Optics vs Copper for In-chassis Connections @ 56-112 Gbps: is copper still a viable solution?

Eric Bogatin, Dean, Teledyne LeCroy Signal Integrity Academy, Moderator
Mitchell Fields, Avago Technologies, VP Fiber Optics Product Division
Marc Verdiell, Samtec, CTO, Samtec Optical Group
Joel Goergen, Cisco, Distinguished Engineer
Pavel Zivny, Domain Expert, Tektronix
Scott McMorrow, R&D Consultant, Teraspeed Consulting, Division of Samtec
In 2018, internet traffic will be 2.4 Petabits/sec = all movies, ever made, crossing the internet every minute.

Every one of those 2.4 Pbps pass through multiple circuit boards.
Attenuation Limited Length – Bit rate Tradeoff

Assuming:

1. **Everything else** is done exactly right
2. Can recover -25 dB attenuation at the Nyquist with equalization
3. NRZ PAM2 signaling
4. Low encoding overhead
5. Cross talk not a problem
OPTICS VS. COPPER FOR IN-CHASSIS CONNECTIONS @ 56-112GBPS: IS COPPER STILL A Viable SOLUTION?

Mitchell Fields, Ph. D. (VP/GM Fiber Optics Products Division, Avago Technologies)
Has anyone been crazy enough to use optics in a chassis?

~2000
Sycamore Networks SN16000
2.5G blade to blade

~2010
IBM POWER775 HPC system
10G drawer to drawer
At every new speed bump, we ask if optics will displace copper…

Metrics drive the decision

Bandwidth
Cost
Density
Reach
Power
Why are we even discussing optics?

Common mantra:
Electrical interconnects don’t scale… …or do they?

For in-chassis, electronics have already won at 56G (would not have stated this 2 years ago).

What will happen at 100G?
What does this stuff look like?

Bottom line: today’s optics are too complicated

Significant innovation happening on both sides, with many more $’s poured into electronics.
The “holy grail” solution:
Remove the redundant functionality between the ASIC and optical module to simplify the electrical interface

Photonic devices on same substrate as ASIC (VCSEL or silicon photonics) to avoid CDRs and strong EQ.

Significant challenges:
Device design
Coupling of light
Thermal management

CMOS ASIC with hundreds of SERDES, electrical interface optimized for power and performance to directly drive photonic devices

*Better yet, integrate photonics with CMOS process, but that is WAY off...
FUTURE HIGH SPEED INTERCONNECTS: FIBER VS. COPPER?

Marc Verdiell – CTO, Samtec Optical Group
Historical perspective

• Ultra-long distance: submarine communications
  – All optical since TAT-8 (1988), 2.5 Gb/s. Now >100 Tb/s.
• Long distance terrestrial
  – >95% optical since mid 90s, what’s left is microwave, no copper
• Metropolitan area networks
  – all turned optical in the 2000’s
• Fiber to the home/premises
  – still ongoing
• Datacenters
  – Now all optical between racks, 10, 40 and 100 Gb/s
  – All supercomputers rely on optical links
• Board interconnects
  – Happening now at 10, 16 and 28 Gb/s per lane
Increased dual Copper/Optics solutions

- QSFP (usually: copper < 5-10m < optical)

- PCIe Gen2/3 (usually: copper < 3m < optical)

- Onboard Interconnects (Copper/Optical)
  - Optics being used for very short, wide links
Mid-Board Optics Interconnect System

• Supports Copper and Optical (e.g. FireFly Flyover)
  – x4, x12 at 14G and 28G
  – Can be arranged in large, dense arrays
Active Copper Solutions

- Copper reach with (i.e. Firefly)
  - Regular Micro-coax
  - Passive Equalized
  - Active Equalized
## Copper vs. Optical

### Reasons to choose optical
- Distance
- Bandwidth density
- Cabling density
- Improved signal integrity/ much easier design

