

Introduction To CXL 2.0 Memory

Article

What is CXL?

Compute eXpress Link (CXL), is a protocol which runs over PCIe without affecting normal PCIe functionality. CXL allows expansion of memory in non-DIMM form factors, for example, DRAM or PMem packaged as EDSFF E1.s or E3.s drives. Add-in-Cards (AICs) can be inserted in standard PCIe slots, and can hold standard DIMMs. Note that, at time of publication, AICs are not supported on Lenovo servers.

The images below show examples of an add-in card and an E3.S memory expansion module from SMART Modular Technologies.

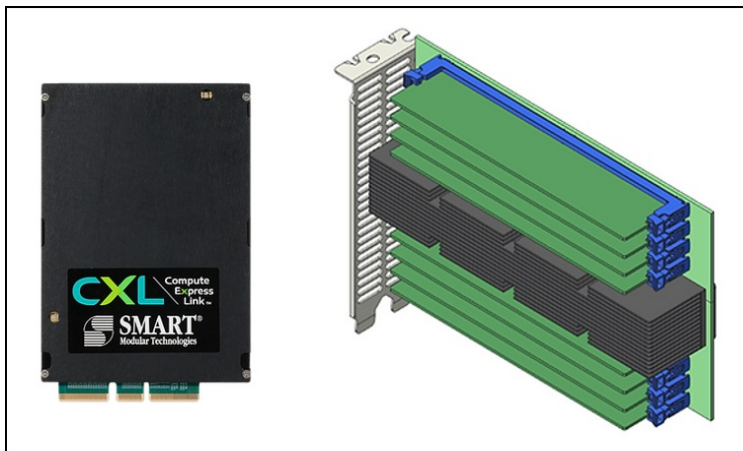


Figure 1. SMART Modular Technologies E3.S CXL Memory Module and 8-DIMM Add-In-Card (images used with permission)

Why CXL?

Three memory challenges hinder performance and increase total cost of ownership in data centers:

- The first of these is the limitation of current server memory subsystems. DRAM and Solid-State Drive (SSD) storage have a massive latency difference – around 1000x. For example, the Micron 96 GB DDR5-4800 MT/s DIMM has a write latency of 15.44 ns (nanoseconds) and the Micron 7450 NVMe SSD has a write latency of 15 us (microseconds). If a processor runs out of capacity in DRAM and starts paging to SSD, the CPU spends a great deal of time waiting for data, with a significant decrease in compute performance.
- Secondly, the number of channels in memory controllers is not keeping pace with increased core counts in multi-core processors. Processor cores can be starved for memory bandwidth, minimizing the benefit of higher core counts. For example, Intel Xeon 6 processors can have up to 288 cores served by 12 memory channels.

- Finally, with accelerators having their own memory, there is a problem of underutilized or isolated memory resources.

The solution to these memory challenges is a flexible technology that provides more memory bandwidth and capacity to processors. Compute Express Link (CXL) was designed to provide low-latency, cache-coherent links between processors, accelerators, and memory.

CXL is a once-in-a-decade technology that will transform data center architectures. It is supported by a variety of industry players including system OEMs, platform and module makers, software developers, and chip makers. Its rapid development and adoption is a testimony to its tremendous value.

Lenovo ThinkSystem V4 servers with Intel Xeon 6 processors support CXL 2.0 memory. It is expected that future AMD EPYC processors will also support CXL memory.

CXL device classes

CXL technology maintains coherency between the CPU memory space and the 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 lists three classes (not the CXL revision number, but a device type) of devices that use the new interconnect. Devices use two or more of the CXL protocols – CXL.io, CXL.mem, and CXL.cache – to communicate.

Type 1 Devices: Accelerators without local memory – these can communicate with the host processor DRAM, shown here as DDR5/MRDIMM, via CXL.

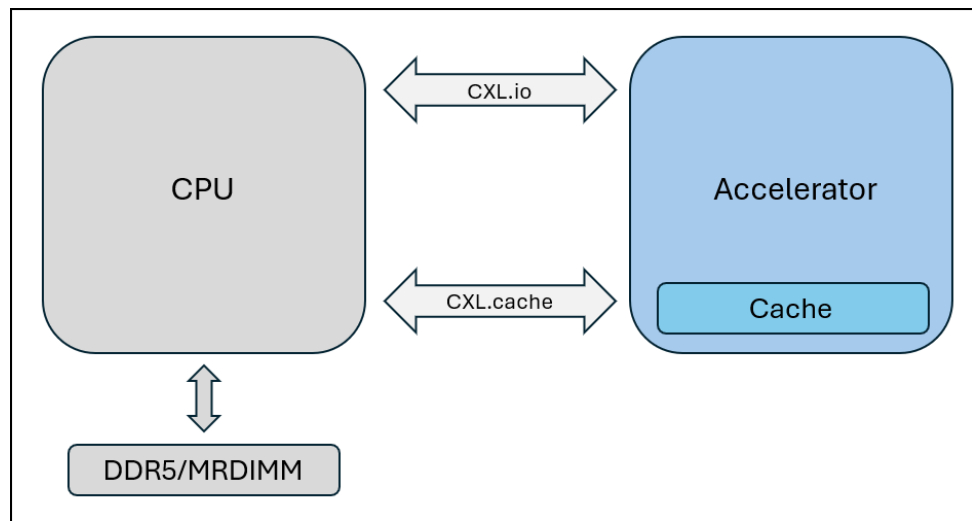


Figure 2. Type 1 device

Type 2 Devices: GPUs, ASICs, and FPGAs equipped with DDR or HBM memory (high-bandwidth memory) can use CXL to make host processor memory locally available to the accelerator, and the accelerator local memory available to the host CPU.

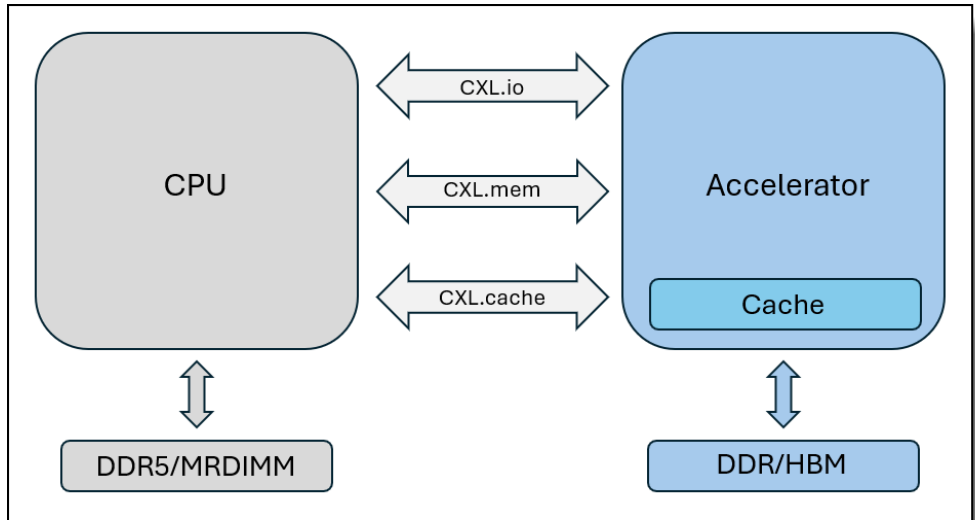


Figure 3. