SAS3_EYEOPENING update

(6) Mladen Luksic: SATA-IO: IW12 Roadmap Update

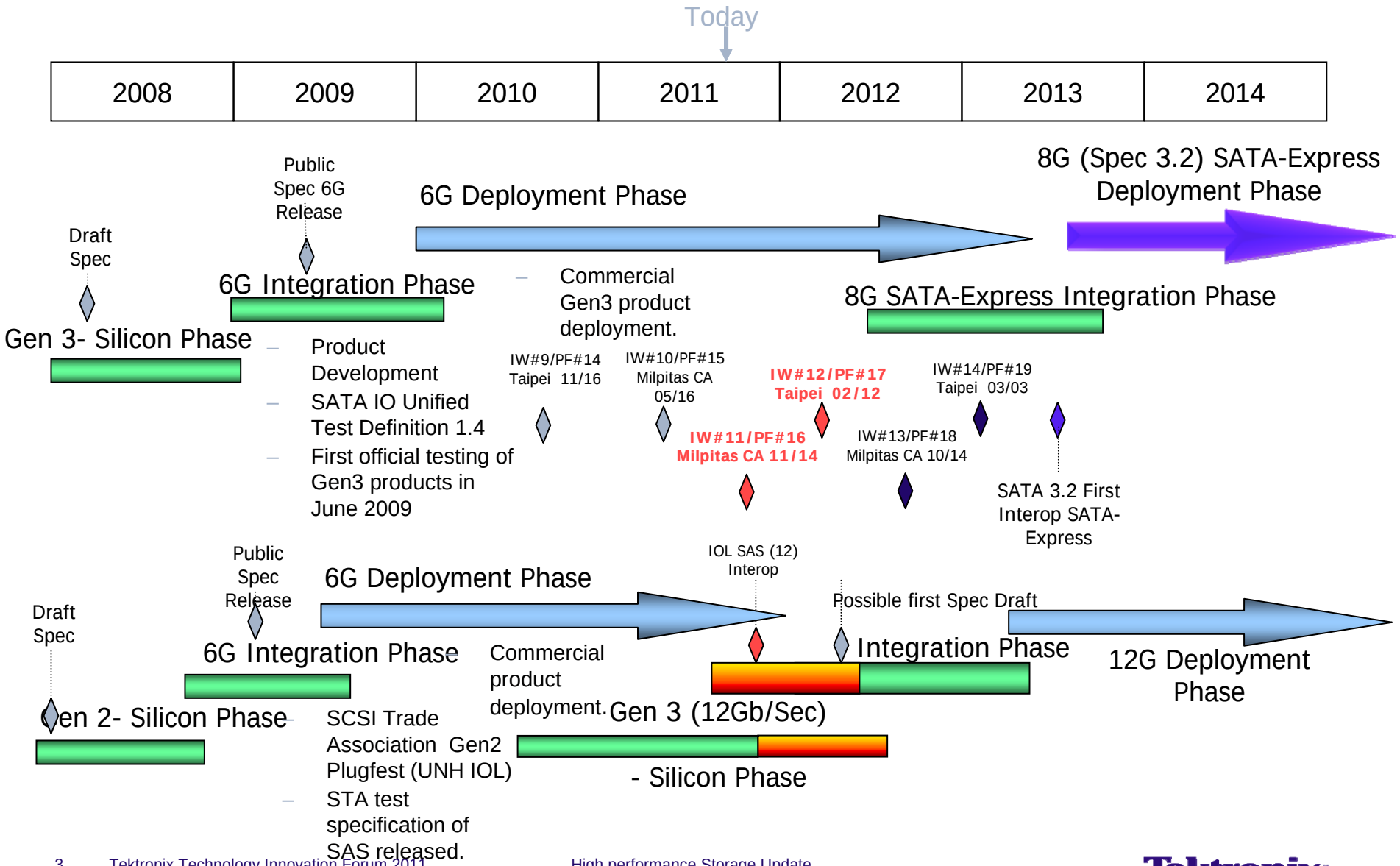
General REFERENCES:

[1] IEEE is 25Gb/s on-board signaling Viable? KAM et al.: IEEE Transactions on advanced Packaging, Vol. 32, No. 2, May 2009.

[2] IEEE CMOS SerDes core with feed-forward and decision-feedback equalization . T. Beukem et al.: IEEE J. Solid-State Circuits, vol. 40, no. 12, pp. 2633–2645

[3] Alcatel-Lucent: DSP & FEC: Towards the Shannon Limit Timo Pfau ECOC'09 | WS1: DSP & FEC | Sept. 20, 2009

Storage Timelines and Solutions Development



SATA 3.1/ UTD 1.4.2 Spec Revisions

It's the measurements ..

SATA UTD 1.4.2 Test Requirements

Phy Transmit Signal Requirements	
TSG-01 : Differential Output Voltage	
TSG-02 : Rise/Fall Time	revised
TSG-03 : Differential Skew	
TSG-04 : AC Common Mode Voltage	revised
TSG-05 : Rise/Fall Imbalance	
TSG-06 : Amplitude Imbalance	obsolete
TSG-07 : Gen1 (1.5Gb/s) TJ at Connector, Clock to Data, fBAUD/10	
TSG-08: Gen1 (1.5Gb/s) DJ at Connector, Clock to Data, fBAUD/10	
TSG-09 : Gen1 (1.5Gb/s) TJ at Connector, Clock to Data, fBAUD/500	
TSG-10 : Gen1 (1.5Gb/s) DJ at Connector, Clock to Data, fBAUD/500	
TSG-11 : Gen2 (3Gb/s) TJ at Connector, Clock to Data, fBAUD/500	
TSG-12 : Gen2 (3Gb/s) DJ at Connector, Clock to Data, fBAUD/500	
TSG-13: Gen3 (6Gb/s) Transmit Jitter w/wo CIC	revised
TSG-14 : Gen3 (6Gb/s)TX Maximum Differential Voltage Amplitude	
TSG-15 : Gen3 (6Gb/s) TX Minimum Differential Voltage Amplitude	ECN-50 Revised
TSG-16 : Gen3 (6Gb/s) Tx AC Common Mode Voltage	revised

Phy Receive Signal Requirement	
RSG-01 : Gen1 (1.5Gb/s) Receiver Jitter Tolerance Test (Normative)	
RSG-02 : Gen2 (3Gb/s) Receiver Jitter Tolerance Test (Normative)	
RSG-03 : Gen3 (6Gb/s) Receiver Jitter Tolerance Test	ECN-50/ECN-51 Revised
RSG-05 : Gen1 Asynchronous Receiver Stress Test at +350ppm	
RSG-06 : Gen1 Asynchronous Receiver Stress Test With SSC	

SI General Requirements
SI-1:8 : Cable Characterization
SI-09 : Inter-Symbol Interference

Phy General Requirements
PHY-01 : Unit Interval
PHY-02 : Frequency Long Term Stability
PHY-03 : Spread-Spectrum Modulation Frequency
PHY-04 : Spread-Spectrum Modulation Deviation

Phy OOB Requirements
OOB-01 : OOB Signal Detection Threshold
OOB-02 : UI During OOB Signaling
OOB-03 : COMINIT/RESET and COMWAKE Transmit Burst Length
OOB-04 : COMINIT/RESET Transmit Gap Length
OOB-05 : COMWAKE Transmit Gap Length

Phy Receiver/Transmitter Channel Reqs	
RX/TX-01 : Pair Differential Impedance	
RX/TX-02 : Single-Ended Impedance (Obsolete)	
RX/TX-03 : Gen2 (3Gb/s) Differential Mode Return Loss	informative
RX/TX-04 : Gen2 (3Gb/s) Common Mode Return Loss	
RX/TX-05 : Gen2 (3Gb/s) Impedance Balance	
RX/TX-06 : Gen1 (1.5Gb/s) Differential Mode Return Loss	
RX/TX-07 : Gen3 (6Gb/s) Differential Mode Return Loss	
RX/TX-08 : Gen3 (6Gb/s) Impedance Balance	

SATA Measurement Legends:

- No change from previous UTD 1.3 spec version
- Revised methodology from UTD1.3 to UTD 1.4
- New test definitions in UTD 1.4
- Obsolete

Summary: TSG05/06 have been classified as EMI related and moved to an obsolete status. TSG15 will use an eye height methodology and will have different limits depending on the DUT being a Host or Device

SATA ECN 50: Asymmetric Amplitude and Measurement Methodology

- Host and Device Transmitter signal amplitude asymmetry has now been instituted in the SATA 3.1 specification.

Parameters	Units	Limit	Electrical Specification					Detail Cross-Ref Section	Measurement Cross-Ref Section
			Gen1i	Gen1m	Gen2i	Gen2m	Gen3i		
SSC_{tol}, Spread-Spectrum Modulation Rate	ppm/ usec	Max	1250		1250	1250-		7.2.2.1.6 7.3.3	7.4.15

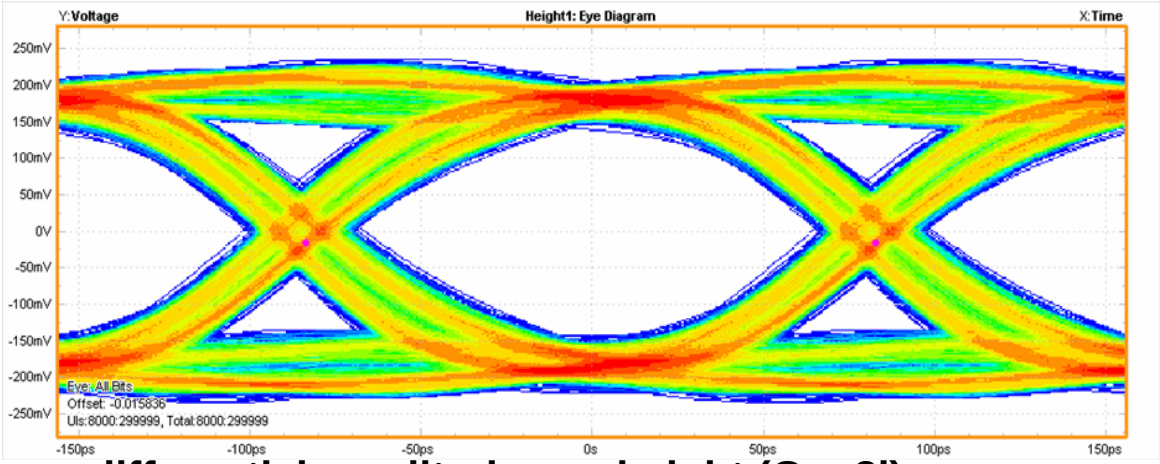
Parameter	Units	Limit	Electrical Specification					Detail Cross-Ref Section	Measurement Cross-Ref Section
			Gen1i	Gen1m	Gen2i	Gen2m	Gen3i		
$V_{diffTXdevice}$, TX Differential Device Output Voltage	mVppd	Min	400		400		-	7.2.2.3	7.4.6
		Min	-		-		240		7.4.4 7.4.4.3
		Nom	500		-		-		7.4.6
		Max	600		700		-		
		Max	-		-		900		7.4.4
$V_{diffTXhost}$, TX Differential Host Output Voltage	mVppd	Min	400		400		-	7.2.2.3	7.4.6
			-		-		200240		7.4.4 7.4.4.3
		Nom	500		-		-		7.4.6
		Max	600		700		-		
		Max	-		-		900		7.4.4
UI_{VmmTX}, TX Minimum Voltage Measurement Interval	UI		0.45-0.55		0.45-0.55		-	7.2.2.3.2	7.4.6
			-		-		0.50 0.45-0.55		7.4.6.2

- Hosts may signal as low as 200mV (40mV lower than previously allowed)
- Devices must retain the original 240mV levels.
- Measurement methodology has been revised from a vertical BER contour to a simple Eye Height measurement.

Tektronix

SATA ECN 50: Asymmetric Amplitude and Measurement Methodology (cont)

- Due to the changes to the minimum Tx levels receiver test protocols have now been revised to calibrate for Host and Device amplitudes differently.



7.4.3.3 Minimum differential amplitude eye height (Gen3i)

4 Sigma Eye Diagram constructed from a minimum 5E6 number of unit intervals with either mean based PLL or an explicit clock or mean based clock equivalent.

