

CXL (Compute Express Link) Technology

Santhosh Nagaraj Nag

Intel, San Jose, California, USA

Email: Santhoshnag@gmail.com

How to cite this paper: Nag, S.N. (2023)
CXL (Compute Express Link) Technolog,
Journal of Computer and Communications,
11, 94-102.

<https://doi.org/10.4236/jcc.2023.116006>

Received: May 19, 2023

Accepted: June 27, 2023

Published: June 30, 2023

Copyright © 2023 by author(s) and
Scientific Research Publishing Inc.

This work is licensed under the Creative
Commons Attribution International
License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

CXL (Compute Express Link) technology is a relatively new high-speed interconnect standard that was developed to enable faster communication between CPUs, GPUs, and other high-performance components in data center systems. This paper aims to provide a comprehensive technical overview of CXL technology, including its features, advantages, and potential applications in the modern data center environment. CXL Technology Research: CXL technology is based on Peripheral Component Interconnect Express (PCIe) and its extensions. CXL 1.0 is a switch-based interconnect architecture that operates on PCIe Gen5 electrical signaling, achieving data speeds of up to 32 Giga transfers per second (GT/s) per lane. CXL technology provides hardware-based support for cache coherency and memory semantics. CXL technology architecture consists of three main components: 1) CXL Devices: Devices that are compatible with the CXL interface can include processors, accelerators such as Graphics Processing Units (GPUs), and Smart Storage Devices; 2) CXL Switch: The switch enables communication between devices that support CXL. The switch can be external or embedded, allowing for more complex topologies; 3) CXL Memory: CXL memory devices support the CXL protocol for the efficient sharing of System memory.

Keywords

CXL, Interconnect, Protocol, Memory

1. Introduction

The increasing demand for high-performance computing has led to the development of new technologies that can deliver higher bandwidth and lower latency for data-intensive workloads. CXL technology [1]. is one such technology that was developed specifically to address the challenges of modern workloads in the data center. CXL is an open standard developed by a consortium of leading companies, including Intel, Google, and Cisco, among others, to define a high-

speed, low-latency interconnect protocol that can support a wide range of applications and workloads.

2. CXL Overview

CXL is a high-speed interconnect standard that enables communication between CPUs, GPUs, and other high-performance components in data center systems. CXL is based on the PCIe 5.0 physical layer, which provides high bandwidth and low latency for data transfers. CXL also includes several features that are designed to support a wide range of applications and workloads, including memory expansion [2]. Coherent memory access, and cache coherence [3].

Real-time Questions and Answers:

Will storage devices such as SSDs require “native” CXL controller support, or can they use the PCIe interface?

Since SSDs are block devices (and not random-access load store devices) they don't need CXL controllers. PCIe will continue to be used in the same way as today.

Does CXL address memory interleaving administrative controls or is that the choice of CXL device vendors?

Memory interleaving choices and implementations are outside the scope of the CXL specification.

How will applications deal with different latencies for different memory types?

Most likely, applications will not be aware of what memory they are using and therefore the different latencies. The OS/kernel will have the responsibility to allocate the correct memory type to an application.

How do you see applications evolving with CXL being like a far NUMA-like node?

Applications will generally not be latency aware. In theory, it is possible to create a malloc function that can specify whether it can use a higher latency memory pool and the operating system services it accordingly. You could apply existing NUMA-like approaches with CXL as well.

How are atomics supported over CXL?

Since CXL memory is cache-coherent this should be the same as CPU/direct-attached memory.

Is PCI Express 5.0 technology the “transport” for all things CXL? Or will there be CXL connectivity between devices that do not require PCI Express?

In CXL 1.1 the transport is PCIe 5.0. Other transports might be developed/specified in the future.

Can we expect to see CXL 1.1 memory expanders using non-volatile memory, or do we have to wait for CXL 2.0?

CXL 1.1 can support memory expansion devices but might need special software/driver support for persistence, RAS [4]. And other features.

What are your thoughts on the first adoption of CXL—first only directly attached memory in a system or in pools of memory used by many systems?

CXL 1.1 does not support pooled memory.

Are there any CXL memory expansion devices in development and if so, when do you expect to see servers being built with the new topology, of course requiring CPUs with CXL IOs as well?

