End-to-End System-Level Simulations with Repeaters for PCIe Gen4: A How-To Guide

Casey Morrison, Texas Instruments
Cindy Cui, Keysight

Yongyao Li (Huawei), Casey Morrison (Texas Instruments), Fangyi Rao (Keysight), Cindy Cui (Keysight), Geoff Zhang (Xilinx)

SPEAKERS

Casey Morrison

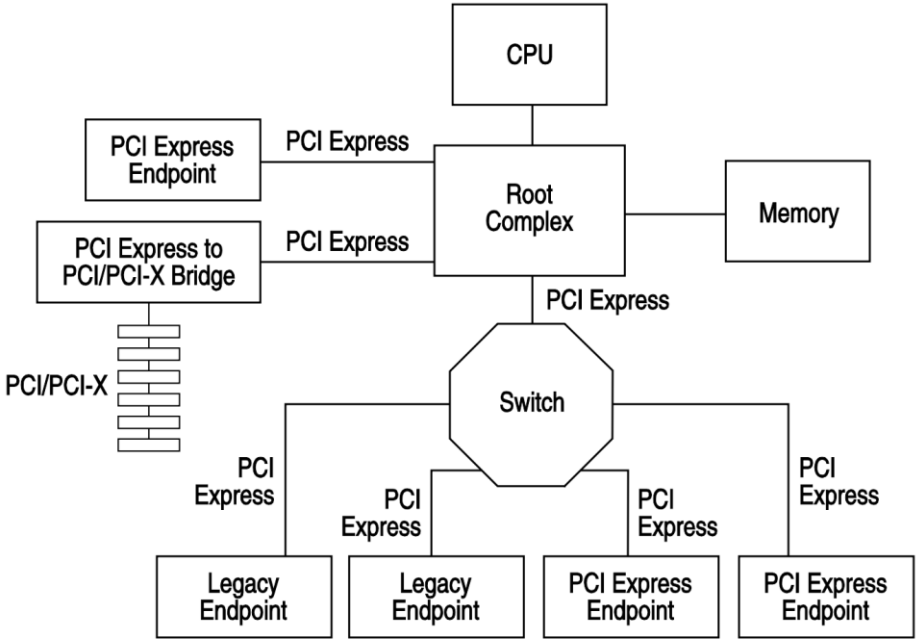
Systems Engineering Manager, Texas Instruments
cmorrison@ti.com | www.linkedin.com/in/casey-morrison | @CaseyTMorrison

Cindy Cui

Application Engineer, Keysight Technologies
Cindy_cui@keysight.com

PCI-Express Gen-4 Overview

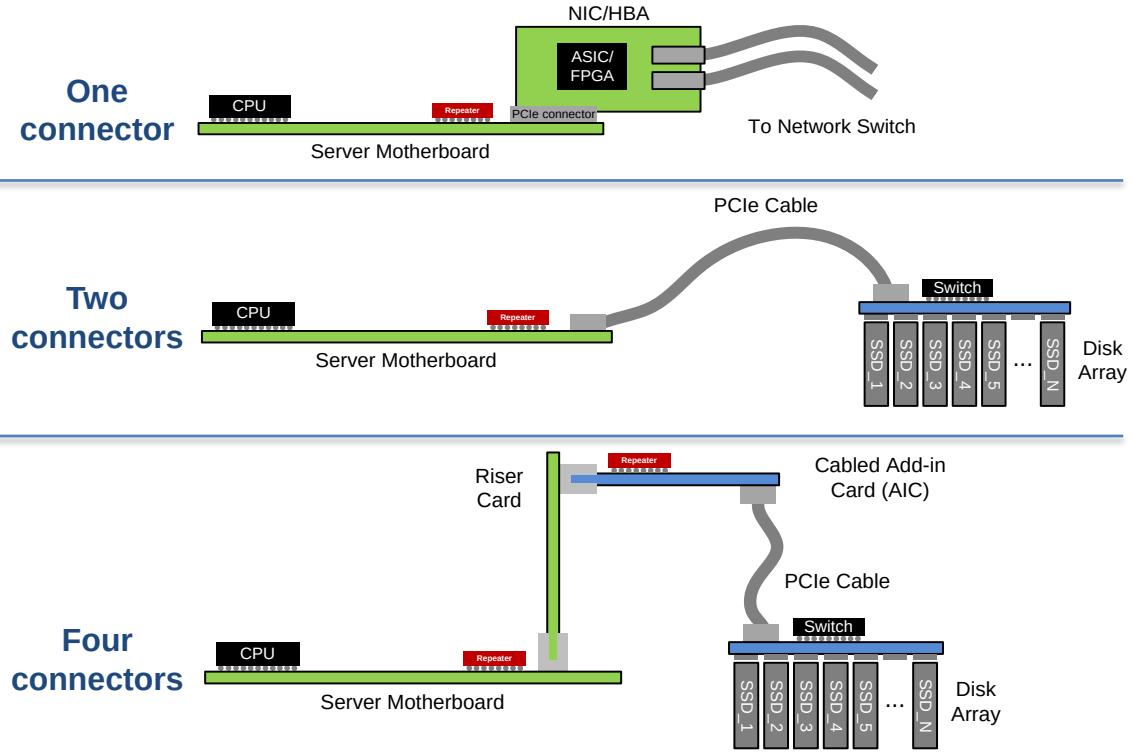
- **PCI-Express Gen-4:**
  - Base Specification revision 4.0 expected to reach version 1.0 in 2017
  - Maximum speed: 16 Gbps / Lane / direction
  - Single- or multi-lane links to scale aggregate bandwidth: x1, x2, x4, x8, x16, and x32
- **Applications:**
  - Servers: CPU-to-network and CPU-to-storage interconnects
  - Client compute: CPU-to-peripheral (i.e. graphics card) interconnect
  - High-performance compute / Compute clusters: CPU-to-CPU interconnect