Receiver Test Calibration levels and method details.

200mV for Drive side receiver (Eye Height + CIC + Explicit Clock)

240mV for the Host side receiver (Eye Height + CIC + Explicit Clock)

Transmitter Test levels and method details.

240mV for Drive side transmitter minimum (Eye Height + CIC + JTF PLL)

200mV for the Host side transmitter minimum (Eye Height + CIC + JTF PLL)

The Eye Height is to be evaluated at the 50% UI location.

SATA ECN 51: Change of receiver test pattern to FCOMP

The receiver tolerance test shall be conducted over variations in parameters SSC on and off, minimum and maximum amplitude, common mode interference over the specified frequency range, the test pattern **FCOMP** described in section 7.2.4.3.7, and jitter which includes the maximum random and deterministic jitter of various types: data dependent, periodic, duty cycle distortion.

7.2.4.3.7 Framed Composite Pattern (FCOMP)

The Framed Composite Pattern is equivalent to the COMP pattern in section 7.2.4.3.6 with the following structured changes:

1. In-line with section 7.2.4.2 the COMP pattern is framed. 2 ALIGN primitives inserted every 256 Dwords.
2. A short Inter Gap region is introduced before and after the SOF/EOF to ensure that when repeated sequentially by a generator the 256 Dword ALIGN primitives are perfectly and uniformly spaced 256 Dwords apart even after wrap-around by the generator.

1.1.1.1.1 Framed Composite Pattern (FCOMP)

Table FF – Framed Composite Pattern (FCOMP)

	Transmission Order →										
	K28.5(BCh)+		D10.2(4Ah)-			D10.2(4Ah)-			D27.3(7Bh)-		
+	1100	0001	0101	0101	0101	0101	0101	0111	0110	0011	+
	C	1	5	5	5	5	5	7	6	3	
	Above Dword is repeated a total of 2 times. 2 DW ALIGNp.										

Tektronix

SATA UTD 1.4.2 Pending Revisions
TSG-13: Gen3 (6Gb/s) Transmit Jitter

2.15.13. TSG-13: Gen3 (6Gb/s) Transmit Jitter

- **Device/Host Expected Behavior**

See sections 7.2.2.2.18 and 7.3 of Serial ATA Revision 3.0.

Measurement Requirements

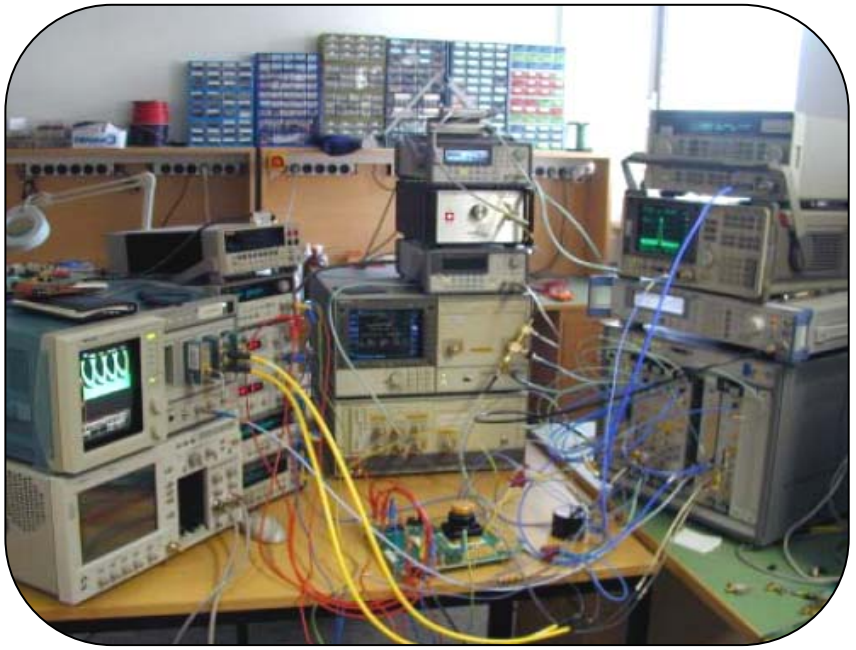
- See sections 7.4.8 and 7.4.10 of Serial ATA Revision 3.0.
- This test requirement is only applicable to products running at 6Gb/s.
- The Jitter Transfer Function (JTF) for the Jitter Measurement Device (JMD) is required to be per section 7.3.2.4 of Serial ATA Revision 3.0.
- The Compliance Interconnect Channel (CIC) for the second Random Jitter (RJ) and the Total Jitter (TJ) measurements is required to be per section 7.2.7 of Serial ATA Revision 3.0.
- The methods used for Gen3i Transmit Jitter testing are intended to minimize RJ measurement error and allow for the TJ to be verified by a full population BERT scan as described in Section 7.4.8. This method also puts an upper limit on both DJ and RJ so neither may dominate the TJ.
- The Transmit Jitter values specified in Table 31 of Serial ATA Revision 3.0, refer to the output signal from the unit under test (UUT) at the mated connector into a Laboratory Load (LL) when measuring Random Jitter (RJ) or from the unit under test through a Compliance Interconnect Channel (CIC) into a Laboratory Load when measuring Total Jitter (TJ). All the interconnect characteristics of the transmitter, package, printed circuit board traces, and mated connector pair are included in the measured transmitter jitter. Since the SATA adapter is also included as part of the measurement, good matching and low loss in the adapter are desirable to minimize its contributions to the measured transmitter jitter.
- The Random Jitter is measured with a MFTP pattern and the Total Jitter is measured with each of the specified patterns in section 7.2.4.1 (LBP, HFTP, MFTP and LFTP) and section 7.2.4.3.4 (SSOP).
- The minimum rise time allowed by Gen3, (33psec [+- 5%] Gen3 ~~and [assuming the TP033 is approved]~~ 50psec to Gen1 and Gen2 products) should be incident to the Rx channel when the device is in loopback (BIST-L) as measured at the reference plane (to the SMA's of the test fixture).
- Amplitude of the incident signal will be the Tx maximum (900mV ppd)

Industry Productivity

Compliance Testing – An ***Industry*** Productivity Issue

Greater speed means greater design challenges, with implications…

1. Greater test complexity
   - More instruments, configurations, and setup time
2. Breadth of tests demands highly experienced, ***senior equipment users*** to perform and interpret results.
3. Text complexity is high
   - Highly specialized – e.g., SSC modulation analysis, advanced receiver testing, Frequency domain S-Parameter measurements.
4. Days to perform effective, repeatable and reliable product validation tests

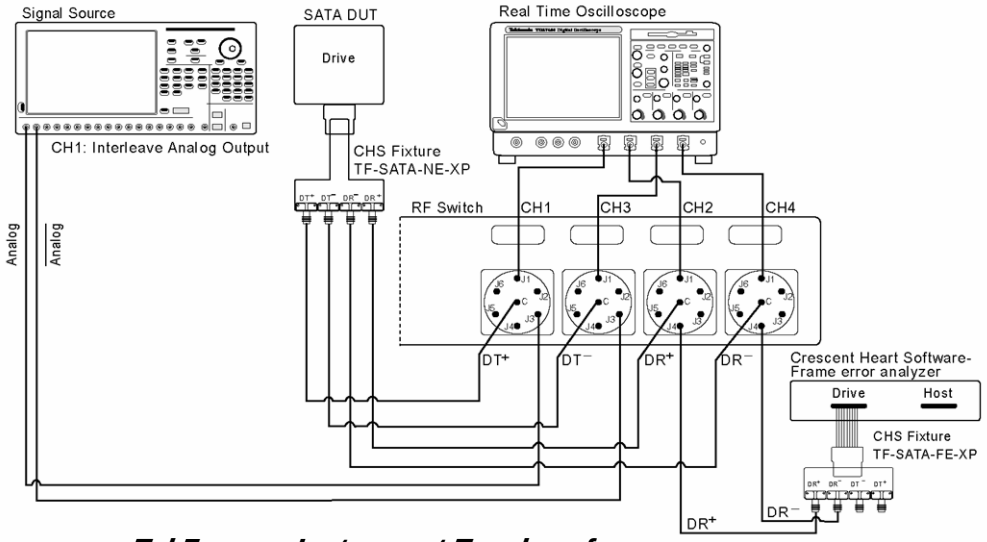


"Banner specs are no longer the gating issue. The latest equipment provides ample raw performance. What's needed is greater ease of use, setup and automation."

- Customer feedback

Integrated and Automated Test Control

TekExpress™ Test Automation Framework

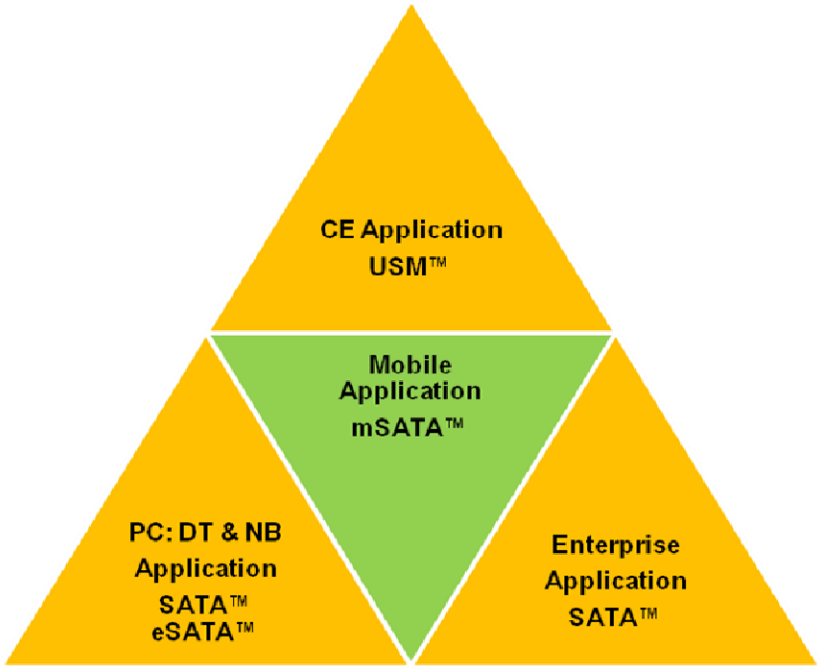


TekExpress Instrument Topology for TSG/PHY/OOB, RSG testing

- Simplifies Complex Measurements
- Improves Engineering Productivity
- Repeatable and Consistent results
- Automatic with no user intervention

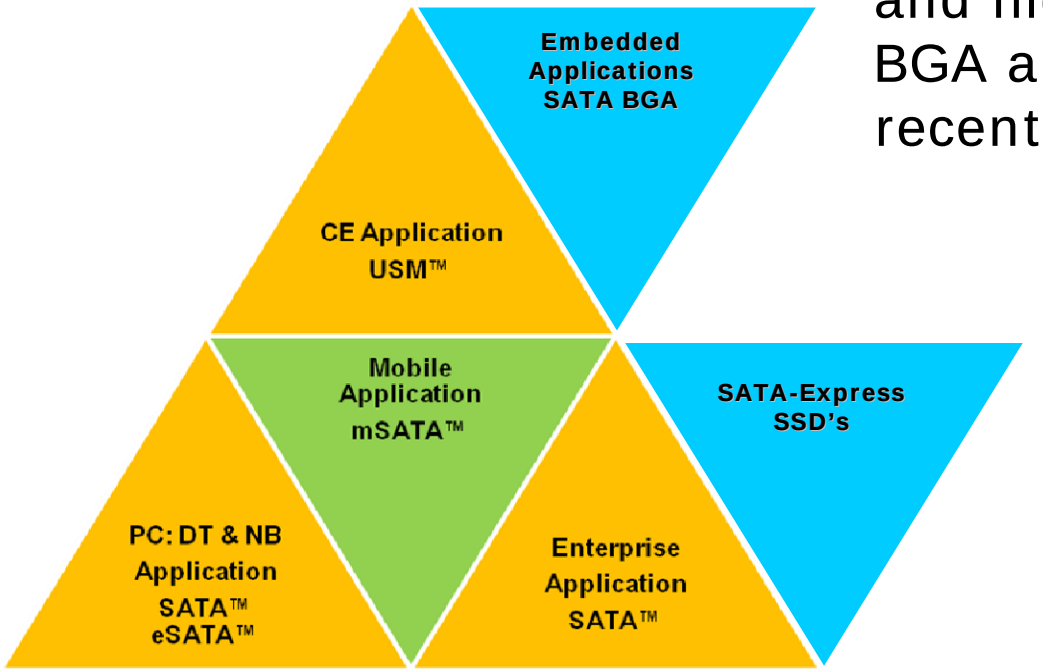
- Complete offering includes:
  - Leading portfolio of Tektronix test instruments
    - Oscilloscopes
    - Signal generators
  - 3rd party integration with RF switch (Keithley), fixtures, API (NI), cabling, deskew, etc.
- Auto discovery of instruments using GPIB, USB, and LAN
  - 1GbE networking is used for data/waveform transport.
  - GPIB/488.2 is used for RF Switch and Power supply communications
- Test sequence automation of **all physical layer aspects of SATA**. The only solution in the industry today which can claim this.
- RF-Switch automation of Keithely microwave switch matrix.
- One button control
- Reporting

