Redefining Computing Architecture Boundaries with Off Package Chiplets
An Energy Centric Computing Perspective

Allan Cantle - 6/16/2022
What’s driving Computing Architecture
Power & Cost Efficient Domain Specific Architectures, DSA

Sustainability
- Energy Recovery
- Energy Efficiency
- Performance Increase

Open Architecture

COST

DSA Composability
- Fixed Architecture
- Hard Physical Composability
- Dynamic Software Composability
- Pluggable Physical Composability
What’s driving Computing Architecture
Power & Cost Efficient Domain Specific Architectures, DSA

Fixed Architecture
Hard Physical Composability
DSA Composability
Dynamic Software Composability
Pluggable Physical Composability
What is the Motivation for Chiplets?

Efficiently Combining Compute & Data in Pursuit of Performance

- Computer Architecture is JUST Compute & Data!
- & Circumventing the Von-Neumann Bottleneck
- Improving Data to Compute Movement Efficiency
  - Moores Law.....shrinking everything
  - Bringing Memory on chip - Caching
  - Heterogeneous Processors
  - Moving everything closer together
    - Chiplets in a package.......2D, 2.5D & 3D
But what about Off Package IO?
System size rarely shrinks with higher density packaging!

• From ~0.5pJ/bit chiplet IO Energy

• To ~10pJ/bit System IO Energy
  • A 20x off package increase!

• But there is the Promise of Co-Packaged Optics

• Ultimately we want Highest System Performance at the Lowest Energy & Cost

• Lets take a closer look at the data
## Popular PHY IO Standards

<table>
<thead>
<tr>
<th>PHY Name</th>
<th>Substrate Type</th>
<th>Ch. Speed Gb/s</th>
<th>Tbits/s/mm Beachfront Tbits/s</th>
<th>Reach mm</th>
<th>Energy pJ/bit</th>
<th>Possible Chiplet IO</th>
<th>Possible Off Package IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>✤ Traditional QSFPxx Optical IO</td>
<td>Organic</td>
<td>112</td>
<td>Same as PHY</td>
<td>100,000’s</td>
<td>15</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>* PCIe-G5 Reference</td>
<td>Organic</td>
<td>32</td>
<td>0.22</td>
<td>973</td>
<td>7</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>▲ DDR Memory Ch. Reference</td>
<td>Organic</td>
<td>6.4</td>
<td>0.06</td>
<td>300</td>
<td>5</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Silicon Photonics</td>
<td>Any</td>
<td>Tb/s</td>
<td>Same as PHY</td>
<td>100,000’s</td>
<td>1</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>* OIF - VSR</td>
<td>Organic</td>
<td>40</td>
<td>1.1</td>
<td>160</td>
<td>1.5</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>* BoW Fast</td>
<td>Organic</td>
<td>16</td>
<td>1.0</td>
<td>25</td>
<td>&lt;0.5</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>† UCle - Standard</td>
<td>Organic</td>
<td>32</td>
<td>1.8</td>
<td>&lt;25</td>
<td>0.5</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>* OpenHBI</td>
<td>Silicon Interposer</td>
<td>8</td>
<td>3.4</td>
<td>8</td>
<td>0.5</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>† UCle - Advanced</td>
<td>Silicon Interposer</td>
<td>32</td>
<td>10.5</td>
<td>&lt;2</td>
<td>0.25</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

Potential Off Package Chiplet PHY

* ODSA PHY Spreadsheet  † UCle Whitepaper  

[Image] MCHPC’19 - Prospects for Memory - J. Thomas Pawlowski - Slide 29
## On and Off Package Memory Configurations

### Traditional Local Memory + RDMA Example

<table>
<thead>
<tr>
<th>Example</th>
<th>80GB</th>
<th>3TB</th>
<th>8TB</th>
<th>16TB+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bandwidth TBytes/s</td>
<td>Power Watts</td>
<td>Bandwidth TBytes/s</td>
<td>Power Watts</td>
</tr>
<tr>
<td>Traditional Local Memory + RDMA Example</td>
<td>0.6</td>
<td>49</td>
<td>0.6</td>
<td>49</td>
</tr>
</tbody>
</table>

Data Movement Energy shown in Red - Assume NIC & Ethernet Switch Power insignificant compared to their IO power.
# On and Off Package Memory Configurations

## CXL Shared Memory Example

<table>
<thead>
<tr>
<th>Example</th>
<th>80GB</th>
<th>3TB</th>
<th>8TB</th>
<th>16TB+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bandwidth</td>
<td>Power</td>
<td>Bandwidth</td>
<td>Power</td>
</tr>
<tr>
<td>Traditional Local Memory + RDMA Example</td>
<td>0.6 TBytes/s</td>
<td>49 Watts</td>
<td>0.6 TBytes/s</td>
<td>49 Watts</td>
</tr>
<tr>
<td>CXL Shared Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>848</td>
</tr>
</tbody>
</table>

Data Movement Energy shown in Red - Assume Retimer & CXL Switch Power insignificant compared to their IO power
# On and Off Package Memory Configurations