Example PCIe Topology (From Base Specification)

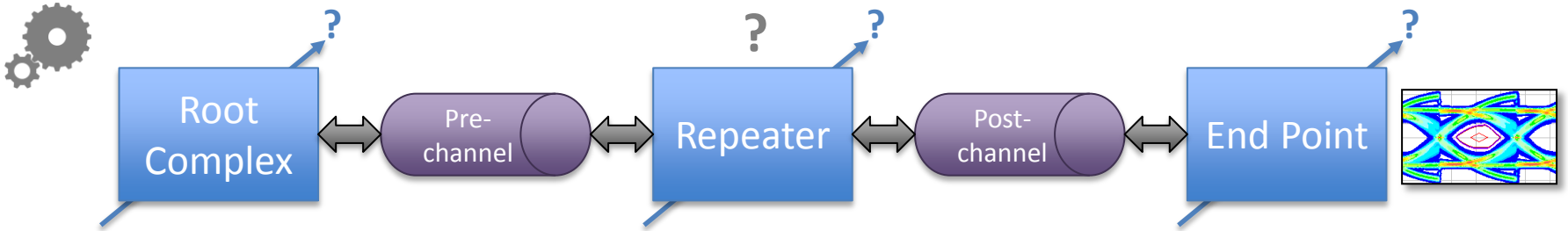
PCI-Express Gen-4 System Topologies

- **System designers anticipate various topologies ranging from *one to five connectors*.**
- **High channel attenuation:**
  - ~5 dB from CPU package
  - ~3 dB from End Point package
  - ~1 dB from each connector
  - ~1 dB / inch from PCB
- ***Linear Repeaters* are commonly used to achieve reach extension while minimizing added *latency, cost, and power* consumption.**
- **System designers need a way to gain confidence in their chosen topology and its viability.**
- ***System-level simulations are one way to achieve this goal.***



End-to-End System-Level Simulations

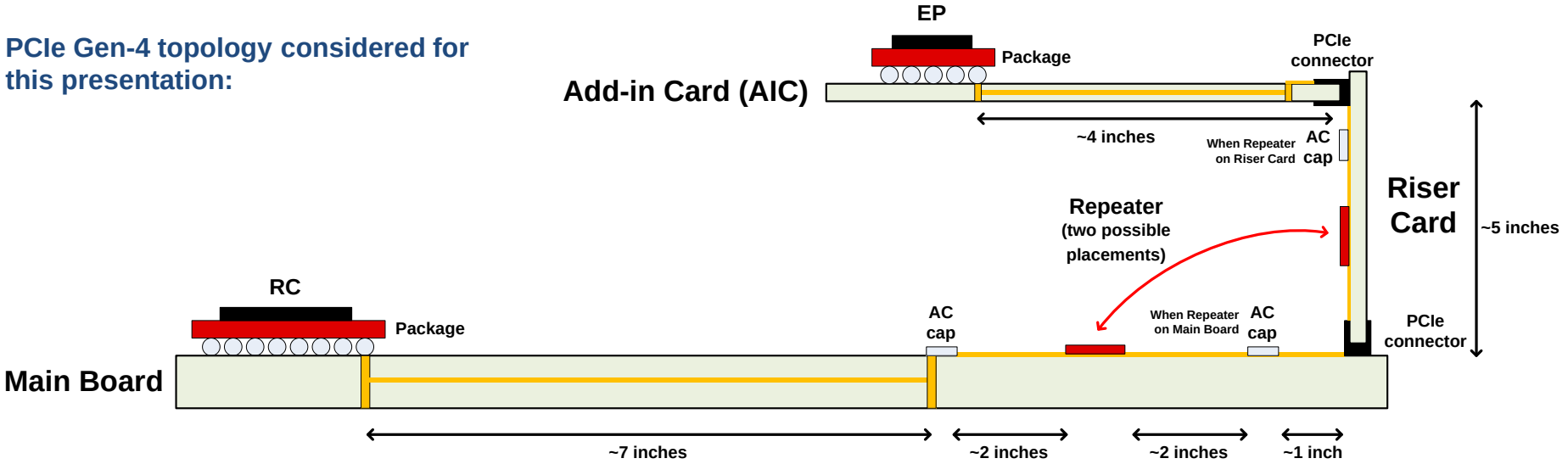
- **The proposed methodology for simulating a *Tx+Repeater+Rx system* in the context of PCIe Gen-4:**
  1. Determine if a Repeater is required.
  2. Define a simulation space.
  3. Define evaluation criteria.
  4. Execute the simulation matrix and analyze the results.
- **<u>Goal</u>: Reach a conclusion regarding the optimum configuration of the system in an efficient and timely manner.**