The SATA Ecosystem: Last Year



SATA covers 95% of the storage ecosystems spanning HDDs, ODDs, SSD, and hybrid HDDs in client PC, mobile, enterprise, CE, and embedded storage markets

The SATA Ecosystem: Now



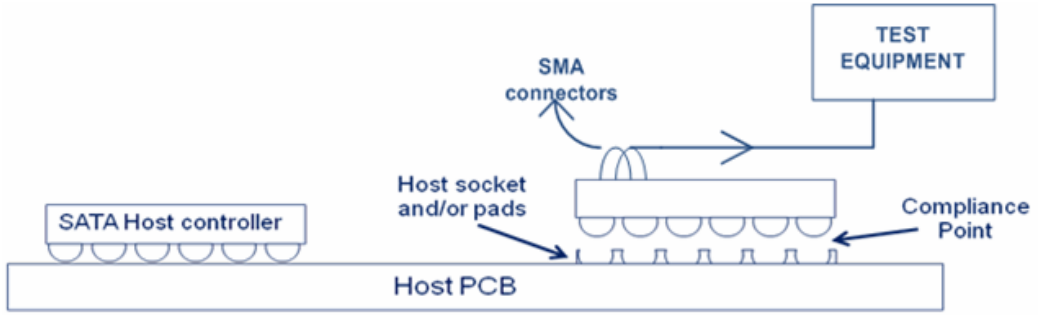
Today, SATA is expanding in specialized low power, compact and high performance areas with BGA and SATA-Express Solutions recently approved by SATA-IO.

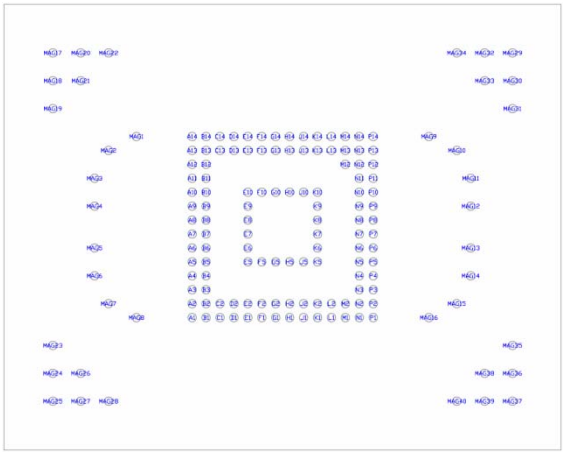
SATA uBGA SSD's present unique test chalenges

New SSD's approved by SATA-IO

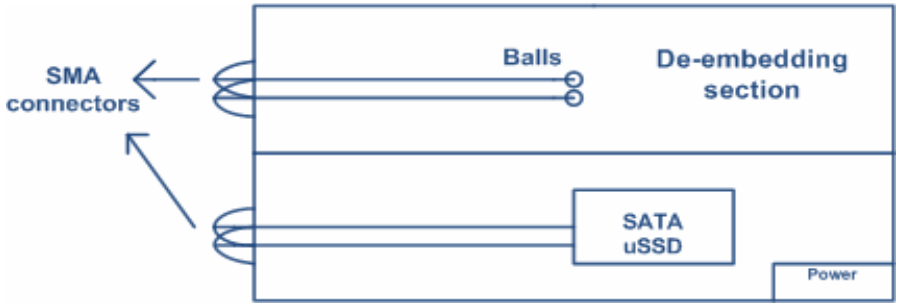
- Test Strategy: Two BGA based fixtures are being commissioned with Wilder through SATA-IO. Target available through reference sell by Q4

  - Device Fixture: A carrier with a uBGA Socket Connector.





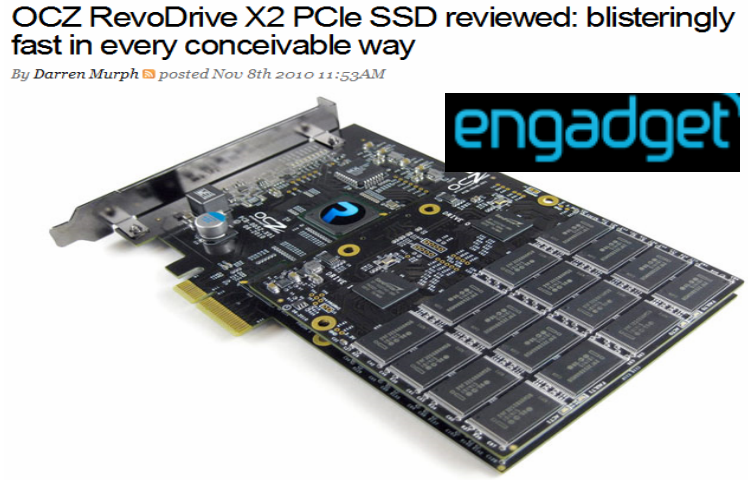
  - Host Fixture: A BGA breakout to SMA's which will be re-flowed onto the motherboard.

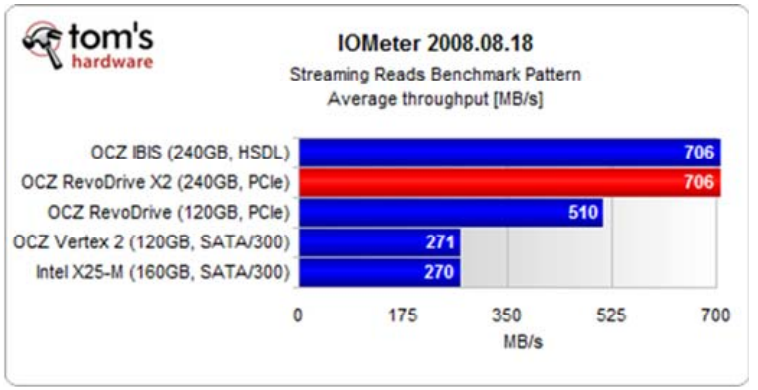


Tektronix

The Need For More Speed: SSD

- SSD speed is dependent on the number of NAND channels
  - ONFI 3.0: 400 MB/s
- SSDs can saturate the 12G SATA
- Development of 12G SATA would be costly, take 2-3 years
  - New connectors/cables, not backward compatible
- Cost not in line with SATA-IO mission





No 12G SATA development planned **at this time**

PCIe

- PCIe is multi-lane → scalable
  - 1 to 16 lanes
- PCIe (Gen3) each lane is 8 Gbps → 1 GB/sec
  - no 8b/10b encoding
- It is already available
- PCIe is **NOT** a replacement for SATA

PCIe is the obvious choice

SATA Express: Tomorrow's Vision

SATA Device: Remains the same

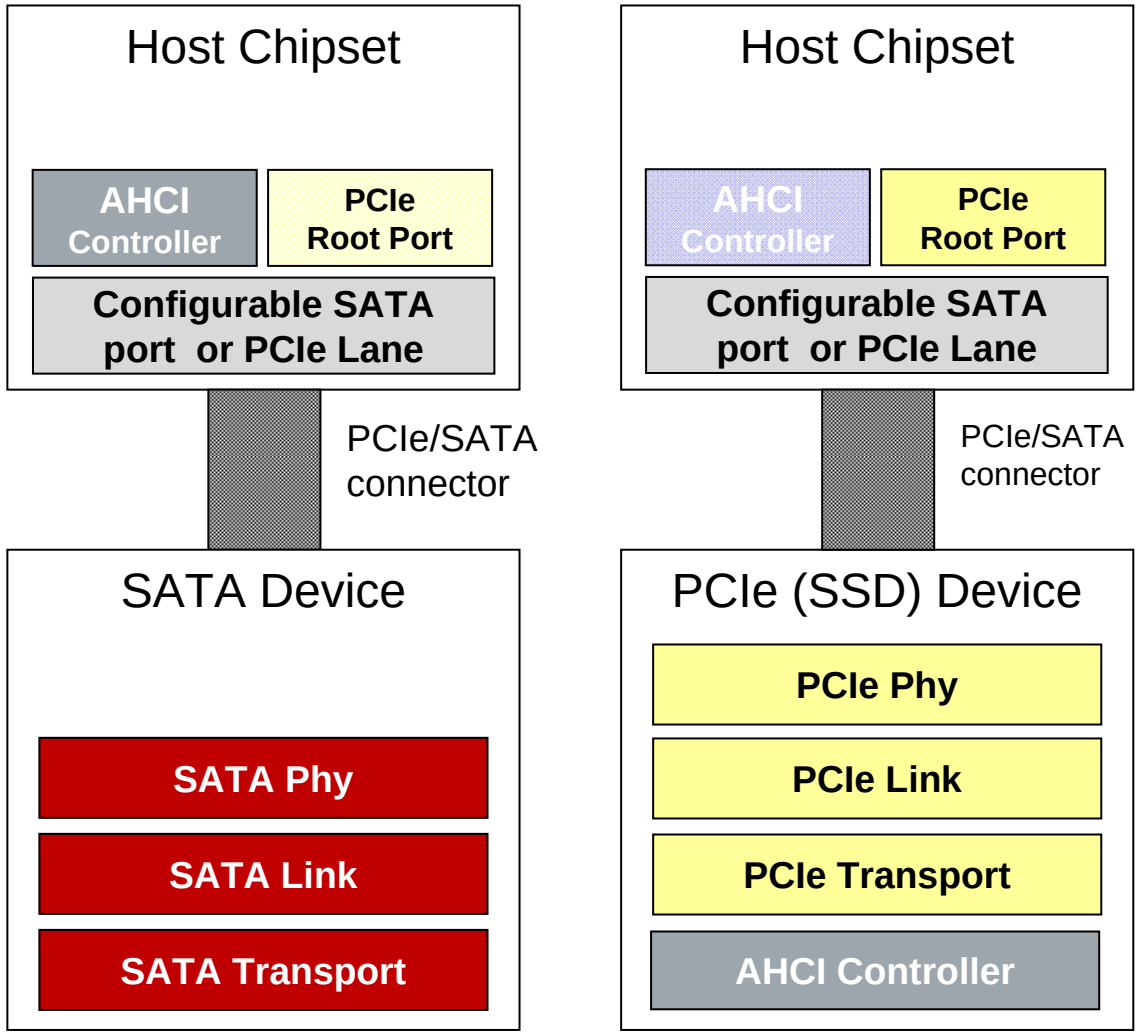
PCIe (SSD) Device: Exposed as AHCI controller for software compatibility

- Will eventually transition to optimized interface (*e.g.* NVMExpress)

