PCI Express® 4.0 Electrical Previews

Dean Gonzales
Advanced Micro Devices
Disclaimer

The information in this presentation refers to specifications still in the development process. This presentation reflects the current thinking of various PCI-SIG® workgroups, but all material is subject to change before the specifications are released.
Outline

- PCIe® 4.0 Motivation & Overview
- PCIe Channel Description
- Transmitter
- Receiver
- Reference Clock & SRIS
- Design & Simulation
PCIe 4.0 Motivations and Assumptions

- We continue to see a requirement to increase PCIe bandwidth
  - Networking, Storage, High Performance Computing
- Motivations for PCIe 2.x->3.0 apply equally for 3.0->4.0
- Eco-system impact of a new generation drives requirement for ≥2x increase in delivered bandwidth
- Desirable to extend PCIe 3.0 infrastructure and PHY architecture for another generation
  - Moving to a new infrastructure such as electrical or optical waveguides likely breaks backwards compatibility
  - Highly desirable to preserve current usage models
  - With incremental improvements 3.0 PHY architecture is capable of higher data rates
- CEM form factor is the most important usage model of PCIe
  - Can be extended another generation with incremental improvements
PCIe 4.0 Overview

- Key attributes of PCIe 4.0
  - 16 GT/s, using scrambling, same as 8GT/s
  - Maintains backward compatibility with installed base of PCIe devices
  - Limited channel reach: approx. 12” one connector
  - Longer channels require retimers or lower loss channels

- New features
  - Uniform spec methodology applied across all data rates (as possible)
  - Support for independent Refclk clocking mode with SSC (SRIS)
  - Integration of Retimer ECN

- This presentation focuses on items adopted in the 0.5 specification
  - Transmitter
  - Reference Clock
  - Retimer
  - Receiver
  - Channel
Channel
PCI Express Channels

- Card Electromechanical (CEM) form factor
  - Most widely adopted PCIe implementation
    - CEM spec sets limits and compliance measurement boundaries
  - Client CEM
    - Short/medium (3-12”), reflection and crosstalk dominated
  - Server CEM
    - Medium/long (20”) loss dominated

2-multiplier server
- Typ 8-12 layers
- Typ 92mil thick

CEM client
- Typ 4-6 layers
- Typ 62mil thick
Channel Loss Characteristics

- Package substrate & PCB loss per inch model used for channel study
  - Tan delta [0.015 to 0.025]
  - Copper conductivity & roughness
  - Temperature & humidity variation

- PCIE3 CEM cal channel RH & temp variation measurement @ 8GHz
  - 12.5” Riser
  - 4” Mainboard
  - 2” CLB
- Impedance discontinuities apparent in TDR response
- Frequency domain response shows complex die-channel interaction
PCIe Connector/Card Enablers

- CEM connector performance criteria for 16GT/s
  - Any hardware change must preserve backwards compatibility for all data rates (2.5, 5, 8, & 16 GT/s)
  - Continue to enable standard plated thru-hole configurations
  - Qualify a common footprint for surface mount style connectors

Model of current CEM connector:

Frequency Domain Response:
PCIe Connector/Card Enablers

- Plated-thru hole connector
  - Resonance near 8GHz causes excessive crosstalk & reflection
  - Agreement between model and measurement
  - Connector fixture de-embedding reference plane is PCB topside

Full-wave connector model s-parameters analyzed in SEASIM.
Note: connector performance boundary is top surface of PCB.
The CEM form factor is the most important usage model for PCIe
- Can be extended by another generation with improvements to CEM connector launch

Current CEM channels are electrically very complex beyond 8GT/s
- Discontinuities and crosstalk from packages, sockets, vias, etch, coupling capacitors, CEM connector
- Non-monotonic frequency domain behavior yields unpredictable data rate scaling

To extend current infrastructure requires enabling SIG membership to design and build ‘cleaner’ channels
- Tuning via launches, minimizing layer transitions, careful layer choices
- For longer reach channels, back-drilling, lower loss materials and repeaters/re-timers will be required

Target max length PCIe 3.0 (8 Gb/s) server channel is ~20” with 1 or 2 connectors

Path finding for 16GT/s shows ~12” with 1 connector
Frequency Domain
Channel Parameters

Normative

Informative
Transmitter
Transmitter Specification

- Preset definition
  - Retain P0-P10 with same definition as PCIe 3.0 at 8GT/s

- Package loss (ps21TX)
  - Informative for root complex devices, normative for AIC devices

- Architecture Specific Post Processing
  - Embedded vs. non-embedded, Common vs. Independent Refclk architectures

- Jitter parameters
  - Applied uniformly for all four data rates
  - Number of normative parameters reduced
  - Informative parameters added