Step 1: Determine if a Repeater is Necessary

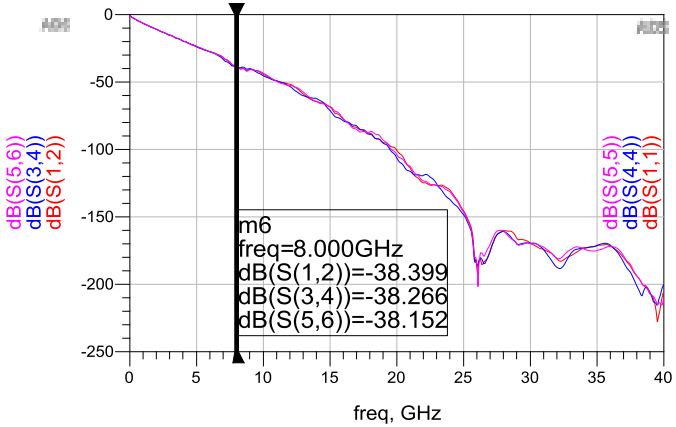
Two ways to determine if a Repeater is necessary:	
1. Compare end-to-end channel loss to PCIe channel requirements	2. Simulate end-to-end channel s-parameter in Seasim

PCIe Gen-4 topology considered for this presentation:

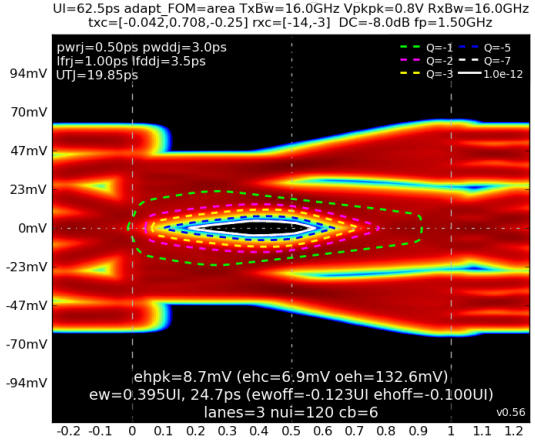


Step 1: Determine if a Repeater is Necessary

S_{DD21} Insertion Loss



Seasim Analysis



Channel analysis method	Value	PCIe Requirement	Conclusion
1. End-to-end channel insertion loss	38.8 dB at 8 GHz	≤ 20.5 dB at 8 GHz	Repeater is required
2. Channel simulation with behavioral Tx, Rx, and package	EH = 8.7 mV EW = 0.395 UI	EH ≥ 15 mV EW ≥ 0.3 UI	Repeater is required

Step 2: Define a Simulation Space

The system-level simulation task is broken down into two sequential phases:

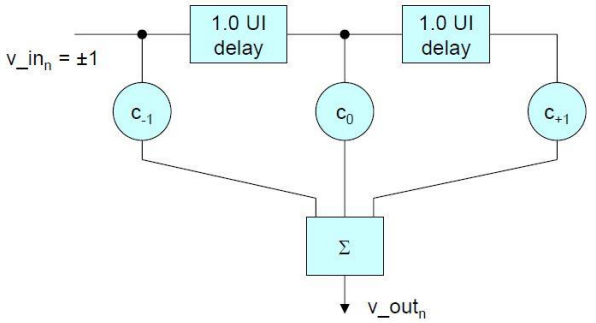
1. **Initial link performance analysis**. Analyze the impact of Repeater placement and Tx/Repeater/Rx settings on link performance.
2. **Sensitivity analysis**. Quantify the sensitivity of link performance to process/voltage/temperature (PVT) variation and to variations in Repeater placement.

Simulation Phase	RC Tx Parameters	Repeater Parameters	EP Rx Parameters	Channel Parameters
1. Initial link performance analysis	**Presets:** 0, 1, ..., 9 **VOD:** 1000 mVppd	**Boost:** Sweep six values **Wide-band gain:** -1 dB	**Rx parameters automatically adaptive**	**Channel topologies considered:** **1. Repeater placed on Main Board** **2. Repeater placed on Riser Card**
2. Sensitivity analysis	**Presets:** 0, 1, ..., 9 **VOD:** 1000 mVppd	**Boost:** Optimum setting from Phase 1 **Wide-band gain:** ±4 dB	**Rx parameters automatically adaptive**	**Focus on optimum topology. Vary specific Repeater placement by ±2 inch to assess sensitivity to placement.**

Root Complex Tx Parameters

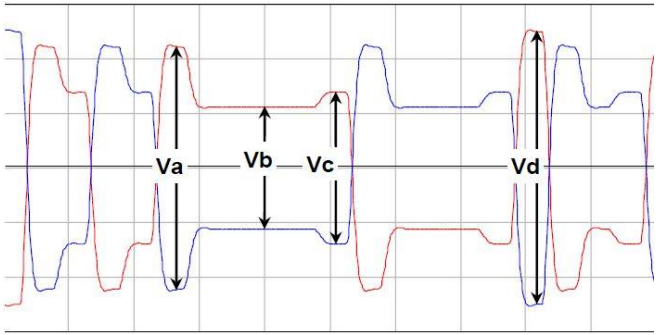
- **A Xilinx FPGA SerDes is used as the Root Complex Tx in this analysis.**
- **It implements a *three-tap FIR filter* which can be configured to achieve any of the ten *PCIe Presets*.**