Vendors are working on memory expansion solutions. For specific product plans and roadmaps, you'll need to talk directly with member companies.

Will CXL eventually replace DDR due to higher bandwidth per pin?

CXL latency addition is a concern, but it is certainly possible that CXL could replace DDR in the future.

Is 3-way interleaving only supported at the proprietary cross-host-bridge level? Or can it be supported at each of the levels (for example the CXL Host-bridge, USP, Device)?

CXL didn't want to affect switches and have them implement an awkward 3-way operation that adds latency and complexity. Therefore, limited the 3-way math to only the host in the host propriety logic and the device. There is no support for 3-way interleaving at the USP.

What effect does no RCRB requirement have on RCEC requirements?

RCEC is required if the device is exposed as an RCIEP (Root Complex Integrated End Point). On slide 11, that aspect has not changed. The device on the right-hand side needs to be exposed as an RCIEP, and the link is still not visible to the legacy software because the downstream port has the same register layout as the CXL 1.1 specification. Therefore, you still need the host to implement the RCEC and the errors will always go with the RCEC. There is no difference in that behavior.

Does the CXL 2.0 ECN support OS (Linux) for Integrity and Data Encryption (IDE)?

Not aware of any code in the upstream right now that enables IDE. However, we are expecting that it will be available soon as we start seeing devices and designs take advantage of IDE.

3. CXL Features

CXL includes several key features that make it an attractive technology for high-performance computing in the data center.

First, CXL provides high bandwidth and low latency for data transfers, which is critical for data-intensive workloads such as AI, machine learning, and big data analytics.

Second, CXL enables memory expansion, which allows systems to scale up memory capacity without adding additional memory channels.

Third, CXL supports coherent memory access, which enables multiple processors to access shared memory in a coherent manner. Finally, CXL provides cache coherence, which ensures that all processors have a consistent view of the data in the cache.

Forth, Memory Semantics

CXL technology supports memory semantics, enabling devices to share mem-

ory through the CXL interface.

4. Advantages of CXL

CXL offers several advantages over other interconnect technologies in the data center.

First, CXL provides significantly higher bandwidth and lower latency than other interconnect technologies, such as Ethernet or InfiniBand;

Second, CXL enables memory expansion, which can significantly improve the performance of memory-intensive workloads;

Third, CXL supports coherent memory access, which enables multiple processors to access shared memory in a coherent manner without the need for additional hardware;

Finally, CXL provides cache coherence, which ensures that all processors have a consistent view of the data in the cache, which can reduce the likelihood of data inconsistencies and errors.

5. Applications of CXL

CXL has the potential to be used in a wide range of applications, and new application scenarios with significant performance benefits including

1) High-Performance Computing Clusters

By providing interconnectivity between many processing elements, CXL technology is ideal for High-Performance Computing (HPC) clusters and their applications such as computational fluid dynamics and financial modeling.

2) Artificial Intelligence

Artificial Intelligence applications require massive amounts of data and computation. CXL technology's low latency and High Bandwidth capabilities make it an ideal interface for faster processing of AI applications.

3) Smart Storage

CXL Technology can connect Smart Storage Devices directly to a Processor, allowing for faster data access and storage.

6. CXL Interconnect

CXL builds upon the physical and electrical interfaces of PCIe with protocols that establish coherency, simplify the software stack, and maintain compatibility with existing standards.

Specifically, CXL leverages a PCIe 5 feature that allows alternate protocols to use the physical PCIe layer.

- CXL is helping data centers more efficiently handle the yottabytes of data generated by artificial intelligence (AI) and machine learning (ML) applications.
- We discuss how CXL technology maintains memory coherency between the CPU memory space and memory on attached devices to enable resource sharing (or pooling).

- We also detail how CXL builds upon the physical and electrical interfaces of PCI Express (PCIe) with protocols that establish coherency, simplify the software stack, and maintain compatibility with existing standards.

To continue to advance the performance, servers are moving increasingly to a heterogeneous computing architecture with purpose-built accelerators offloading specialized workloads from CPUs. The memory cache coherency of CXL allows for sharing of memory resources between CPUs and accelerators.