Firmware: A SATA (SSD) and PCIe (SSD) exposed as AHCI are equivalent to FW

- ATA command set
- Leverage existing investment in SW

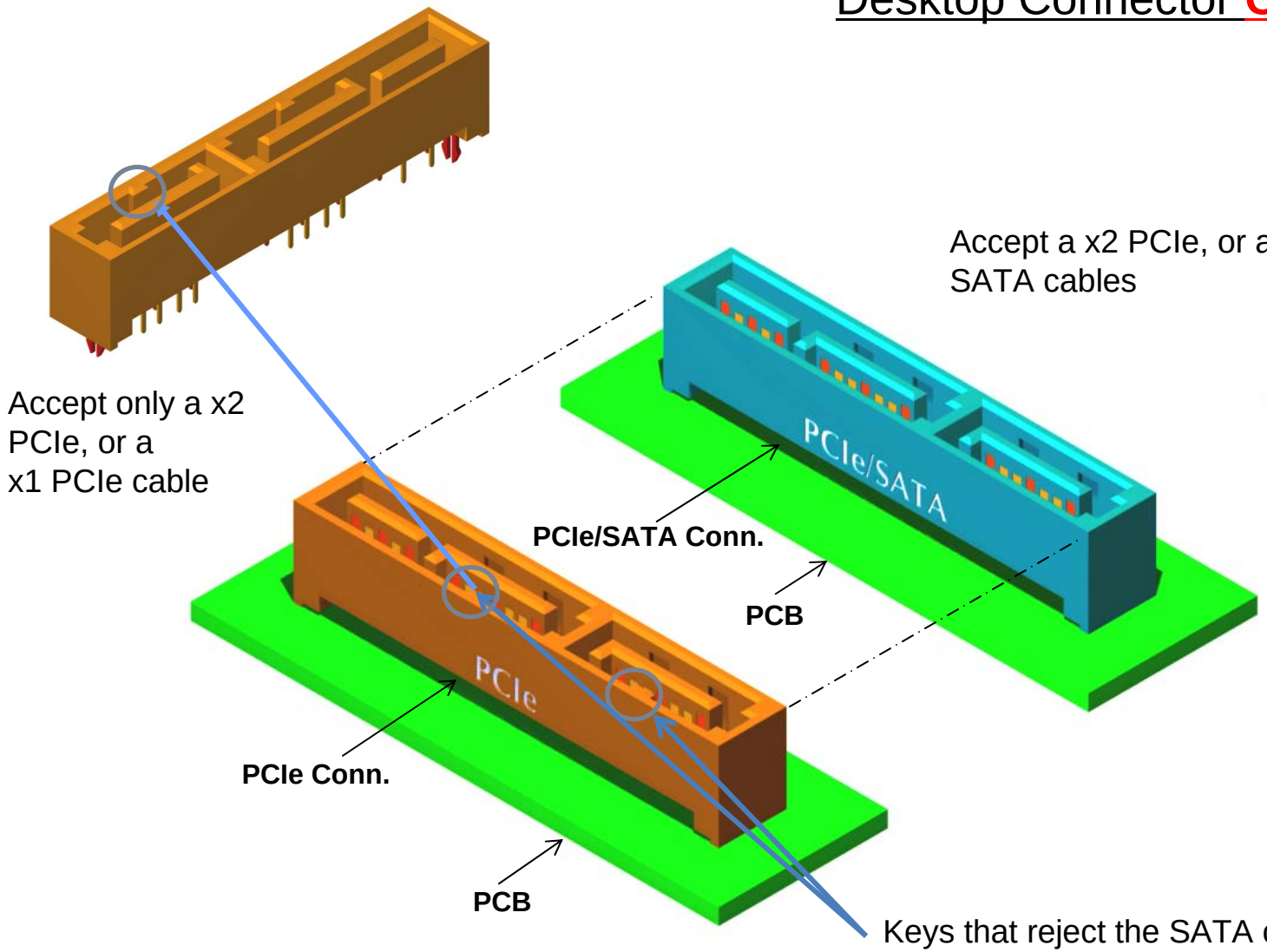
Architecture Tomorrow



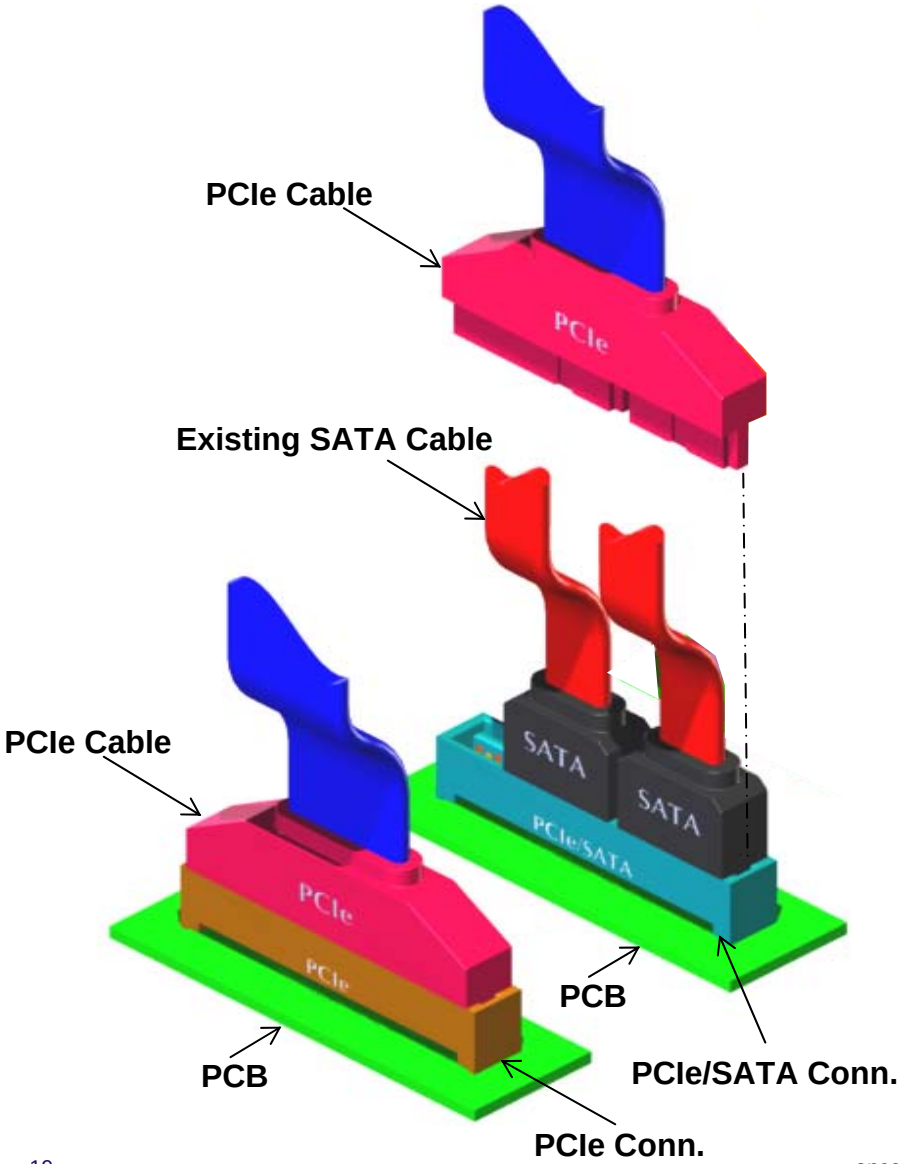
Tektronix

Enabling the New SATA Express Ecosystem

Desktop Connector **Concept**



Enabling the New SATA Express Ecosystem



Desktop Cables **Concept**

SATA devices will coexist with next generation PCIe devices

SATA cost/performance benefits

Requires a connector that supports both PCIe and SATA

- Allows a single motherboard (backplane) connector to support both interfaces

HDD-compatible form factors to be defined for PCIe devices

- Enables system-level mechanical compatibility
- Preserves high-capacity storage

SATA-IO CabCon has been chartered to develop SATA compatible connectors and form factors for PCIe SSD/hybrid drives

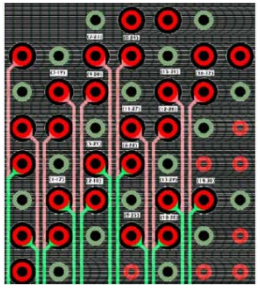
Tektronix

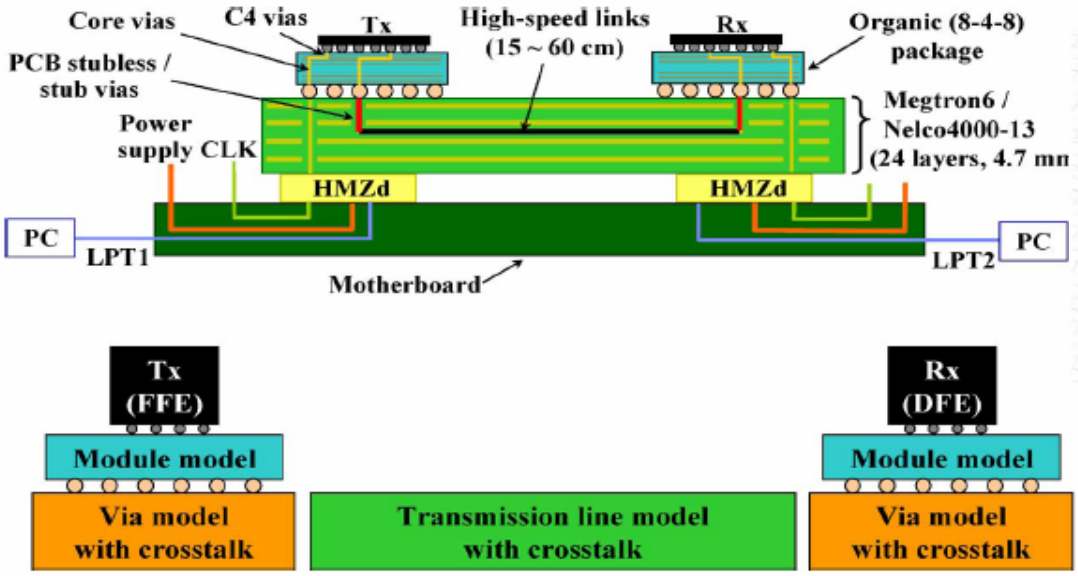
Fundamentals of 12G SAS characterization

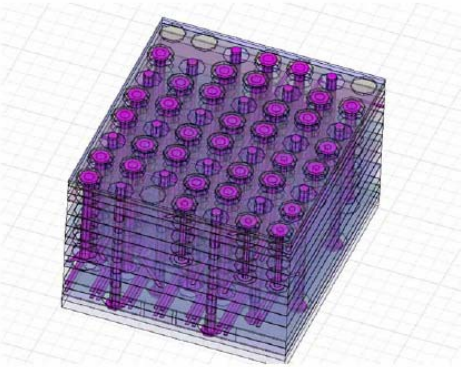
12+G Design Problem:
1000mV, FFE, Crosstalk, Crosstalk, Crosstalk, DFE, 50mV

- Significant advances in high tap count Decision Feedback Equalization are key to operating at 12+G.
- Mitigating the complex Channel Crosstalk and Signal loss problems which 12+G designs present, is the largest design challenge today.
- Typical Escape Structure bandwidth is 18GHz.
- Crosstalk is often beyond the capability of current equalization architectures to combat, and needs to be quantified if accurate performance projections are to be made based on experimental measurements. For short channels, NEXT may be less of an issue since the insertion loss is not as severe; however, in longer links and at higher data rates it has the potential to become a dominant design consideration. Ref:[1]