**CXL Shared Memory Example with Co-Packaged Optics**

Data Movement Energy shown in Red - Assume CXL Switch Power insignificant compared to its IO power

<table>
<thead>
<tr>
<th>Example</th>
<th>80GB</th>
<th>3TB</th>
<th>8TB</th>
<th>16TB+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bandwidth TBytes/s</td>
<td>Power Watts</td>
<td>Bandwidth TBytes/s</td>
<td>Power Watts</td>
</tr>
<tr>
<td>Traditional Local Memory + RDMA Example</td>
<td>0.6</td>
<td>49</td>
<td>0.6</td>
<td>49</td>
</tr>
<tr>
<td>CXL Shared Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>848</td>
</tr>
<tr>
<td>SiP CXL Shared Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>448</td>
</tr>
</tbody>
</table>
# On and Off Package Memory Configurations

CXL Shareable, Local Memory using OIF-VSR Example

---

**Example** | **80GB** | **3TB** | **8TB** | **16TB+**
--- | --- | --- | --- | ---
| **Bandwidth TBytes/s** | **Power Watts** | **Bandwidth TBytes/s** | **Power Watts** | **Bandwidth TBytes/s** | **Power Watts** | **Bandwidth TBytes/s** | **Power Watts**

**Traditional Local Memory + RDMA Example**

- 0.6 | 49
- 0.6 | 49
- 0.6 | 373+
- 0.6 | 373+

**CXL Shared Memory Example**

- 3 | 24
- 2 | 848
- 2 | 848
- 2 | 848

**SiP CXL Shared Memory Example**

- 3 | 24
- 2 | 448
- 2 | 448
- 2 | 448

**CXL Shareable, Local Memory Example**

- 3 | 24
- 2 | 208
- 2 | 208
- 2 | 656

---

Data Movement Energy shown in Red - Assume CXL Shared Memory Buffer & CXL Switch Power insignificant compared to their IO power

---

NALLASWAY
### On and Off Package Memory Configurations

**Optical, CXL Shareable, Local Memory using OIF-VSR Example**

Data Movement Energy shown in Red - Assume CXL Shared Memory Buffer & CXL Switch Power insignificant compared to their IO power.

<table>
<thead>
<tr>
<th>Example</th>
<th>80GB</th>
<th>3TB</th>
<th>8TB</th>
<th>16TB+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bandwidth TBytes/s</td>
<td>Power Watts</td>
<td>Bandwidth TBytes/s</td>
<td>Power Watts</td>
</tr>
<tr>
<td>Traditional Local Memory + RDMA Example</td>
<td>0.6</td>
<td>49</td>
<td>0.6</td>
<td>49</td>
</tr>
<tr>
<td>CXL Shared Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>848</td>
</tr>
<tr>
<td>SiP CXL Shared Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>448</td>
</tr>
<tr>
<td>CXL Shareable, Local Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>208</td>
</tr>
<tr>
<td>SiP CXL Shareable, Local Memory Example</td>
<td>3</td>
<td>24</td>
<td>2</td>
<td>208</td>
</tr>
</tbody>
</table>

 Champetized CPU

32x CXLx8 Shared Memory Channels

SiP CXL Shared Memory Example

CXL Shared Memory Example

CXL Shareable, Local Memory Example

SiP CXL Shareable, Local Memory Example
Watts/TByte/s Memory Sharing Comparison
Energy Advantage of Off-Package Chiplet Shared Memory Buffers

Passive Copper Interconnect

Silicon Photonic Interconnect

Lower is Better

Watts per TByte/s

Memory Capacity

80GB 3TB 8TB 16TB+

Energy Advantage of Off-Package Chiplet Shared Memory Buffers

Watts per TByte/s

Memory Capacity

80GB 3TB 8TB 16TB+

Lower is Better
Can UCIe Standard or BoW support HBM?
BoW & UCIe-Standard for HBM with Organic packages

- HBM = 11mm beachfront + roadmap to 1.1TB/s
- Requires 1.1TB/s * 8 / 11mm = 0.8Tb/s/mm
- UCIe-Standard & BoW can support HBM using organic substrates
Proposed Chiplet Industry Standardization
Accessible to Tier 2 and 3 Industry Players

• Utilize BoW & OIF-VSR PHYs for On & Off Package Chiplet IO’s respectively
• Leverage Cost effective Organic Packaging
• Unifying Off Package IO for ALL Processors WILL Democratize Compute
Proposed Chiplet Industry Standardization
Accessible to Tier 2 and 3 Industry Players

- Off-Package Chiplets are simply IO Translators!
- For differing IO Use Cases
- OCP HPCM - 3D System Level Packaging
Summary

• Today’s Disaggregation Models are taking Energy Efficiency in the wrong direction
  • Insight through Energy Centric Compute Analysis
• Support of an off-package Chiplet IO Standard will allow:
  • Graceful Power and latency increase
• This Presentation ONLY assessed Off-Package Chiplets from an Energy Saving Perspective. Several other significant benefits accrue:
  • Low Latency to Large Capacities at High Bandwidths
  • Flexible, Modular and Composable
  • Lower cost and improved manufacturability
  • Simpler, Sustainable and Competitive Ecosystem Business Model