Further, CXL enables the deployment of new memory tiers that can bridge the latency gap between main memory and SSD storage. These new memory tiers will add bandwidth, capacity, increased efficiency, and lower total cost of ownership (TCO). With these many benefits, the industry is decisively converging on CXL as a cache-coherent interconnect for processors, memory, and accelerators.

CXL is an open standard industry-supported cache-coherent interconnect for processors, memory expansion, and accelerators. Essentially, CXL technology maintains memory coherency between the CPU memory space and memory on attached devices. This enables resource sharing (or pooling) for higher performance, reduces software stack complexity, and lowers overall system cost. The CXL Consortium has identified three primary classes of devices [5]. That will employ the new interconnect:

- **Type 1 Devices:** Accelerators such as smart NICs typically lack local memory. Via CXL, these devices can communicate with the host processor's DDR memory.
- **Type 2 Devices:** GPUs, ASICs, and FPGAs are all equipped with DDR or HBM memory and can use CXL to make the host processor's memory locally available to the accelerator—and the accelerator's memory locally available to the CPU. They have also co-located in the same cache coherent domain and help boost heterogeneous workloads.
- **Type 3 Devices:** Memory devices can be attached via CXL to provide additional bandwidth and capacity to host processors. The type of memory is independent of the host's main memory.

7. CXL Consortium

The CXL Consortium is an open industry standard group formed to develop technical specifications that facilitate breakthrough performance for emerging usage models while supporting an open ecosystem for data center accelerators and other high-speed enhancements.

8. CXL Protocols & Standards

The CXL standard supports a variety of use cases via three protocols: CXL.io, CXL.cache, and CXL.memory.

- **CXL.io:** This protocol is functionally equivalent to the PCIe protocol—and utilizes the broad industry adoption and familiarity of PCIe. As the founda-

tional communication protocol, CXL.io is versatile and addresses a wide range of use cases.

- **CXL.cache:** This protocol, which is designed for more specific applications, enables accelerators to efficiently access and cache host memory for optimized performance.
- **CXL.memory:** This protocol enables a host, such as a processor, to access device-attached memory using load/store commands.

Together, these three protocols facilitate the coherent sharing of memory resources between computing devices, e.g., a CPU host and an AI accelerator. Essentially, this simplifies programming by enabling communication through shared memory. The protocols used to interconnect devices and hosts are as follows:

- Type 1 Devices: CXL.io + CXL.cache
- Type 2 Devices: CXL.io + CXL.cache + CXL.memory
- Type 3 Devices: CXL.io + CXL.memory

9. Compute Express Link vs PCIe

CXL builds upon the physical and electrical interfaces of PCIe with protocols that establish coherency, simplify the software stack, and maintain compatibility with existing standards. Specifically, CXL leverages a PCIe 5 feature that allows alternate protocols to use the physical PCIe layer. When a CXL-enabled accelerator is plugged into a x16 slot, the device negotiates with the host processor's port at default PCI Express 1.0 transfer rates of 2.5 giga transfers per second (GT/s). CXL transaction protocols are activated only if both sides support CXL. Otherwise, they operate as PCIe devices.

CXL 1.1 and 2.0 use the PCIe 5.0 physical layer, allowing data transfers at 32 GT/s, or up to 64 gigabytes per second (GB/s) in each direction over a 16-lane link.

CXL 3.0 uses the PCIe 6.0 physical layer to scale data transfers to 64 GT/s supporting up to 128 GB/s bi-directional communication over a x16 link.

10. CXL Features and Benefits

Streamlining and improving low-latency connectivity and memory coherency significantly bolsters computing performance and efficiency while lowering TCO. Moreover, CXL memory expansion capabilities enable additional capacity and bandwidth above and beyond the direct-attach DIMM slots in today's servers. CXL makes it possible to add more memory to a CPU host processor through a CXL-attached device. When paired with persistent memory, the low-latency CXL link allows the CPU host to use this additional memory in conjunction with DRAM memory. The performance of high-capacity workloads depends on large memory capacities such as AI. Considering that these are the types of workloads most businesses and data-center operators are investing in, the advantages of CXL are clear. (**Figure 1**)

