System Composability Using CXL

Kurtis Bowman, Marketing Working Group Co-Chair, CXL Consortium
Mitch Wright, CXL System Architect, Liqid
Data Center: Expanding Scope of CXL

CXL 3.0/3.1
Composable Fabric growth for disaggregation/pooling/accelerator

CXL 2.0
Memory/Accelerator Pooling with Single Logical Devices

Multiple Nodes inside a Rack/Chassis supporting pooling of resources

Compute Express Link® and CXL® are registered trademarks of the Compute Express Link Consortium.
## CXL Specification Feature Summary

<table>
<thead>
<tr>
<th>Features</th>
<th>CXL 1.0 / 1.1</th>
<th>CXL 2.0</th>
<th>CXL 3.0</th>
<th>CXL 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Release date</td>
<td>2019</td>
<td>2020</td>
<td>August 2022</td>
<td>November 2023</td>
</tr>
<tr>
<td>Max link rate</td>
<td>32GTs</td>
<td>32GTs</td>
<td>64GTs</td>
<td>64GTs</td>
</tr>
<tr>
<td>Flit 68 byte (up to 32 GTs)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Flit 256 byte (up to 64 GTs)</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Type 1, Type 2 and Type 3 Devices</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Memory Pooling w/ MLDs</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Global Persistent Flush</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CXL IDE</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Switching (Single-level)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Switching (Multi-level)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Direct memory access for peer-to-peer</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Enhanced coherency (256 byte flit)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Memory sharing (256 byte flit)</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Multiple Type 1/Type 2 devices per root port</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Fabric capabilities (256 byte flit)</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Fabric Manager API definition for PBR Switch</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Host-to-Host communication with Global Integrated Memory (GIM) concept</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Trusted-Execution-Environment (TEE) Security Protocol</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Memory expander enhancements (up to 34-bit of meta data, RAS capability enhancements)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Compute Express Link® and CXL® are registered trademarks of the Compute Express Link Consortium
CXL 3.1 Fabric Management Features
CXL 3.1 Fabric Management Features

- Fabric Decode/Routing requirements
- Host-to-Host communication with Global Integrated Memory (GIM) concept
- Direct P2P .mem support through PBR Switches
  - Adds symmetric Link Layer definition
  - Enables direct caching of CXL.mem for an accelerator, which is not possible with Unordered IO (UIO)
- Fabric Manager API definition for PBR Switch
CXL 3.1: Fabric Management Features

Port-Based Routing (PBR) compared to Hierarchy-Based Routing (HBR)

Support *fabric* topologies other than *tree* topologies that HBR switches offer
- Address-based, non-prescriptive routing for large memory fabrics
- Supports tree, mesh, ring, star, butterfly, and multi-dimensional topologies
Host to Fabric-Attached Memory (FAM) communication with Global Integrated Memory (GIM)

- Multiple Hosts mapping CXL Fabric-Attached Memory devices
- Hosts and FAM devices can initiate cross-domain accesses to GIM

Cross-domain access

CXL Fabric

Compute Express Link® and CXL® are registered trademarks of the Compute Express Link Consortium.
CXL 3.1: Fabric Management Features

Direct peer-to-peer (P2P) mem support for Accelerators through PBR Switches

Direct P2P CXL.mem for Accelerators

- Enables accelerator access to peer Type-3 memory
- The accelerator and Type-3 device must each be directly connected to an Edge Downstream Port (DSP)
- Utilizes port-based routing (PBR) for transactions
CXL 3.1: Fabric Management Features

Fabric Manager (FM) API definition for PBR Switch
CXL 3.1 Fabric Management Features

• CXL Fabric Management Integration with Industry Standards
  • DMTF Redfish®
    • Redfish Support for CXL
    • CXL to Redfish Mapping Specification
  • SNIA Swordfish®
    • SNIA Swordfish Specification
    • What is the SNIA Swordfish Standard
  • OFA Sunfish™
    • OFA Sunfish

Compute Express Link® and CXL® are registered trademarks of the Compute Express Link Consortium.
Thank You

www.ComputeExpressLink.org