- Return Loss extended up to 8GHz
  - Same limits as at 4.0 GHz
  - T-coils likely required to meet limits
- CEM Spec Defines Tx Requirements for Chip + Interconnect
- No Separate Tx Chip Or Interconnect Only Requirements.
Summary of Base vs. CEM Differences for Tx Testing

- CEM Tx Testing is at the end of the CEM reference channel
  - Eye can already be closed at CEM connector with long channel motherboards
  - Waveform-based test with reference equalizer application in post-processing was chosen as the only option to assess overall Tx interoperability of channel plus silicon

- CEM Tx test is an eye test @ BER $\leq 10^{-12}$ after applying the reference equalizer
  - No jitter decomposition beyond Rj/Dj due to end of channel reference point

- CEM Tx eye test only required to pass with “best” preset
Summary of Base vs. CEM Differences for Tx Testing

- Motherboard Tx test is done with real motherboard clock (not a clean/lab reference clock)
  - Do not want to add cost/complexity and require a motherboard to provide method for external clock source
  - Want a method that can test real, off-the-shelf motherboards

- Motherboard Tx test is done by sampling data lane under test and 100 MHz reference clock simultaneously (dual port methodology)

- Explore consistency between Base (4.0) and CEM (4.0)
  - Specified/Recommended measurement/calibration method for eye after reference equalizer
  - Study possible test/reference channel commonality
Receiver
Receiver Specification

- Stressed eye methodology applied to all data rates
  - Stressed jitter and voltage as a single test
- Calibration channel defined by data rate dependent mask
  - Current direction to make variable at 16GT/s
  - Minimize Rj/Sj/DM variation across different set-ups
- Separate Root Complex and AIC behav pkg models
- Behavioral Rx equalization data rate dependent
  - 2.5 and 5.0G: none
  - 8.0 and 16.0G: CTLE and DFE (8G: 1 tap, 16G, 2 taps)
- Eye height minimum reduced to 15mV for 16G
Receiver Linear Equalizer

\[ H(s) = \frac{sC_1R_1 + R_2}{sC_1R_1R_2 + R_1 + R_2} \cdot \frac{1}{sC_2R_3 + 1} \]

\[ G_{DC} = \frac{R_2}{R_1 + R_2} \]
\[ \omega_2 = \frac{1}{C_1R_1} = G_{DC}\omega_p \]
\[ \omega_p = \frac{R_1 + R_2}{C_1R_1R_2} \]
\[ \omega_p = \frac{1}{C_2R_3} \]

Graphs showing the magnitude and phase response of the equalizer for different frequencies.
EQ Tuning – High Loss Two Connector Server
EQ Tuning – Low Loss Two Connector Server

PCI-SIG Developers Conference

Copyright © 2015, PCI-SIG, All Rights Reserved
Equalization Sweep

![Graphs showing Peak Eye Height, Eye Width, Centered Eye Height, and Eye Horizontal Offset against various Tx Tap Resolution and Channel PCB Variation settings.](image)

Optimal CTLE Settings

- DC Gain [dB]
- Fpole [GHz]

Tx Tap Resolution
- [1/24, 1/32, 1/64, 1/96]

Channel PCB Variation
- [Tline1_Tline2_Tline3]
Receiver
Stressed Eye Calibration

Generator swing is 2x amplitude due to 6dB reduction of splitter.

TP2P’ stressed eye has reduced instrument noise compared to the TP2P eye.
Reference Clock
Refclk Specifications

- Architecture independent parameters
- Architecture dependent parameters
  - Common Clock (CC) and Independent Reference Clock (IR) filter functions
    - IR with SSC (SRIS) defined in 3.0 ECN and 4.0 specification
  - Explicit listing of all combinations of PLL and CDR limits that need to be evaluated
- Normative limits for CC
- Informative limits for IR (may be removed)
- A PHY may support one or more modes
### Architecture Independent Parameters

- Note that $T_{\text{SSC\_MAX\_PHASE\_SKEW}}$ is no longer defined.
- There is now an explicit freq vs. amplitude mask for SSC profile phase jitter.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Limits</th>
<th>Units</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{\text{REFCLK}}$</td>
<td>Refclk Frequency</td>
<td>99.97 (min), 100.03 (max)</td>
<td>MHz</td>
<td></td>
</tr>
<tr>
<td>$F_{\text{SSC}}$</td>
<td>SSC frequency range</td>
<td>30 (min), 33 (max)</td>
<td>KHz</td>
<td></td>
</tr>
<tr>
<td>$T_{\text{SSC-FREQ-DEV}}$</td>
<td>SSC deviation</td>
<td>-0.5 (min), 0.0 (max)</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>$T_{\text{TRANSPORT-DELAY}}$</td>
<td>Tx-Rx transport delay</td>
<td>12 (max)</td>
<td>Nsec</td>
<td>1</td>
</tr>
<tr>
<td>$T_{\text{SSC-MAX-FREQ-SLEW}}$</td>
<td>Max SSC df/dt</td>
<td>1250</td>
<td>ppm/usec</td>
<td>2</td>
</tr>
</tbody>
</table>
Low Frequency Reference Clock Jitter Limits (Mask)

