Introducing the CXL 3.1 Specification

Presented by
Rob Blankenship and Mahesh Wagh
CXL Board of Directors

255+ Member Companies

Industry Open Standard for High Speed Communications
Industry Landscape

- Proliferation of Cloud Computing
- Growth of AI & Analytics
- Cloudification of the Network & Edge
CXL Ecosystem

Growth of CXL ecosystem since its inception

- Ecosystem that meets the ever-increasing performance and scale requirements
- Fully backwards compatible
- Lower overall system cost
- Comprehensive compliance and testing support
Data Center: Expanding Scope of CXL

CXL 3.0/3.1

Composable Fabric growth for disaggregation/pooling/accelerator

CXL 2.0

Multiple Nodes inside a Rack/Chassis supporting pooling of resources

Memory/Accelerator Pooling with Single Logical Devices

Memory Pooling with Multiple Logical Devices

CXL Switch

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium.
Representative CXL Usages

Caching Devices / Accelerators

- **TYPE 1**
  - **CXL**
    - PROTOCOLS: CXL.io, CXL.cache, CXL.mem (for direct P2P)
    - USAGES: PGAS NIC, NIC atomics

- **Accelerator**
  - **Nic**
    - **Cache**

- **Processor**

Accelerators with Memory

- **TYPE 2**
  - **CXL**
    - PROTOCOLS: CXL.io, CXL.cache, CXL.mem
    - USAGES: GP GPU, Dense computation

- **Accelerator**
  - **Cache**

- **Memory Controller**

Memory Buffers

- **TYPE 3**
  - **CXL**
    - PROTOCOLS: CXL.io, CXL.mem
    - USAGES: Memory BW expansion, Memory capacity expansion, Storage class memory

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium.
CXL Specification Release Timeline

- March 2019: CXL 1.0 Specification Released
- September 2019: CXL Consortium Officially Incorporates CXL 1.1 Specification Released
- November 2020: CXL 2.0 Specification Released
- August 2022: CXL 3.0 Specification Released
- November 2023: CXL 3.1 Specification Released
# 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>
</tr>
</thead>
<tbody>
<tr>
<td>Release date</td>
<td>2019</td>
<td>2020</td>
<td>August 2022</td>
</tr>
<tr>
<td>Max link rate</td>
<td>32GTs</td>
<td>32GTs</td>
<td>64GTs</td>
</tr>
<tr>
<td>Flit 68 byte (up to 32 GTs)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Flit 256 byte (up to 64 GTs)</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>
</tr>
<tr>
<td>Memory Pooling w/ MLDs</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Global Persistent Flush</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>CXL IDE</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Switching (Single-level)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Switching (Multi-level)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct memory access for peer-to-peer</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Enhanced coherency (256 byte flit)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Memory sharing (256 byte flit)</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>
</tr>
<tr>
<td>Fabric capabilities (256 byte flit)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium.
RECAP: CXL 3.0: POOLING & SHARING

1 Expanded use case showing memory sharing and pooling
2 CXL Fabric Manager is available to setup, deploy, and modify the environment

CXL Switch(es)
Standardized CXL Fabric Manager

H1
S1 Copy
S2 Copy

H2
S2 Copy
S3 Copy

H3
S1 Copy

H4
S1 Copy
S3 Copy

D1
S2

D2

D3

D4
S1

D#
S3

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium.
CXL 3.1
Feature enhancements
The CXL specification continues to evolve to meet new usage models

New features introduced in the CXL 3.1 specification:

- **CXL Fabric Improvements/Extensions**
  - Scale-out of CXL fabrics using PBR (Port Based Routing)
- **Trusted-Execution-Environment Security Protocol (TSP)**
  - Allows for Virtualization-based Trusted Execution Environments (TEEs) to host Confidential Computing Workloads
- **Memory Expander Improvements**
  - Up to 32-bit of meta data and RAS capability enhancements
CXL Fabric
Improvements/
Extensions
CXL Fabric Improvements/Extensions

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

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
CXL 3.1: Fabric Enhancement 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
- Target memory device can be dedicated or shared with the host.