New in CXL 2.0 and 3.0

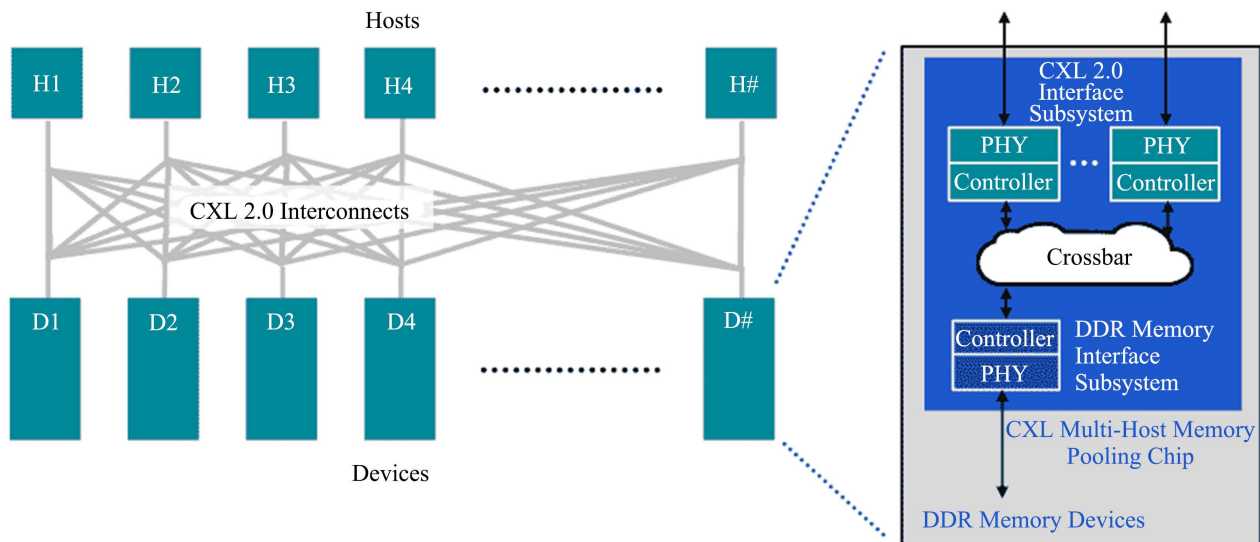


Figure 1. CXL memory pooling through direct connection.

11. Memory Pooling

CXL 2.0 supports switching to enable memory pooling [6]. With a CXL 2.0 switch, a host can access one or more devices from the pool. Although the *hosts* must be CXL 2.0-enabled to leverage this capability, the memory *devices* can be a mix of CXL 1.0, 1.1, and 2.0-enabled hardware. At 1.0/1.1, a device is limited to behaving as a single logical device accessible by only one host at a time. However, a 2.0-level device can be partitioned as multiple logical devices, allowing up to 16 hosts to simultaneously access different portions of the memory.

As an example, host 1 (H1) can use half the memory in device 1 (D1) and a quarter of the memory in device 2 (D2) to finely match the memory requirements of its workload to the available capacity in the memory pool. The remaining capacity in devices D1 and D2 can be used by one or more of the other hosts up to a maximum of 16. Devices D3 and D4, CXL 1.0 and 1.1-enabled respectively, can be used by only one host at a time.

CXL 3.0 introduces peer-to-peer direct memory access and enhancements to memory pooling where multiple hosts can coherently share a memory space on a CXL 3.0 device. These features enable new use models and increased flexibility in data center architectures.

12. Switching

By moving to a CXL 2.0 direct-connect architecture [7]. Data centers can achieve the performance benefits of main memory expansion—and the efficiency and total cost of ownership (TCO) benefits of pooled memory. Assuming all hosts and devices are CXL 2.0-enabled, “switching” is incorporated into the memory devices via a crossbar in the CXL memory pooling chip. This keeps latency low but requires a more powerful chip since it is now responsible for the control plane functionality performed by the switch. With low-latency direct connec-

tions, attached memory devices can employ DDR DRAM to provide expansion of host's main memory. This can be done on a very flexible basis, as a host is able to access all—or portions of—the capacity of as many devices as needed to tackle a specific workload.