Preset #	Pre-shoot (dB)	De-emphasis (dB)	c_{-1}	c_{+1}	Va/Vd	Vb/Vd	Vc/Vd
P4	0.0	0.0	0.000	0.000	1.000	1.000	1.000
P1	0.0	-3.5 ± 1	0.000	-0.167	1.000	0.668	0.668
P0	0.0	-6.0 ± 1.5	0.000	-0.250	1.000	0.500	0.500
P9	3.5 ± 1	0.0	-0.166	0.000	0.668	0.668	1.000
P8	3.5 ± 1	-3.5 ± 1	-0.125	-0.125	0.750	0.500	0.750
P7	3.5 ± 1	-6.0 ± 1.5	-0.100	-0.200	0.800	0.400	0.600
P5	1.9 ± 1	0.0	-0.100	0.000	0.800	0.800	1.000
P6	2.5 ± 1	0.0	-0.125	0.000	0.750	0.750	1.000
P3	0.0	-2.5 ± 1	0.000	-0.125	1.000	0.750	0.750
P2	0.0	-4.4 ± 1.5	0.000	-0.200	1.000	0.600	0.600



$$v\_out_n = v\_in_{n-1}c_{n-1} + v\_in_n c_n + v\_in_{n+1}c_{n+1}$$

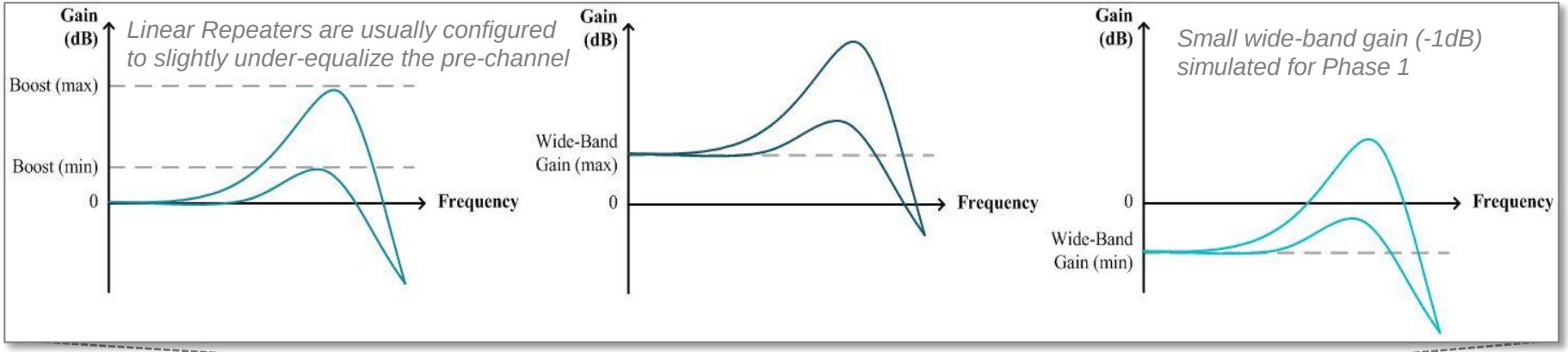
$$|c_{-1}| + |c_0| + |c_{+1}| = 1 \quad c_{+1} \le 0 \quad c_{-1} \le 0$$

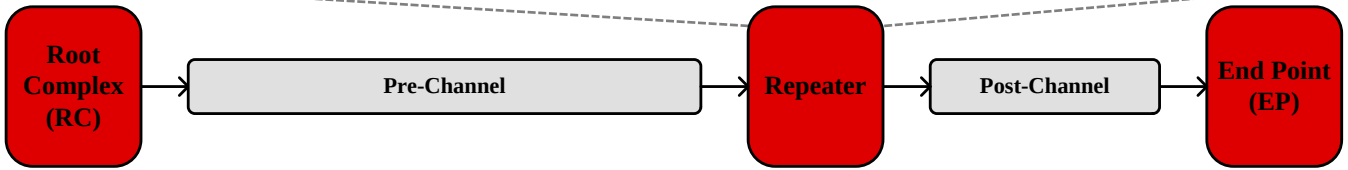


De-emphasis = 20 $\log_{10}$Vb/Va
Preshoot = 20$\log_{10}$Vc/Vb
Boost = 20$\log_{10}$ Vd/Vb

Repeater Parameters

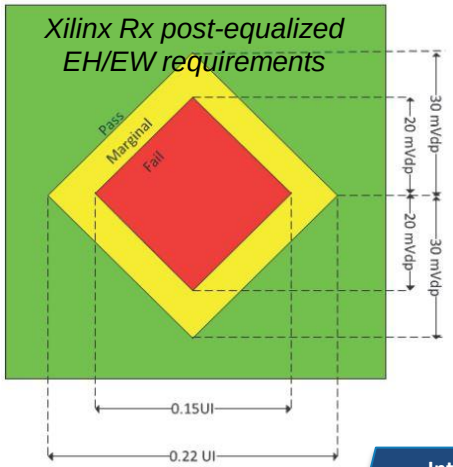
- **A Texas Instruments Linear Repeater is used to extend the reach between the RC and EP.**
- **Linear Repeaters conventionally provide two mechanisms for signal conditioning: *High-frequency boost* and *wide-band amplitude gain.***





Step 3: Define Pass/Fail Criteria

- **Bit error rate (BER) is the ultimate gauge of link performance, but an accurate measure of BER is not possible in relatively short, multi-million-bit simulations.**
- **Instead, the methodology proposed here uses two criteria to establish link performance:**
  1. A link must meet receiver's EH and EW requirements
  2. A link must meet criterion 1 for all Tx Preset settings