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

Fabric Manager (FM) API definition for PBR Switch

Diagram:
- Host
- CXL
- CXL PBR Switch
- Fabric Manager (FM)
- Type-1, 2, 3
Trusted Execution Environment (TEE) & TEE Security Protocol (TSP)
RECAP: CXL 2.0 Security Benefits

CXL 2.0 provides Integrity and Data Encryption (IDE) of traffic across all entities (Root Complex, Switch, Device)

CPU/SoC Root Complex
- Host Memory
- MC
- Home Agent
- Coherent Bridge
- IO Bridge

CXL Device
- MC
- DTLB
- Coherent Cache (Optional)
- Device Memory

Area of Protection
- CXL.memory
- CXL.cache
- CXL.io

CXL 2.0 Switch
CXL 3.1 Trusted Security Protocol (TSP)

Allows for Virtualization-based, Trusted Execution Environments (TEEs) to host Confidential Computing Workloads (CC WL)

**Key Capabilities:**
- Separation between TVM* & CSP’s infrastructure (VMM)
- Configuration of CXL device
- Encryption of sensitive data in both Host/Device memory
- Cryptographically verify correct configuration of trusted computing environment

**Benefits:**
- Freedom to migrate sensitive WLs to TSP-enabled Clouds
- Collaboration with multiple parties for sharing data
- Conform to Compliance & Data sovereignty programs
- Strengthen Application security & Software IP protection

*TVM = Trusted VM

---

Compute Express Link® and CXL® are trademarks of the Compute Express Link Consortium.
### TSP Components for Confidential Computing

- **Trusted Execution State & Access Control**
  - How access to memory is controlled
- **Configuration**
  - Ability to determine the **supported** security features on the device, enable required features, and **lock** the configuration
- **Attestation & Authentication**
  - Trusting who you are talking to
- **Memory Encryption**
  - Encrypting **data-at-rest**
- **Transport Security**
  - Encrypting the link to protect **data-in-flight** and detect/prevent physical attacks

---

#### Confidential Computing (initiator)

<table>
<thead>
<tr>
<th>Trusted Execution State &amp; Access Control</th>
<th>Configuration</th>
<th>Attestation Authentication</th>
<th>Memory Encryption (Data-at-rest)</th>
<th>Transport Security (Data-in-flight)</th>
</tr>
</thead>
</table>

**CXL Trusted Execution Environment Security Protocol (TSP)**

**CXL memory expander (target)**
- HDM-H (CXL 3.1), HDM-DB (Future)
Memory Expander Improvements
The CXL 3.1 specification provides for up to 32-bits of meta data per cacheline. CXL 3.1 supports 2-bits of meta-data for cache coherency that are used to describe if a cache line is shared, exclusive, modified, or invalid and up to 32-bits for additional meta-data.

The extended meta data may be used with Host coherent and Device coherent memory. Extended meta data allows more information about the line to be available. Possible use cases:

- Access control, data type tagging, and memory-tiering algorithms
- Supports proposed DDR6 feature that adds 16-32 bits of meta data for each 64B of data

Improved visibility into CXL memory device errors. Additional information on errors such as correctable error limits. New information on the source of the errors and the transaction occurring during the error condition.

Added control over memory device RAS. Memory-sparing, DDR5 error-check-scrubbing, Media testing, Patrol-scrub, Capacity or performance degradation.

Direct peer-to-peer CXL memory access for accelerators.

Enhancements to CXL-Attached Memory
# 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 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 32-bit of meta data, RAS capability enhancements)</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
CXL 3.1 Summary

• The CXL specification continues to evolve to meet the usage models

• New features introduced in the CXL 3.1 specification:
  • CXL Fabric Improvements/Extensions
    • Scale-out of CXL fabrics using PBR (Port Based Routing)
  • Trusted-Execution-Environment Security Protocol (TSP)
    • Allows for Virtualization-based, Trusted Execution Environments (TEEs) to host Confidential Computing Workloads
  • Memory Expander Improvements
    • Up to 32-bit of meta-data and RAS capability enhancements

www.ComputeExpressLink.org
Q&A

Please share your questions in the Question Box
Thank You

www.ComputeExpressLink.org