### Comparison Table

<table>
<thead>
<tr>
<th></th>
<th>Cost</th>
<th>Reach</th>
<th>Inside Box</th>
<th>Rack to Rack</th>
<th>Rack + 100M</th>
<th>Electrical Isolation Advantage</th>
<th>Thermals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cu FireFly™</strong></td>
<td>&quot;Low&quot;</td>
<td>1M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>&quot;Cold&quot;</td>
</tr>
<tr>
<td><strong>Passive Eq Cu FireFly™</strong></td>
<td>&quot;Low&quot;</td>
<td>3M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>&quot;Cold&quot;</td>
</tr>
<tr>
<td><strong>Active Eq Cu FireFly™</strong></td>
<td>&quot;Medium&quot;</td>
<td>5M</td>
<td>Yes</td>
<td>TBD</td>
<td>No</td>
<td>No</td>
<td>&quot;Warm&quot;</td>
</tr>
<tr>
<td><strong>Optical FireFly™</strong></td>
<td>&quot;High&quot;</td>
<td>~150M</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>&quot;Hot&quot;</td>
</tr>
</tbody>
</table>
Example: Recent Customer FPGA Boards

- New trend – these boards would have been copper 18 month ago
- Better signal integrity, much simplified design
- Can scale up
Conclusion

• Optics starting to penetrate short distance interconnects

• Copper flyover solution evolve to meet the challenge

• New dual copper/optical compatible solutions

• Optics has a lot of dry powder in reserve
  
  – Path to many Tb/s!
Optics vs Copper for In-chassis Connections @ 56-112 Gbps: is copper still a viable solution?

Pavel Zivny, Domain Expert, Tektronix
PAM4 is hard at 26 GBd... it’s very hard at 53 GBd

1. Electrical signaling for *interconnect* at 26 GBd is here – PAM2 for a while, for PAM4 there are several chips in beta.

2. Electrical signaling for *backplane* (ca. 40 dB loss chip-chip) however... PAM2 ships, PAM4 ... not so much.

3. Optical interconnect & backplane: performance exists for *everything* ... but nobody ships *anything* ... well, almost We’ve seen slides of optical backplanes at this forum. But shipments are elusive.

4. To *measure* any of this (electrical or optical) 26 GBd, you need a oscilloscope with about 50 GH BW; achievable in both electrical and optical

Note: graph simplified to show both electrical and optical, PAM2 and PAM4 BW need as the same. Optical: true BW shown (not the electrical BW shown in ORR specifications).
So how hard is 50+ GBd? … and how do you measure it?

At 53 ... 56 GBd the Optical exists in proprietary designs, but standard is underway (400GBASE-DR4). Electrical signaling for interconnect at 26 GBd is here – PAM2 for a while, for PAM4 there are several chips in beta.

1. For optics the bandwidth is do-able: see graph on the right – meets the traditional ORR (Optical Reference Receiver) spec for 56 GBd. You can buy this today. (Tek 80C10C module)

2. To measure the electrical (which is not in any standard), you need a oscilloscope with about 60+ GH BW; achievable ... but hard for a smooth roll-off to 1.5 x that value! (1mm connectors etc.)
50+ GBd measurement examples

Why do you want the smooth rolloff (left) over brick wall (right). Both 70 GHz el. oscilloscopes.

Electrical eyes: PAM2, PAM4. Both 56 GBd

Equipment: Tek 80E11 sampler, DPO77004SX RT oscilloscope
Kalmar Optical impulser into Finisar O/E; Tek PPG into SHF mux, DAC
Optics vs Copper for In-chassis Connections @ 56-112 Gbps: is copper still a viable solution?

Scott McMorrow, R&D Consultant, Teraspeed Consulting, Division of Samtec
Questions?