Reference Clock Phase Jitter Limit

- 30-33 kHz: 2500 ps
- 100 kHz: 1000 ps
- 500 kHz: 25 ps

Jitter Pk-Pk (ps)

Freq (Hz)

10^4 to 10^6
Refclk Topologies

Common Refclk

Data Clocked

Transfer function:

\[ H(s) = \frac{s^2}{s^2 + 2s\zeta_1\omega_n + \omega_n^2} \]

Jitter at Rx latch:

\[ J(s) = X(s) * H_1(s) * H_3(s) \]

Refclk jitter = X(s)

Transfer function:

\[ H(s) = H_1(s) * H_3(s) \]

Jitter at Rx latch:

\[ J(s) = X(s) * H_1(s) * H_3(s) \]
Refclk Topologies (cont.)

IR Reference Clock

Transfer Function: \( H(s) = \left( H_1(s) + H_2(s) \right) * H_3(s) \)

Jitter at Rx latch: \( J(s) = [X_1(s)H_1(s) + X_2(s)H_2(s)] * H_3(s) \)

- Incorrect for Clock Test – Options to fix
  - \( X_1(s)H_1(s) * H_3(s) < .7 \text{ ps} (.5 \text{ ps in ECN}) \)
    - 8 GT/s (same as SRIS ECN) and 16GT/s
  - Define reference worst case \( H_2(s) \)
SRIS/IR Reference Clock Test

- Currently informative in 4.0
- Pessimistic – assumes worst case specification compliant model PLL transfer function
- Difficult to meet for current discrete clock chips – even with improved model CDR
- Should 100 MHz frequency be required/implied for a SRIS/IR only implementation?
- Reference clock test not specified by other standards with similar PHY architectures
  - USB 3.0, 3.1, SATA
- Current direction to remove altogether for SRIS/IR Mode
  - Allow maximum implementation PLL/Transmitter trade-off
Separate Reference Clocks with Independent SSC (SRIS)
Inexpensive Cabling & Independent SSC Reference

- Previous spec releases did not support independent reference clocks with spread spectrum
  - Estimated cable cost with integrated reference clock transmission line ~$1; double the cost of a common SATA cable
- SRIS released in PCIe Base Spec 3.0 ECN
  1) Larger elasticity buffer requirement
  2) Increased insertion frequency of SKIP ordered sets
  3) CDR transfer function spec changes; no impact to transmitter or reference clock requirements
  4) Second ECN updates Model CDRs
  5) Introduces terms for Separate Refclk Modes of Operation
    - 5600ppm (New – SRIS) and 600ppm (Existing - SRNS)
- Creates new form factor opportunities for PCIe
  - SATA Express: Connector for PCIe SSD compatible with SATA
  - Lower cost external cabled PCIe
4.0 SRIS/IR Model CDRs (First 3.x ECN CDR)

For bitrate 2.5GT/s & 5GT/s:

\[ H(s) = \frac{s^2}{s^2 + s\zeta\omega_o s + \omega_o^2}; \quad \zeta = \frac{1}{\sqrt{2}}; \]

if SRIS 2.5GT/s:
\[ \omega_o = 1.5\text{MHz} \times 2\pi \]

elif SRIS 5GT/s:
\[ \omega_o = 5\text{MHz} \times 2\pi \]

For bitrate 8GT/s & 16GT/s:

\[ H(s) = \frac{s^2}{s^2 + 2\zeta s\omega_o s + \omega_o^2} \]
\[ \zeta_1 = \frac{1}{\sqrt{2}}; \quad \zeta_2 = 1; \quad \omega_o = 10^7 \times 2\pi \]

if SRIS 8GT/s:
\[ A = 1e^7 \times 2\pi \]
\[ B = 2.2e^{12} \times (2\pi)^2 \]

elif SRIS 16GT/s:
\[ A = 9.5e^6 \times 2\pi \]
\[ B = 4.36e^{12} \times (2\pi)^2 \]
Design & Simulation
Channel Simulation

2.5G, 5.0G

37
Seasim
Tool Improvements Needed

- Channel response at >8GHz is affected by large number of features in the channel
  - Pre-layout evaluation of topology choices is a complicated multi-dimensional problem
  - Need to be able to quickly build and test many different options
  - Large number of HVM permutations need to be evaluated to determine robustness of solution