KAM *et al.*: IS 25 Gb/s ON-BOARD SIGNALING VIABLE? 329



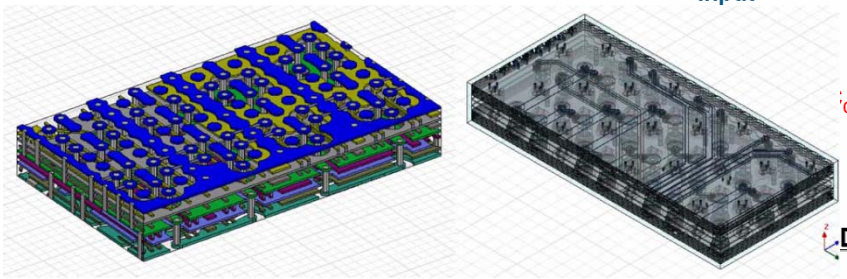


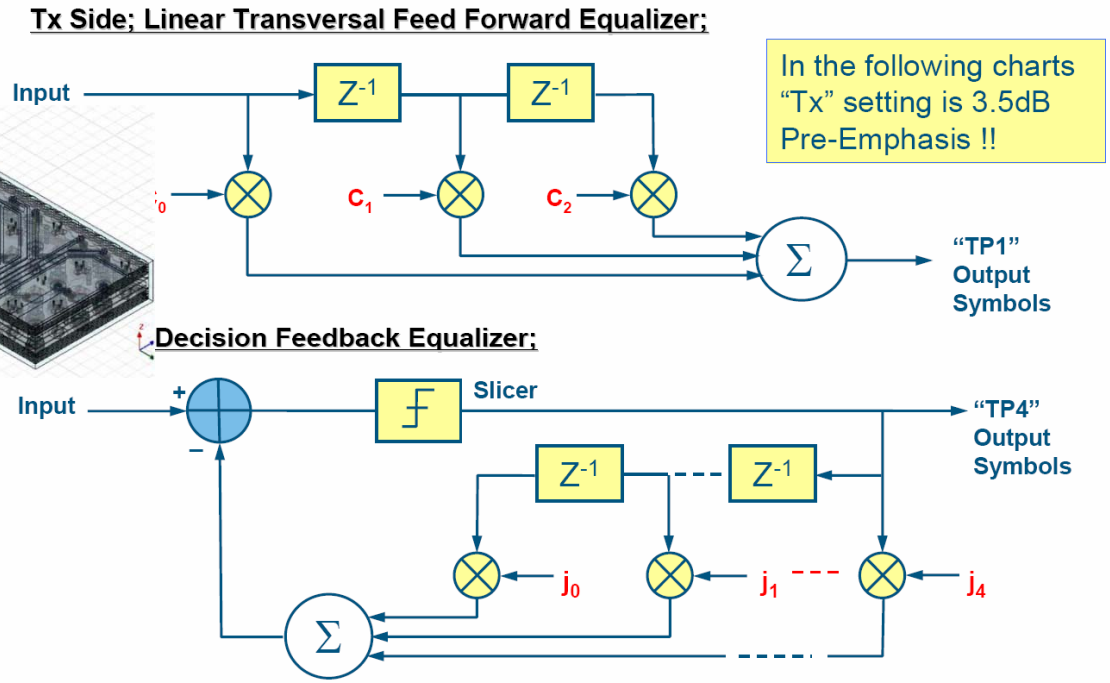


12G Design Problem: FFE, DFE

- A non-recursive DFE can only compensate a fixed time span of ISI. In very low-bandwidth channels, significant post-cursor ISI may fall outside the time span covered by the DFE taps.
- FFE can compensate ISI over a very wide time span since the FFE filter response is convolved with the impulse response of the channel.
- The utility of FFE alone drops off rapidly over complex channels which have spectral nulls (Via stubs, connectors, etc) which require many FFE taps to cancel reflections.
- Optimal solutions exist around 4-tap FFE with 20+ (20*60mW) tap DFE designs. More emphasis is required in the Receiver section of the topology as more aggressive FFE makes crosstalk worse.

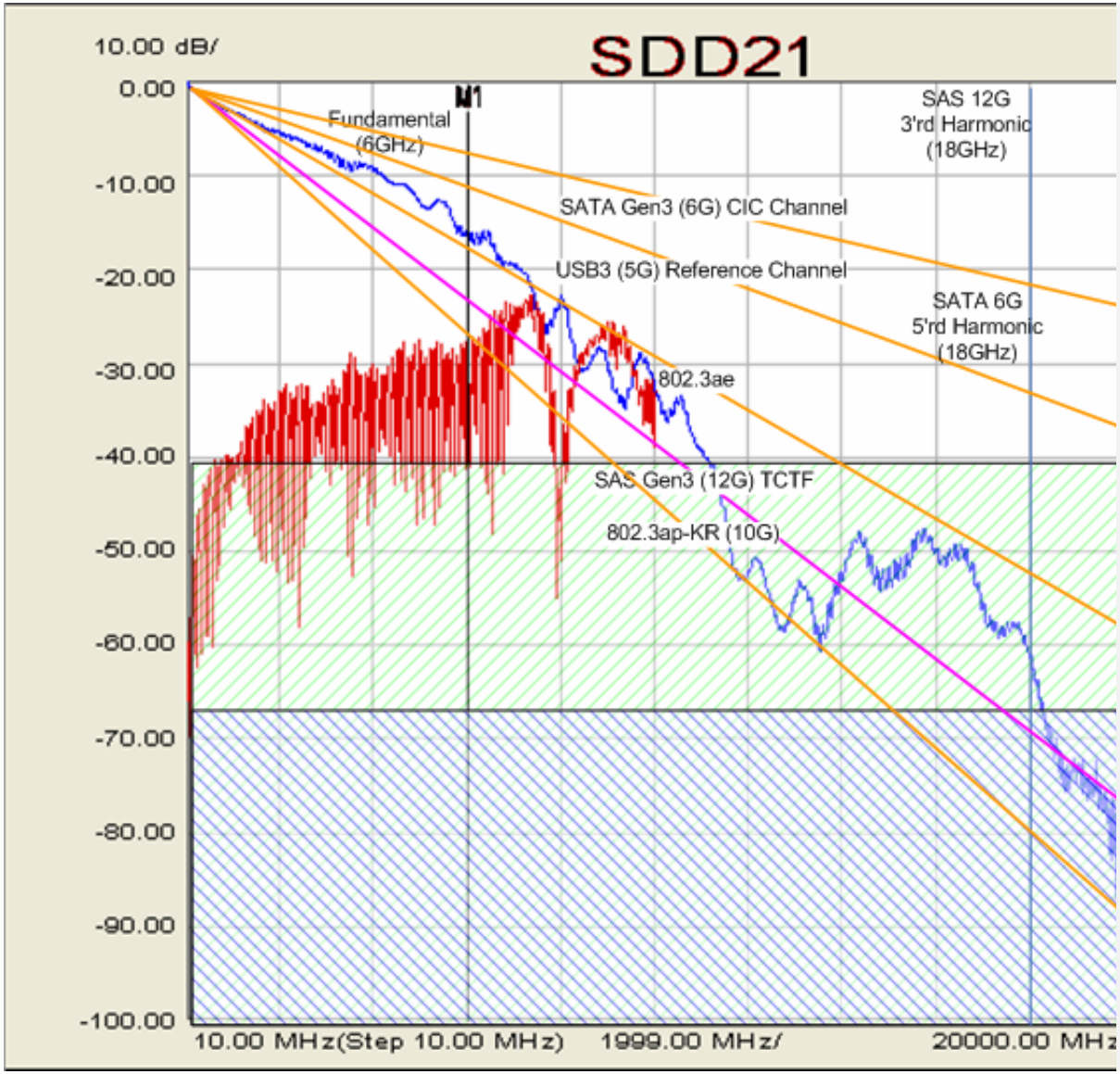
Ref:[1], [2]





Ref:[2]: Tyco Electronics

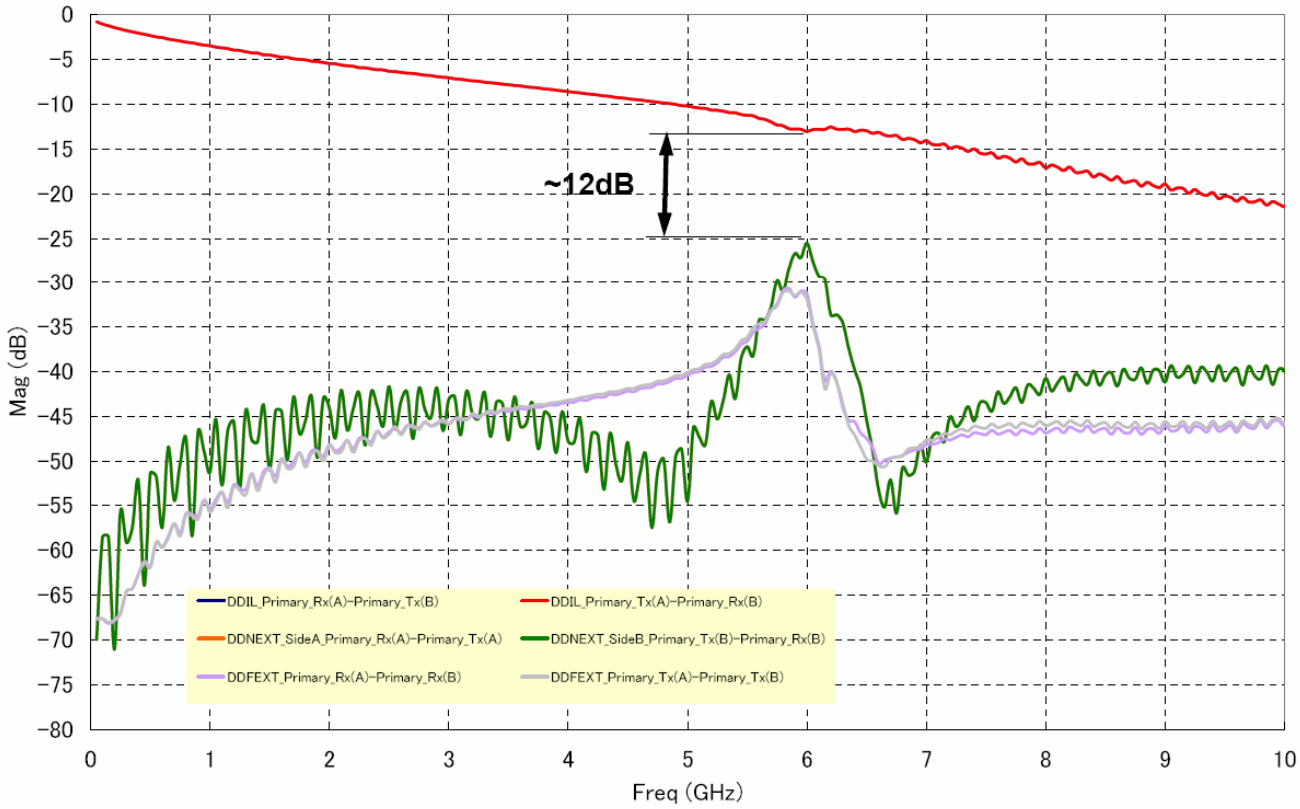
DFE/FFE Advances
-vs- Dynamic Range Range -vs Data Rates –vs- BW



- 10+G datarates are being achieved by both advanced DFE, and FFE systems with link training and adaptive channel compensation.
- Channel bandwidth is largely being held constant from similar 6G solutions, with focus on ***extracting the un-tapped carrier capacity*** in the low dynamic range regions of the channel.
- 6-10 Meter solutions and associated reference channels are shown here.