• How fast can we push signals out of conventional FCBGA packaging?
• When do PCB materials run out of steam?
• How far can we push twin-ax cables?
• Doesn’t it cost more?
Novel Differential Serdes Package Design Cell

• Package trace to ball transitions no longer performance limit
• Transition BW > 50 GHz
• 56 G NRZ / 112 G PAM4
Bump-to-PCB Package Performance From Actual Design

-10 dB

-3 dB

40 GHz

50 GHz
Interconnect Loss per Inch (dB)
OIF LR (Long Reach)

Design Space for Low Cost PCB Materials

Design Space for Cable

High Cost Materials or Cable

High Cost PCB

Low Cost PCB

Megtron 6

ISpeed

FR408
OIF XSR (Extra Short Reach)
## -5 dB Reach

<table>
<thead>
<tr>
<th>Bandwidth\Material\Reach</th>
<th>FR408</th>
<th>MEGTRON 6</th>
<th>Micro-Twinax</th>
<th>Optics</th>
</tr>
</thead>
<tbody>
<tr>
<td>10G</td>
<td>&lt;10”</td>
<td>12”+</td>
<td>12”+</td>
<td>12”+</td>
</tr>
<tr>
<td>14G</td>
<td>&lt;5”</td>
<td>&lt;12”</td>
<td>12”+</td>
<td>12”+</td>
</tr>
<tr>
<td>28G</td>
<td>&lt;2”</td>
<td>&lt;5”</td>
<td>12”+</td>
<td>12”+</td>
</tr>
<tr>
<td>56G</td>
<td>0.0”</td>
<td>&lt;2”</td>
<td>12+”</td>
<td>12”+</td>
</tr>
<tr>
<td>112G</td>
<td>0.0”</td>
<td>0.0”</td>
<td>&lt;6”</td>
<td>12”+</td>
</tr>
</tbody>
</table>
Back to the Future

- 1976 Cray 1 delivered to Los Alamos National Labs
- The interconnect system was coax
- He was ahead of his time.
3.5 ps per Meter 6 Sigma

Eyespeed™ Cable

Ultra-low skew by Design
Design Possibilities
Channel Performance

Coax launch → Host cable to board connector → Twinax Cable → FQSFP → Host compliance board

- Insertion Loss
- FQSFP - 12 inch AWG30 - DCC SMT DIFFERENTIAL RETURN LOSS
- COMMON-MODE RETURN LOSS

<table>
<thead>
<tr>
<th>Pair</th>
<th>NEXT</th>
<th>PEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIMIT</td>
<td>1.80</td>
<td>4.60</td>
</tr>
<tr>
<td>Rx1</td>
<td>0.10</td>
<td>2.07</td>
</tr>
<tr>
<td>Rx2</td>
<td>0.13</td>
<td>2.31</td>
</tr>
<tr>
<td>Rx3</td>
<td>0.14</td>
<td>2.30</td>
</tr>
<tr>
<td>Rx4</td>
<td>0.13</td>
<td>2.01</td>
</tr>
<tr>
<td>Rx5</td>
<td>0.12</td>
<td>2.04</td>
</tr>
<tr>
<td>Rx6</td>
<td>0.11</td>
<td>2.33</td>
</tr>
<tr>
<td>Rx7</td>
<td>0.17</td>
<td>2.27</td>
</tr>
<tr>
<td>Rx8</td>
<td>0.14</td>
<td>2.08</td>
</tr>
</tbody>
</table>
System PCB Cost Reduction by Moving to Cable

PCB Cost per Square Inch
1K Units

Cost

$4.50
$4.00
$3.50
$3.00
$2.50
$2.00
$1.50
$1.00
$0.50
$0.00

Layer Count

2 12 22 32 42 52

FR4 Halogen Free w/FlyUnder
Low Loss Laminate
Thank you!