- Seasim has been enhanced to allow EWG members to efficiently evaluate these options
  - Once validated this tool will be made available to the PCI-SIG membership for 4.0 channel compliance
Improvements to Seasim

- Addition of a GUI form based interface
  - Underlying config file interface to seasim is unchanged
  - A simple form based dialogue tool added
  - Tab based interface to group config controls by context
  - Ability to save and load configurations
  - Launch (and kill) seasim from GUI

- Touchstone channel modeling
  - A set of touchstone files can be cascaded to form die-pad to die-pad channel
  - A vector of left hand and right hand ports define connections between S-parameters
  - Rx port and set of Tx ports define step responses to be generated
  - Tx amplitude and Gaussian bandwidth can be specified
  - Addition of a PCIe 4.0 include file for simulation conditions
Seasim Channel

- Allows what-if analysis on the channel components by changing the touchstone files that are concatenated together for the channel.

- Different analyses can be selected as the channel is ‘tuned’.

- Either a pre-saved config can be loaded or the pcie-gen4.inc for normal sim conditions.

- The other tabs allow simulation conditions to be changed from the default config.
Client HVM Sweeps

- At 16GT/s small changes in the channel have a big impact on eye opening
  - An end-end reflection peaks and nulls when flight time changes by 62.5ps or ~0.42”
  - Impedance variations between motherboard and add-in card introduce low frequency reflections that interact with the adapted EQ solution
  - Different root complex package responses vary significantly
  - Topology differences between top and bottom microstrip routing impact reflections
  - Different connector models impact channel reflections

- To capture solution space sensitivity channel parameters can be swept using seasim
  - Set of component touchstone files built
  - Seasim sweep capability allows large number of cases to be studied
    - In this example ~15,000 cases tested
Seasim HVM Sweep

- Seasim can be used to define HVM sweeps
- The channel model can be swept to represent manufacturing variations of impedance or loss
- To consider the impact of different PCB layout the length of different channel segments can be swept
- Seasim will launch jobs in parallel then collect results and plot them
Example Client Channel Simulation

Touchstone files:
caps_shunt_0.80pf.s12p
g3413_gx4x12p
G143xtk38xp
etch titular_s5_z70_0.60inch.s12p
etch titular_s30_z70_2.00inch.s12p
cap0.80pf.s12p
etch titular_s30_z70_5.00inch.s12p
caps_shunt_0.20pf.s12p
cpx43 committing_recessed up to 1234.8m, NEW s12p
etch titular_s30_z70_4.00inch.s12p
CM_MB_6lyrs_V45_3ports_HFSS.s12p
Tahls56A_SUB1221_PCE_Tx.s12p
caps_shunt_0.40pf.s12p
Cascaded touchstones
Rx-Tx ports = (9,10), (3,4), (9,10), (1,2), (9,10), (5,6)

Left ports:
[5,13,3,9,11]
[1,13,4,5,6]
[1,2,3,4,5,6]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]
[1,13,5,7,9,11]

Right ports:
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[7,8,9,10,11,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]
[2,4,6,8,10,12]

IL (9,10),(3,4)
FXT1 (9,10),(1,2)
FXT2 (9,10),(5,6)
su_mBag
RxRL (9,10),(9,10)
TxRL (3,4),(3,4)
IL Phase Delay

Step = caps_shunt_0.80pf.s12p..13_6590..caps_shunt_0.40pf.s12p
job = case100_11.60
UI = 62.5ps adapt_FOM = area Tx BW = 16.0GHz Rx BW = 16.0GHz
tx = [-0.042, 0.958, 0] rx = [-30, 3] DC = -6.0dB fp = 4.00GHz

PWR = 0.50ps pdd = 3.0ps
If max = 1.00ps If max = 3.5ps
UL = 19.85ps

ehpk = 39.8mV (ehc = 28.2mV oeh = 480.6mV)
ev = 0.282UJ Tj = 44.9ps (ewoff = -0.077UJ ehoff = -0.050UJ)
lanes = 3 nui = 1020 c6 = 6
Client Channel HVM Sweep

- Variables swept:
  - Motherboard:
    - Length: 1-11”
    - Impedance 70/100ohm
  - Root package
    - Length 10-30mm
    - Impedance 80/90ohm
    - Loss hi/lo
  - EP package
    - Length 10/30mm
    - Impedance 80/90
    - Loss hi/lo
  - AIC etch
    - 70/90

- 15,360 cases

Min eye height 47mV
Min eye width 0.38UI
Max IL -21dB
Thank you for attending the PCI-SIG Developers Conference Israel 2015.

For more information please go to www.pcisig.com