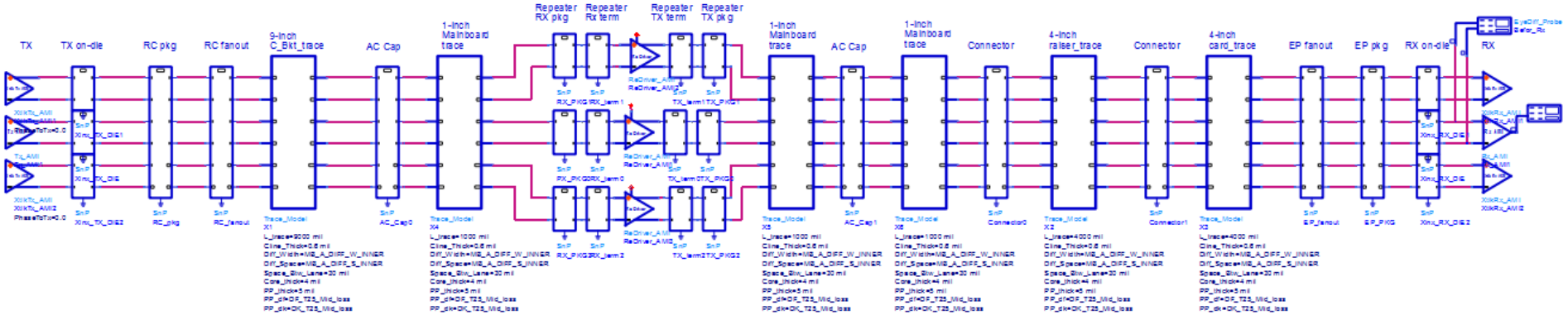


Criterion 1 establishes that the there is a viable set of settings which will result in the desired BER.

Criterion 2 ensures that the link has adequate margin and is not overly-sensitive to the Tx Preset setting.

Step 4: Execute and Analyze

- **IBIS-AMI models are used for each active component: RC and EP SerDes from Xilinx and Linear Repeater from Texas Instruments.**
- **Keysight ADS is used to execute the IBIS-AMI simulations, measure the extrapolated EH and EW, and plot post-equalized eye.**



Simulation Schematic used for this Analysis

Step 4 (Phase 1): Initial Link Performance Analysis

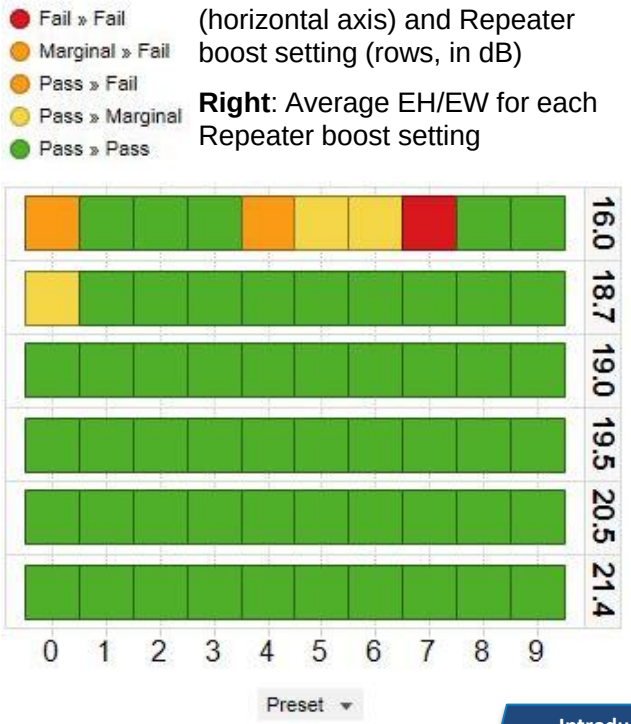
- **The focus of Phase 1 is to run a broad set of relatively short simulations to explore the design solution space.** ***Minimizing the simulation time for each simulation is crucial.***

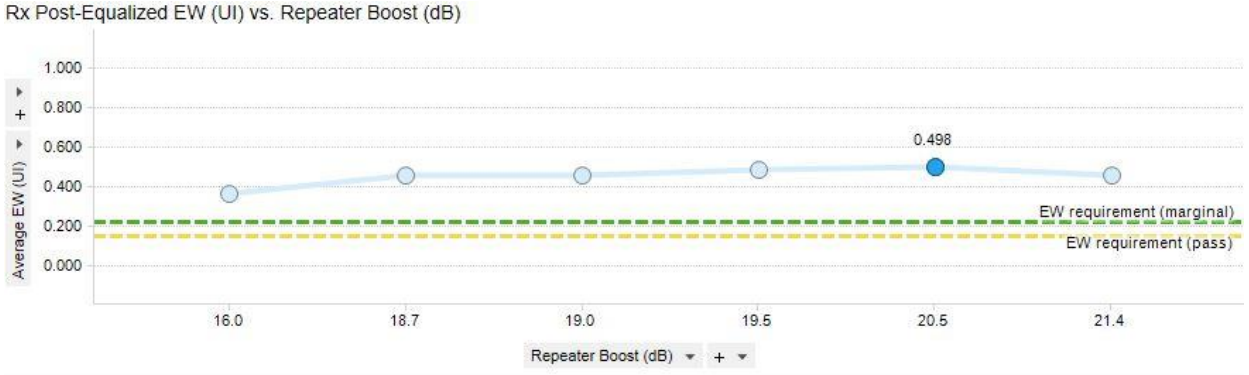
Simulation Parameter	Value
Data Rate	16.0 GT/s
Data Pattern	PRBS31
Total number of bits	1 Million
Crosstalk	Yes. Two far-end crosstalk (FEXT) aggressors.
Ignore_Bits	500k *Note: This is set by the Rx model*
Simulation type	Time domain (a.k.a. bit-by-bit) *Note: Simulations will be faster running in Statistical mode, however non-linear behavior may not be adequately represented.*
Bit-by-bit extrapolation	Enabled *Note: Simulations will be faster without this mode enabled, however RJ will not be accounted for as accurately.*