High Speed FR4 FR4 materials

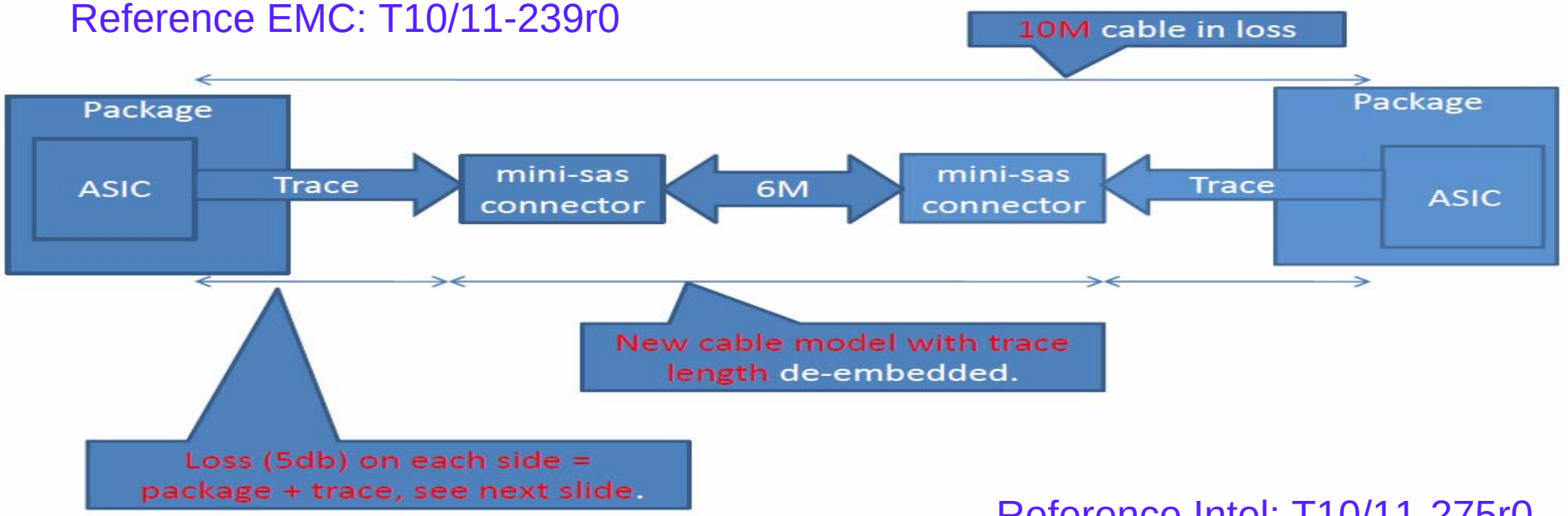
Channel Model 'ICR', Near Top Layer w/ Backdrilling, Primary Port (Device: FR4 Std. / Backplane: 22" FR4 High Speed)



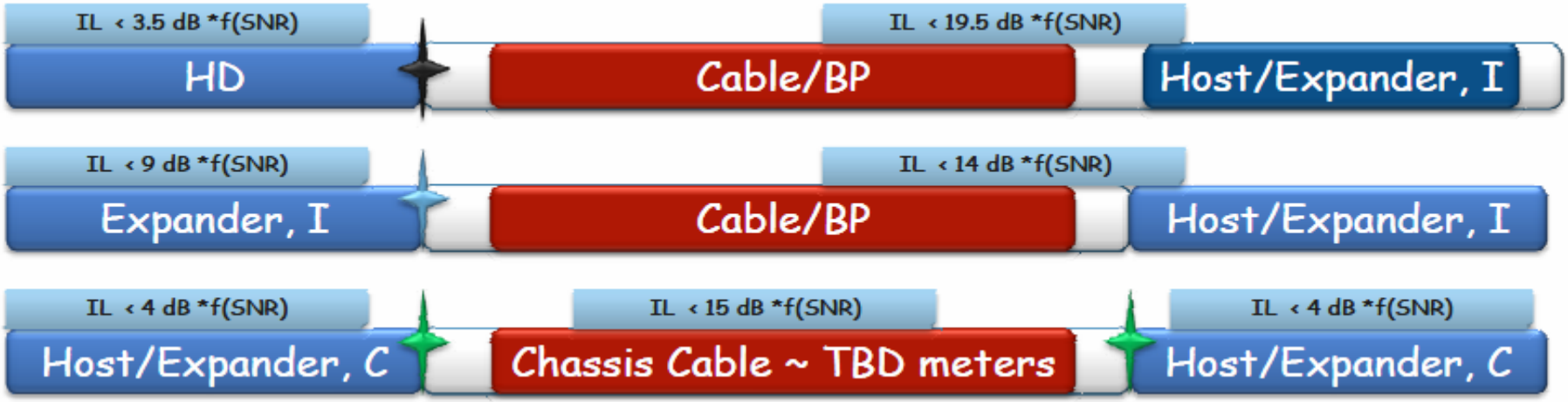
Ref:[2]: Tyco Electronics

Channel allocation model (EMC and Intel proposal) Fixture De-embed and Reference channel embed required

Reference EMC: T10/11-239r0



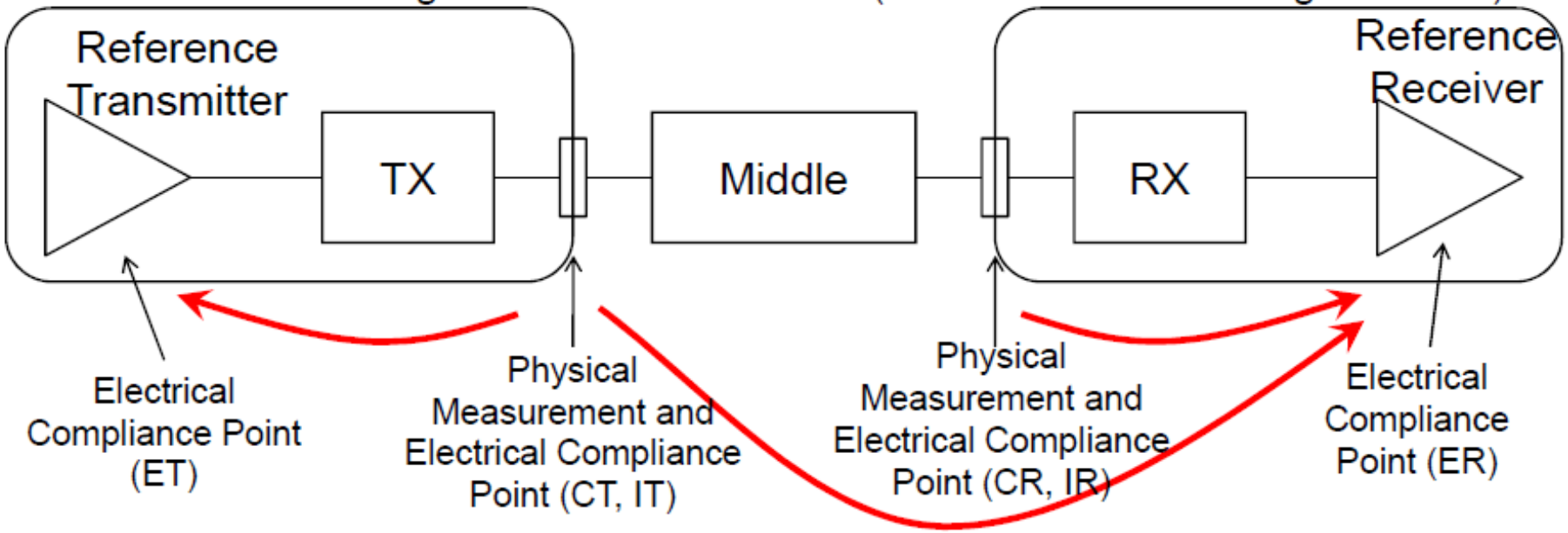
Reference Intel: T10/11-275r0



SAS3_EYEOPENING Usage Review
Compliance Points

PMC-SIERRA
Enabling connectivity. Empowering people.

- SAS3_EYEOPENING is used to transform the measurement from a Physical Compliance point into an Electrical Compliance Point
  - CT/IT into ET for TX Equalization & Amplitude
  - CT/IT into ER for TX ISI and crosstalk
  - CR/IR into ER for crosstalk and RX stress pattern calibration and for Middle section ISI compliance
  - De-embedding of test fixture is allowed (other "Probe Point" usage in SAS2)