---

QUESTIONS?
Interconnects: Where We Came From
Where We Are Going To

Joel Goergen – Cisco Systems, Inc.
Version r1
DesignCon
January 2016 Santa Clara, CA
Overview

• What role does a SERDES play in the platform?
• What are the 5 basic copper reaches today?
• What are the power requirements?
• What defines the end of electrical reach?
• How much further can electrical go with SiP as a partner?
SERDES in the Platform

Past, Present, Future
High Speed Interconnects to the Fabric Platforms in Early Years

Mid 1993 - 1999
2 to 4 Channels per line card to the fabric
4000-13 or Rogers material
Back Plane / Mid Plane

1998 – 2006
10 to 20 Channels per line card to the fabric
N6000-21 or Rogers material
Back Plane / Mid Plane

2003 – 2010
20 to 40 Channels per line card to the fabric
Broad range of materials
Back Plane / Mid Plane

** Channel count is double for TX and RX
** C2C and C2M channels are in addition to the fabric channels

- Platforms range 5 to 8 years
- Major undertaking / Major Cost
- Takes advantage of new SERDES, Materials, and Connector technology
- Involves smaller steps in ASIC geometry

** Channel count is double for TX and RX
** C2C and C2M channels are in addition to the fabric channels
High Speed Interconnects to the Fabric Platforms Here and Now

2009 – 2015
40 to 200 Channels per line card to the fabric
Broad range of materials
Back Plane / Mid Plane / Orthogonal

- Platforms range 5 to 8 years
- Major undertaking / Major Cost
- Takes advantage of new SERDES, Materials, and Connector technology
- Involves smaller steps in ASIC geometry

2011 – 2018
200 to 400 Channels per line card to the fabric
Broad range of materials
Back Plane / Mid Plane / Orthogonal

** Channel count is double for TX and RX
** C2C and C2M channels are in addition to the fabric channels

- 6.25Gbps
- 10.325Gbps
- 15.7Gbps
- 28Gbps
- 10+Gbps
- 15+Gbps
- 28Gbps
- 56Gbps
High Speed Interconnects to the Fabric Platforms to Come

2016 – 2024
400 to 1000 Channels per line card to the fabric
Broad range of materials
Back Plane / Mid Plane / Orthogonal

• Platforms range 5 to 8 years
• Major undertaking / Major Cost
• Takes advantage of new SERDES, Materials, and Connector technology
• Involves smaller steps in ASIC geometry

** Channel count is double for TX and RX
** C2C and C2M channels are in addition to the fabric channels

2018 – 2026
800++ Channels per line card to the fabric
Broad range of materials
Back Plane / Mid Plane / Orthogonal

• Looks to be a problem area
• OIF and IEEE pursuing a 27dB loss line
Reach

Looking at the 5 basic reach definitions talked about most
Electrical Reach - Length, Loss & Applications
Up to 50 Gb/s Signaling Rates

<table>
<thead>
<tr>
<th>IL</th>
<th>USR</th>
<th>XSR</th>
<th>VSR</th>
<th>C2M</th>
<th>MR</th>
<th>LR</th>
<th>C2F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 10mm/0.4in</td>
<td>&lt; 50mm/2.0in</td>
<td>&lt; 200mm/7.9in</td>
<td>&lt; 500mm/19.7in</td>
<td>&lt; 1000mm/39.4in</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.5dB@14GHz</td>
<td>4dB@14GHz</td>
<td>10dB@14GHz</td>
<td>20dB@14GHz</td>
<td>20dB@14GHz</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3dB@28GHz</td>
<td>8dB@28GHz</td>
<td>20dB@28GHz</td>
<td>40dB@28GHz</td>
<td>40dB@28GHz</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35dB@14GHz</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bump-to-bump Inside MCM or 3D Stack</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball Across PCB</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ball-to-ball</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DesignCon 2016 – Santa Clara, CA
## Electrical Reach - Length, Loss & Applications Where We Seem to be Going