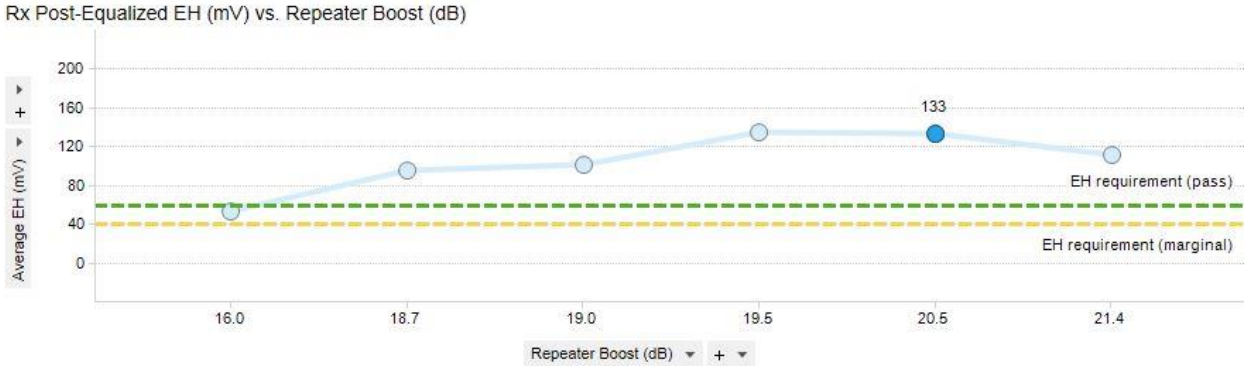
Step 4 (Phase 1): Initial Link Performance Analysis

Bottom-left: EH and EW pass/fail result for each Tx Preset (horizontal axis) and Repeater boost setting (rows, in dB)

Right: Average EH/EW for each Repeater boost setting

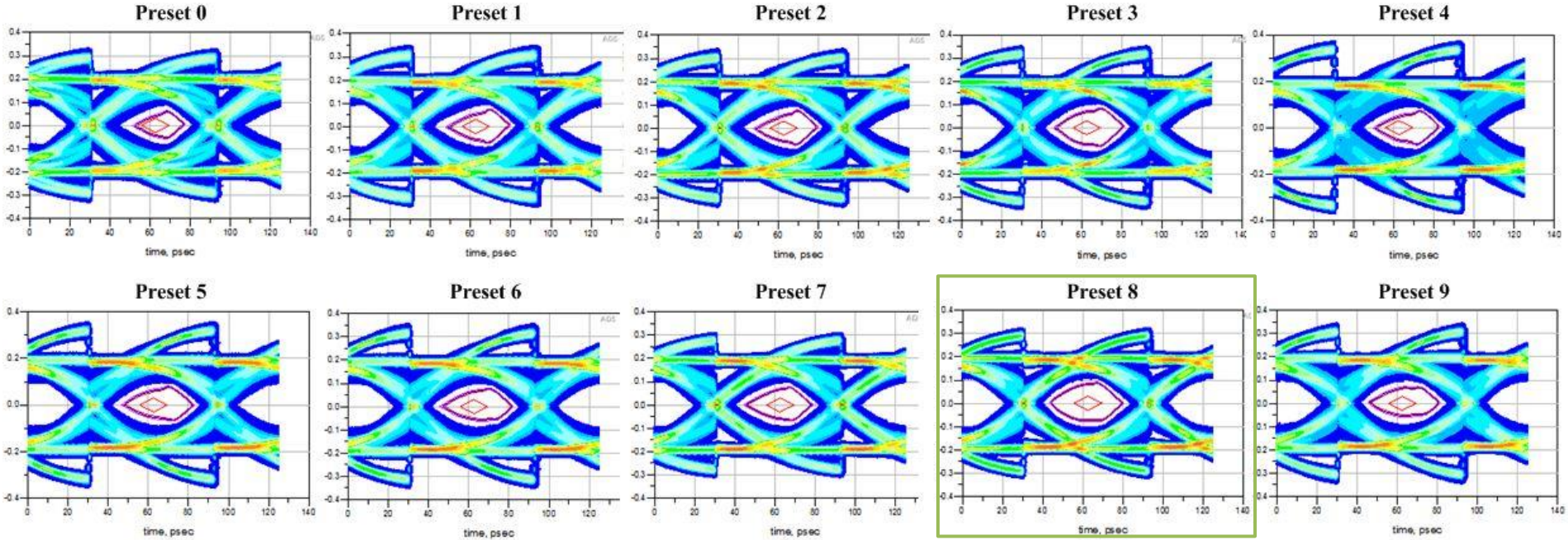






Step 4 (Phase 1): Initial Link Performance Analysis

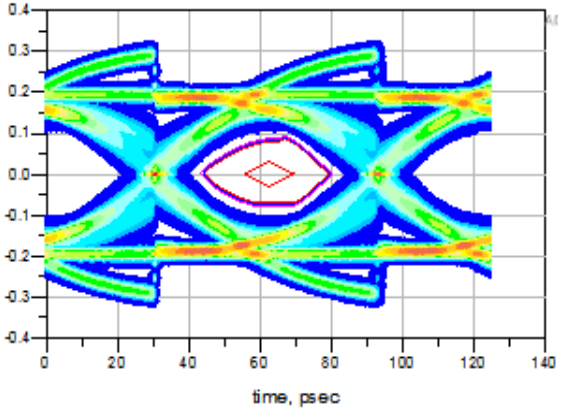
Post-equalized eye diagram for all Tx Preset settings. The optimum Repeater boost (20.5 dB) is used for these cases. Tx Preset 8 yields the largest post-equalized eye opening.