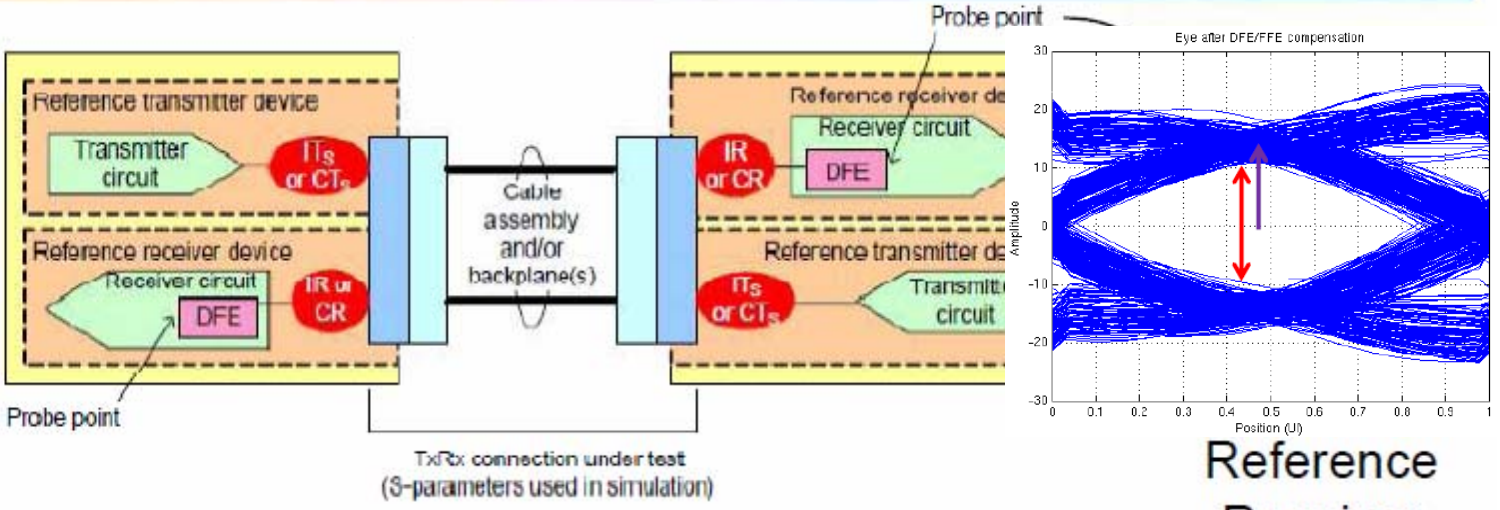


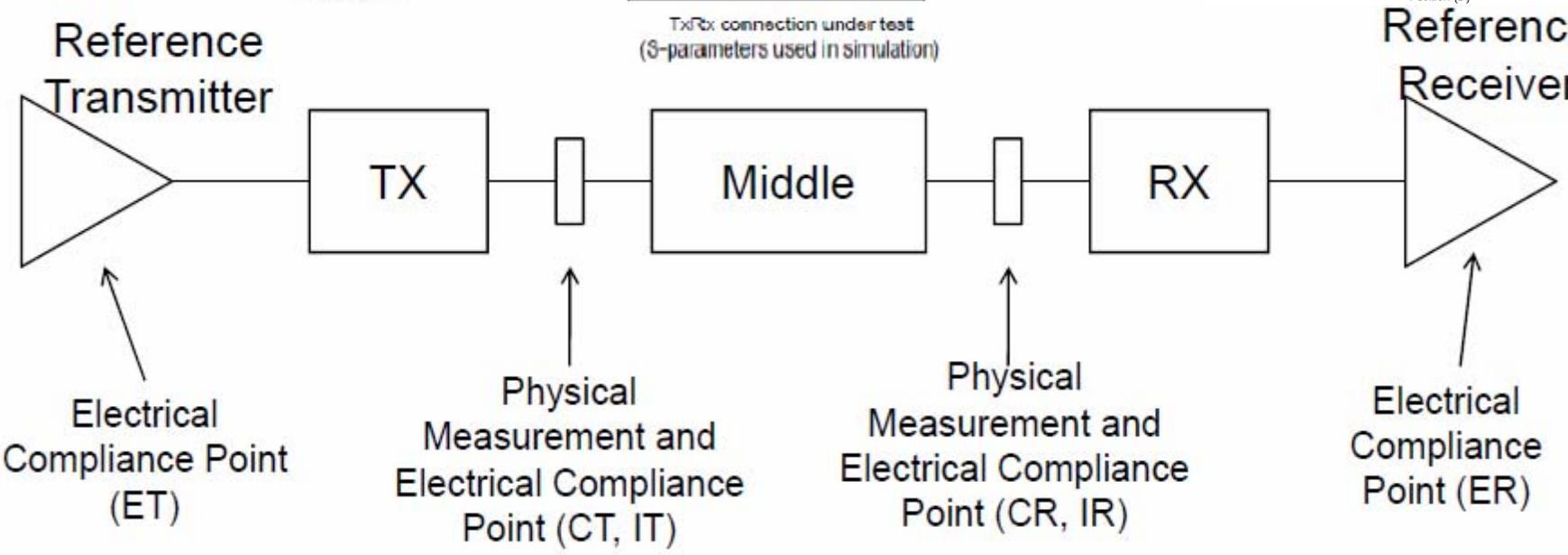
Tektronix

SAS3_EYEOPENING Usage Review
Compliance Points

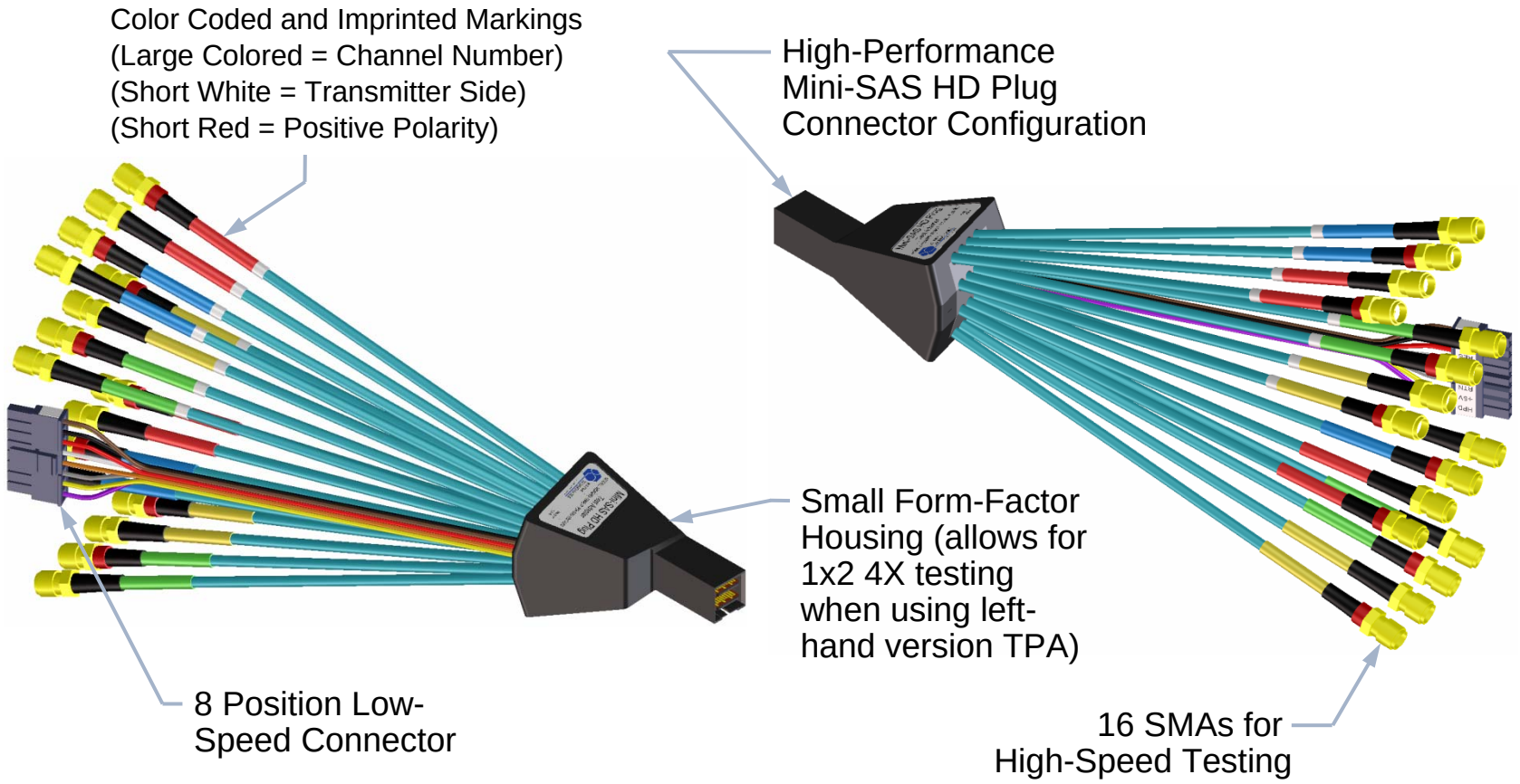
PMC-SIERRA
Enabling connectivity. Empowering people.

Note from Harvey: In SAS2, "Probe point" did not include any channel model, only the DFE model.





Mini-SAS HD Plug Test Adapter (right-hand) Top Views

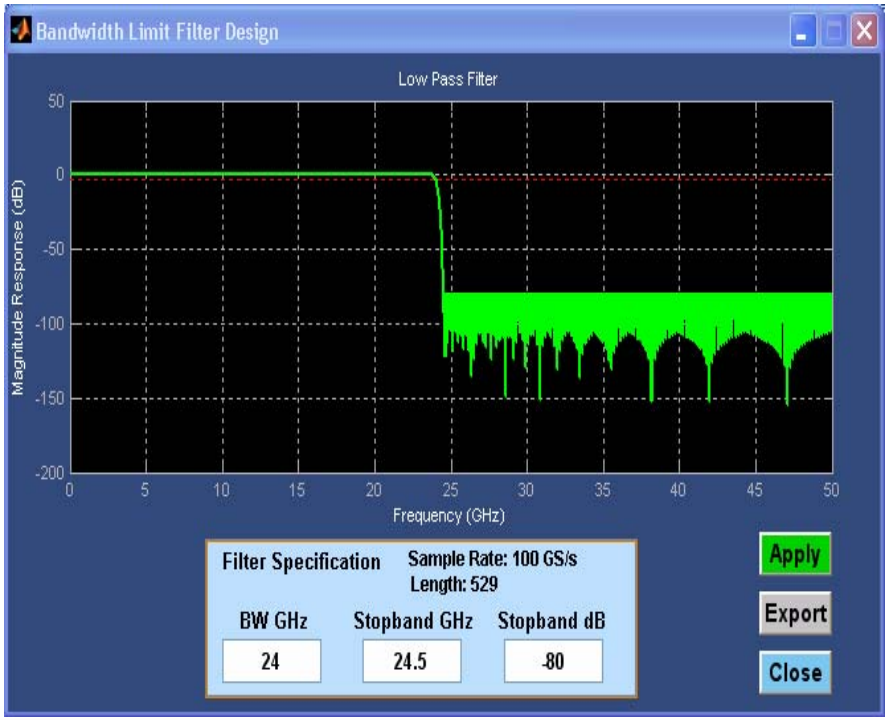


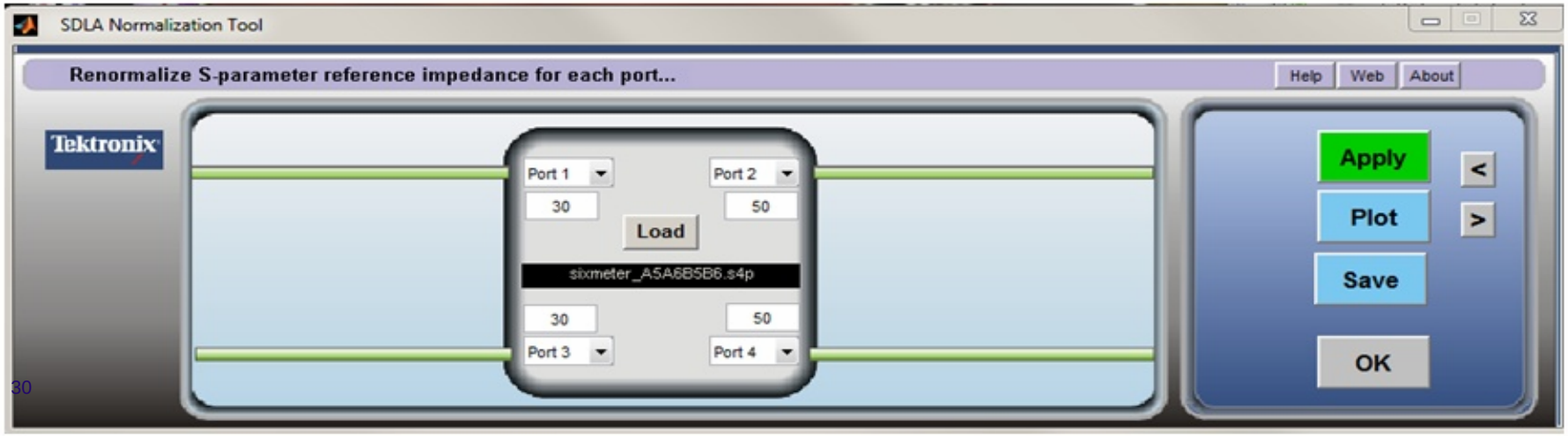
Preliminary de-embed strategy (Subject to change)

- SASHD connectors when peripheral HW and cables are de-embedded have decent performance out past 26G. It will be important to back out the fixture loss contributions from physical layer measurements.
  - SDLA on a Tektronix **25G Scope** Will the principal means of achieving this.
  - Empirical analysis of connectors suggest net fixture/connector losses at +14dB at 25G. De-embedding to the natural null which occurs at the 4'th (24GHz) harmonic results in the best signal quality with the least de-embed filter induced ringing. The added 1GHz allows margin in the instrument stop band for a gradual roll off.
- PMC's SAS3 utility computes Silicon to Physical Layer compliance point de-embed by analyzing differences in specific Tx pattern properties. This tool may automatically take this process into account but it's still un-proven.

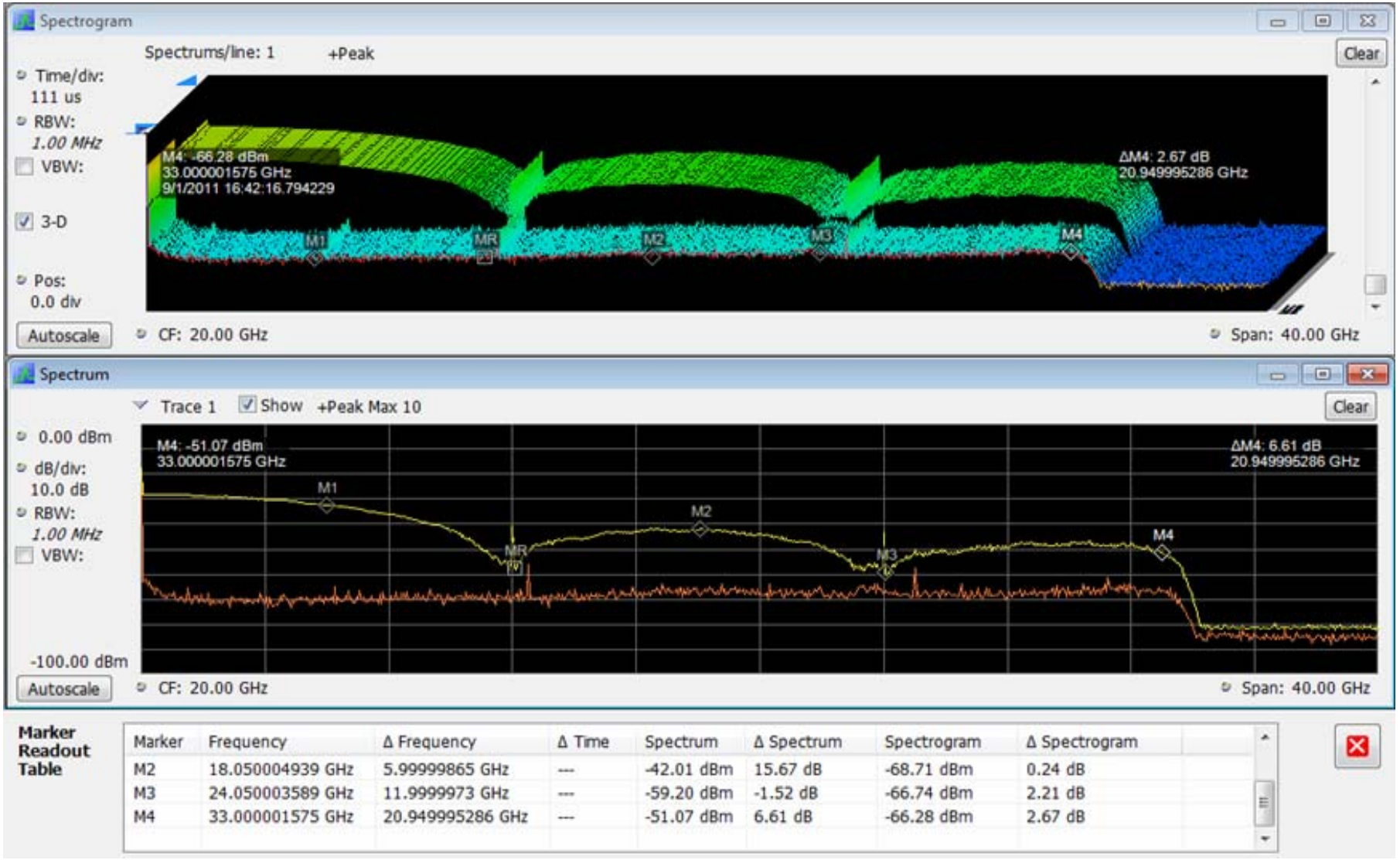
SDLA Filters (de-embed cont)