<table>
<thead>
<tr>
<th>Distance</th>
<th>Loss at 14GHz</th>
<th>Loss at 28GHz</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 10mm/0.4in</td>
<td>1.5dB@14GHz</td>
<td>&lt;3dB@28GHz</td>
<td>Bump-to-bump Inside MCM or 3D Stack</td>
</tr>
<tr>
<td>&lt; 50mm/2.0in</td>
<td>4dB@14GHz</td>
<td>&lt;8dB@28GHz</td>
<td>Ball-to-ball Across PCB</td>
</tr>
<tr>
<td>&lt; 200mm/7.9in</td>
<td>10dB@14GHz</td>
<td>&lt;20dB@28GHz</td>
<td>Ball-to-ball</td>
</tr>
<tr>
<td>&lt; 500mm/19.7in</td>
<td>20dB@14GHz</td>
<td>&lt;27.5dB@28GHz</td>
<td>Ball-to-ball</td>
</tr>
<tr>
<td>&lt; 1000mm/39.4in</td>
<td>&lt;27.5dB@14GHz</td>
<td>&lt;27.5dB@28GHz</td>
<td>Ball-to-ball</td>
</tr>
</tbody>
</table>

DesignCon 2016 – Santa Clara, CA
Back Plane / Chip-to-Fabric Loss

.3bj/25Gbps agreed to -35dB@12.89GHz accepted loss line.

Looks like OIF and IEEE may focus on -27.5dB@14GHz for 56Gbps PAM4 coding.

Doing so sets a new floor of -27.5dB and that may drive -27.5dB@28GHz for 112Gbps coding.

Staying above this limit line will result in re-timers every 20 inches for 56Gbps PAM4.

Staying above this limit line will result in re-timers every 10 to 15 inches for 112Gbps PAM4.

From IEEE802.3 P802d3bj_D3.pdf 100Gbps KR-4 Channel Loss
How to cover reach with the best power coverage of SERDES cores
Simple Math

- 25Gbps NRZ LR SERDES cores do not often require re-timer chips. The percentage of links is less then 15%. Since re-timers are almost the same power as the entire core, the total link power is 1.15X.

- 56Gbps PAM4 LR SERDES cores at the proposed limit line would require a re-timer on almost every link. The percentage is above 90%. With at least one re-timer per link, the total link power is really 2X.

- 112Gbps PAM4 LR SERDES cores at the proposed limit line would require two re-timers per link. The total power per link is now 3X.
Electrical Channel Reach Finally Ends ...
Finally Ends ... Now ... Why???

• Given two re-timers per link at 112Gbps, the required total power now exceeds what could be done with a common light source using optical links.

• In simple terms, unless the re-timer count is maintained at 1.15X, using an electrical link is no longer an effective bit path.

• For years, connector companies, materials companies, and board shops have had to cave into loss budgets. Chip companies have held us hostage. And those same chip companies are now pushing the limit lines once again, resulting in a fast transition to optical interconnects.
Integrated Photonics

The Only Hope
Unified Substrate
Core + SERDES + Optics

Need to isolate the SERDES from the die. Allows for the SERDES technology, the SiP technology, and the Core to progress at optimum geometries for those technology blocks.

Drive SERDES from the core die with parallel I/O.

SERDES I/O is now directly coupled to the SiP via a very short interface, requiring little power.
On Board Optics
Core + SERDES and Optics On Board

Need to place the optics within 10 inches of the core die. This will help to contain power, and maintain the limit lines comfortable to the chip companies.

SERDES core becomes more of a chip-to-module structure.

The Optics is now a component mounted on the circuit board.
Copper will hold its own up to 56Gbps using PAM4.

After 112Gbps, there may exist opportunities for very short reach applications. Everything else has to be optical. Unless the chip companies are willing to compromise on the channel loss limit.

To be fair – in July 2005, I said the copper limit was 25Gbps. Today, I think we can achieve 224Gbps, with a very few 400Gbps opportunities. But limit lines have to be maintained at -35dB.
Thank you!