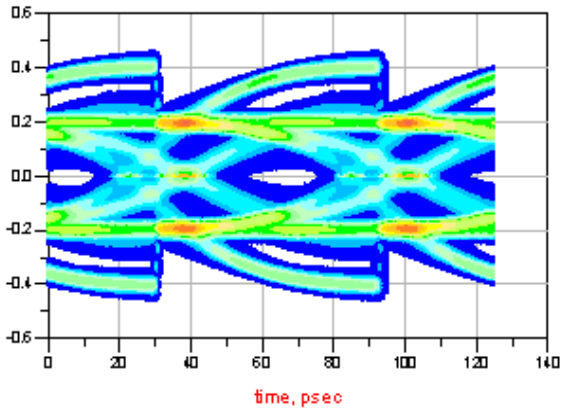
Step 4 (Phase 1): Initial Link Performance Analysis

- A similar analysis is conducted for the alternate placement: Repeater on the Riser Card.
- In this configuration, the pre-channel loss is ~25 dB at 8 GHz; post-channel loss is ~14 GHz at 8 GHz.
- This placement shows consistently reduced performance compared to Main Board placement of the Repeater.

Repeater on Main Board, Preset 8, Boost=20.5 dB

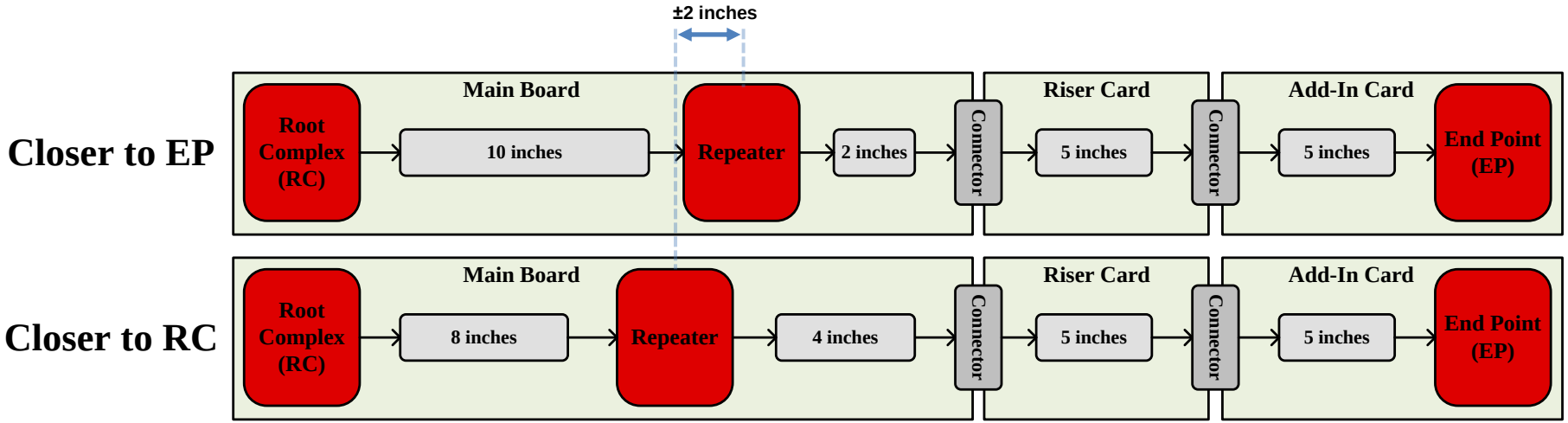


Repeater on Riser Card, Preset 9, Boost=20.5 dB

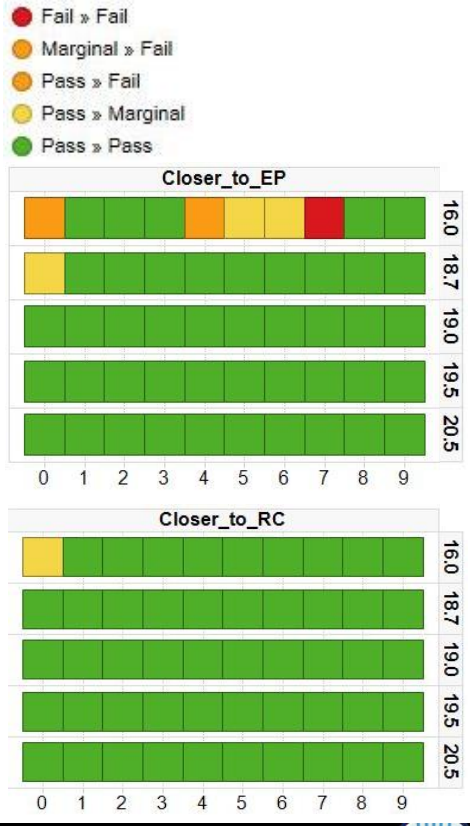


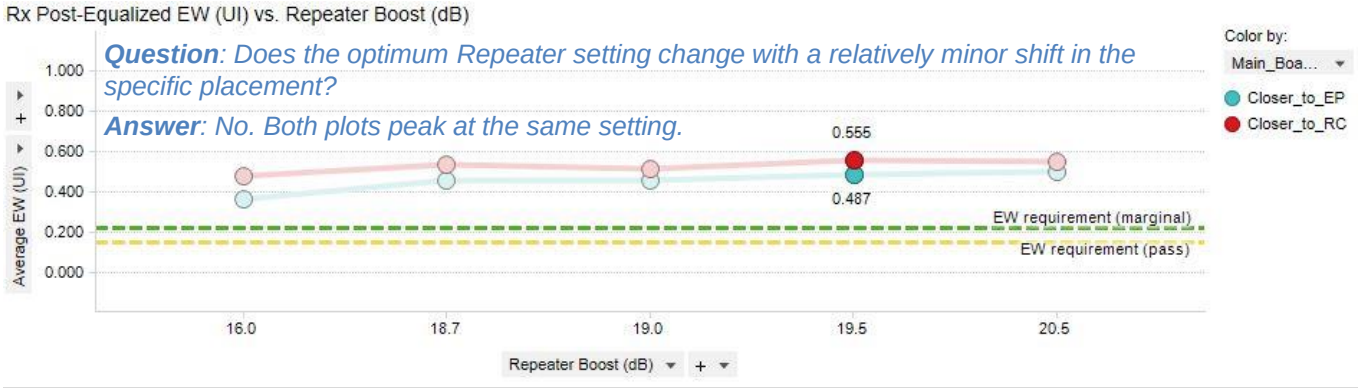
Step 4 (Phase 2): Sensitivity to Placement

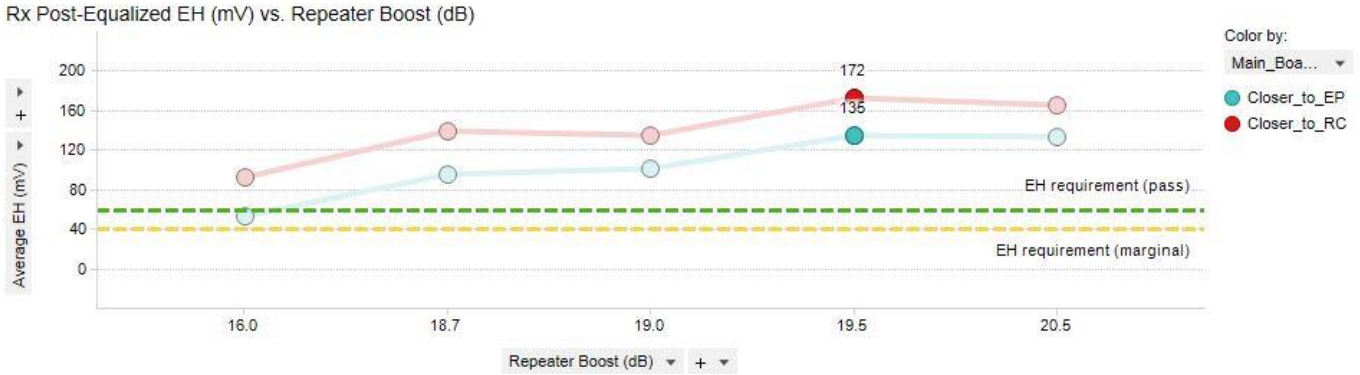
- **Optimum Repeater placement is on the Main Board.**
- **Analyze the sensitivity of link performance due to the specific placement of the Repeater.**
- **Simulations are run with a *±2 inch variation in Repeater placement*.**



Step 4 (Phase 2): Sensitivity to Placement







Step 4 (Phase 2): Sensitivity to PVT Variations

- **The last part of the Phase 2 sensitivity analysis is to look at the sensitivity of link performance to process, voltage, and temperature (PVT) variations.**
- **To exacerbate the effects of PVT variation, the Repeater's wide-band gain is varied by ±4 dB.**
- **Overall link performance is not affected until *both* extremes of PVT variation and wide-band gain are realized.**
- **As long as the Repeater's wide-band gain setting is kept to a reasonable, mid-level value, the link performance will be robust across PVT corners.**

Conclusions

A simple four-step process for evaluating *Root Complex + Repeater + End Point* PCIe Links:

Determine if a Repeater is necessary