- S4P filter/de-embed configurations can be loaded inline with the acquisition channel and provide a virtual channel to the SAS compliance point (At the silicon).
- User defined fast transition equi-ripple filter stop-band performance from -20 to -200dB
- In auto mode the -80dB limit is always used





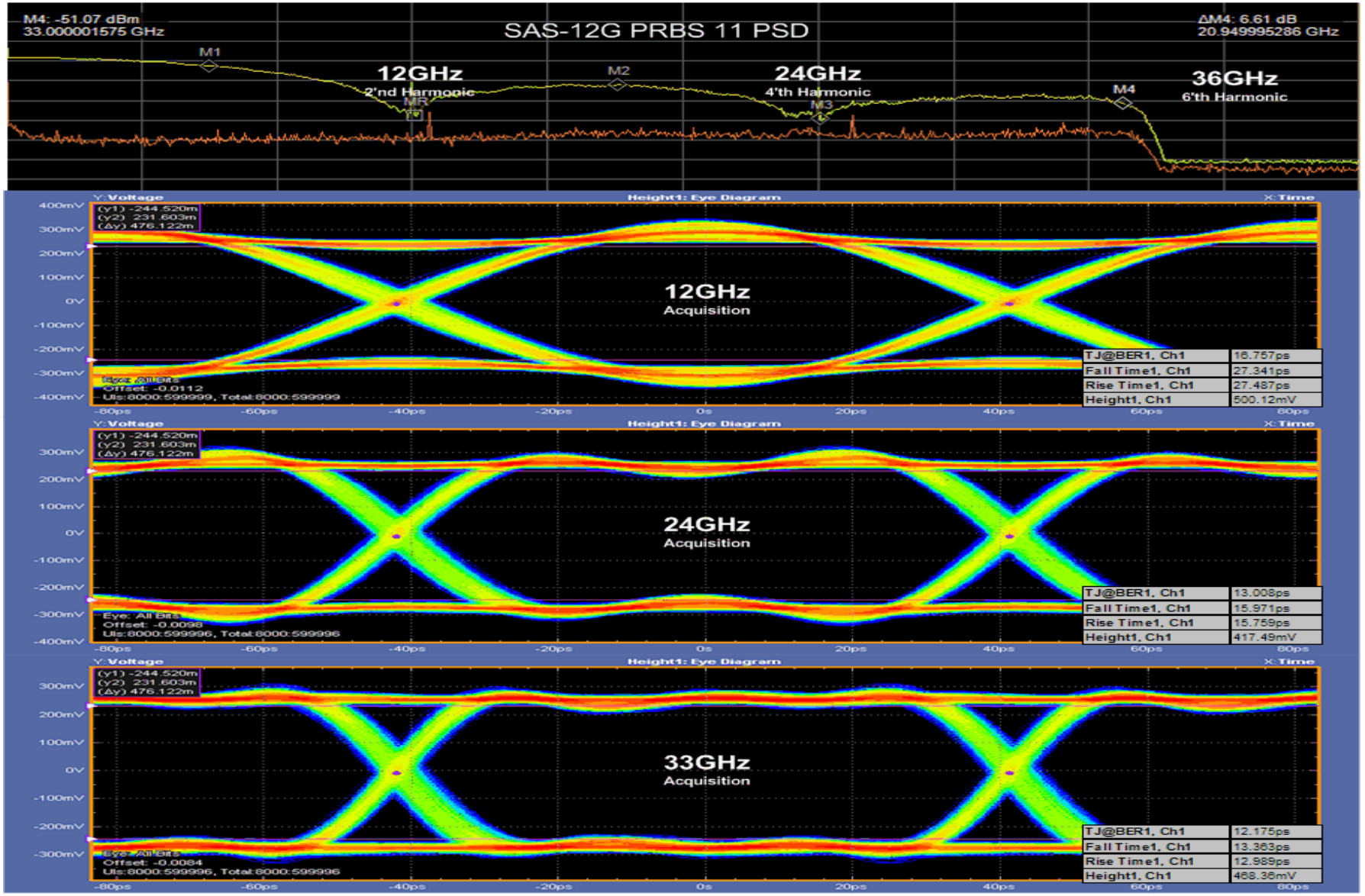
Spectrum
SAS PRBS11 12G NRZ Power Spectrum



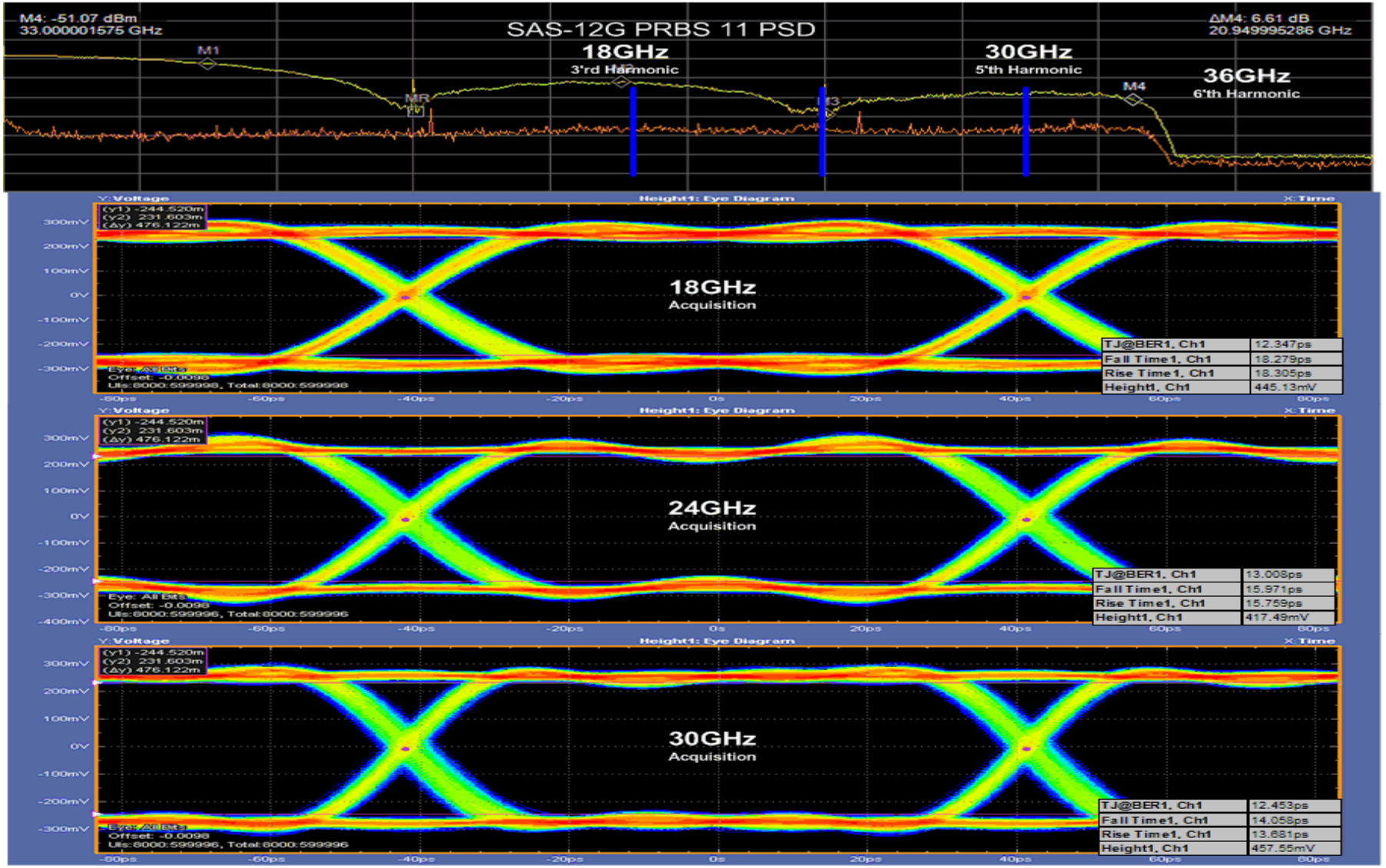
Marker	Frequency	Δ Frequency	Δ Time	Spectrum	Δ Spectrum	Spectrogram	Δ Spectrogram
M2	18.050004939 GHz	5.99999865 GHz	---	-42.01 dBm	15.67 dB	-68.71 dBm	0.24 dB
M3	24.050003589 GHz	11.9999973 GHz	---	-59.20 dBm	-1.52 dB	-66.74 dBm	2.21 dB
M4	33.000001575 GHz	20.949995286 GHz	---	-51.07 dBm	6.61 dB	-66.28 dBm	2.67 dB

Bandwidth

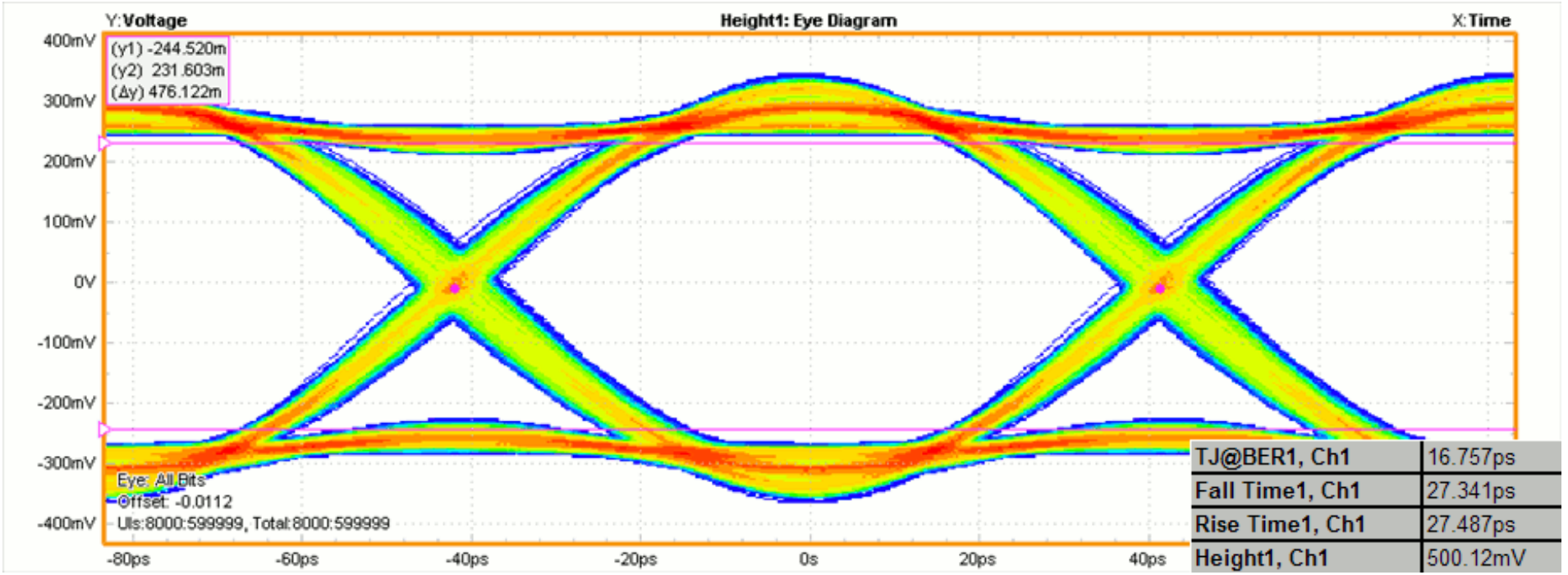
Signal Acquisition and bit rate harmonics



Bandwidth
Signal Acquisition and bit rate harmonics



Transmitter Characterization:
Precise Characterization of Silicon



- 33GHz multi-channel acquisition offers 100GS/sec resolution to resolve edges down to 9psec.
- 250M/channel acquisitions.