CXL 3.0 introduces multi-tiered switching which enables the implementation of switch fabrics. CXL 2.0 enabled a single layer of switching. With CXL 3.0, switch fabrics are enabled, where switches can connect to other switches, vastly increasing the scaling possibilities.

13. The “as Needed” Memory Paradigm

Analogous to ridesharing, CXL 2.0 and 3.0 allocate memory to hosts on an “as needed” basis, thereby delivering greater utilization and efficiency of memory. This architecture provides the option to provision server main memory for nominal workloads (rather than worst case), with the ability to access the pool when needed for high-capacity workloads and offering further benefits for TCO. Ultimately, the CXL memory pooling models can support the fundamental shift to server disaggregation and composability. In this paradigm, discrete units of computing, memory and storage can be composed on-demand to efficiently meet the needs of any workload.

14. Integrity and Data Encryption (IDE)

Disaggregation—or separating the components of server architectures— increases the attack surface. This is precisely why CXL includes a secure-by-design approach. Specifically, all three CXL protocols are secured via Integrity and Data Encryption (IDE) which provides confidentiality, integrity, and replay protection. IDE is implemented in hardware-level secure protocol engines instantiated in the CXL host and device chips to meet the high-speed data rate requirements of CXL without introducing additional latency. It should be noted that CXL chips and systems themselves require safeguards against tampering and cyberattack. A hardware root of trust implemented in the CXL chips can provide this basis for security and support requirements for secure boot and secure firmware download

15. Scaling Signaling to 64 GT/s

CXL 3.0 brings a step function increase in a data rate of the standard. As mentioned earlier, CXL 1.1 and 2.0 use the PCIe 5.0 electricals for their physical layer: NRZ signaling at 32 GT/s. CXL 3.0 keeps that same philosophy of building on broadly adopted PCIe technology and extends it to the latest 6.0 version of the PCIe standard released in early 2022. That boosts CXL 3.0 data rates to 64 GT/s using PAM4 signaling. We cover the details of PAM4 signaling in PCIe 6.

16. Conclusions

CXL is a once-in-a-decade technological force that will transform data center

architectures. Supported by a who's who of industry players including hyper scalers, system OEMs, platform and module makers, chip makers [8]. And IP providers, its rapid development reflects the tremendous value it can deliver.

CXL technology is a promising new interconnect standard that has the potential to improve the performance and scalability of high-performance computing applications in the data center. CXL enables high-bandwidth, low-latency communication between CPUs, GPUs, and other high-performance components, and supports a wide range of applications and workloads. As the adoption of CXL grows, we can expect to see significant improvements in the performance and efficiency of data center systems.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Sharma, D.D. (2019) Compute Express Link. Compute Express Link Consortium, White Paper.
https://docs.wixstatic.com/ugd/0c1418_d9878707bbb7427786b70c3c91d5fbd1.pdf
- [2] Sharma, D.D. (2022) An Open Industry-Standard Interconnect Enabling Heterogeneous Data-Centric Computing. 2022 *IEEE Symposium on High-Performance Interconnects (HOTI)*, California, 17-19 August 2022, 5-12.
<https://ieeexplore.ieee.org/abstract/document/9912551>
- [3] Van Doren, S. (2019) Abstract—HOTI 2019. 2019 *IEEE Symposium on High-Performance Interconnects (HOTI)*, Santa Clara, 14-16 August 2019, 18.
<https://ieeexplore.ieee.org/abstract/document/9070313>
- [4] Academia.edu. https://www.academia.edu/45413748/Compute_Express_Link_RAS
- [5] VIP Experts. Verification Central, “An Introduction to the CXL Device Types” (2020).
<https://blogs.synopsys.com/vip-central/2020/03/04/an-introduction-to-the-cxl-device-types/>
- [6] Li, H., *et al.* (2022) First-Generation Memory Disaggregation for Cloud Platforms.
<https://arxiv.org/abs/2203.00241>
- [7] Compute Express Link: The Breakthrough CPU-to-Device Interconnect (2020).
<https://www.computeexpresslink.org/about-cxl>
- [8] <https://www.marketsandmarkets.com/Market-Reports/compute-express-link-cxl-market-259303213.html>