- Does end-to-end channel exceed PCIe Base Specification?
- Does channel fail Seasim EH/EW analysis?

Understand if reach extension device is needed.

Setup simulation sweep space

- Include all Tx Presets
- Chose Repeater boost around pre-channel loss

Chose limited set of parameters which are most likely to impact system performance.

Define a pass/fail criteria

- Final Receiver's post-equalized eye height (EH) and eye width (EW)
- Achieving Rx EH/EW requirements across *all Tx Presets*

Define evaluation criteria up front to avoid subjective conclusions later on.

Execute simulation matrix

- **Phase 1: Initial link performance analysis (simulating the sweep space)**
- **Phase 2: Understand link's sensitivity to specific Repeater placement and PVT**

Break execution into two phases to minimize overall simulation time.

Thank you!

QUESTIONS?

Appendix: Terminology

- **Root Complex (RC)** A defined System Element that includes at least one Host Bridge, Root Port, or Root Complex Integrated Endpoint.
- **End Point (EP)** One of several defined System Elements. A Function that has a Type 00h Configuration Space header.
- **Linear Repeater** An analog reach extension device which generally provides continuous time linear equalization (CTLE) and wide-band amplitude gain.
- **Link** The collection of two Ports and their interconnecting Lanes. A Link is a dual-simplex communications path between two components.
- **Link Segment** The collection of a Port and a Pseudo Port or two Pseudo Ports and their 45 interconnecting Lanes. A Link Segment is a dual simplex communications path between a Component and a Retimer or between two Retimers (two Pseudo Ports).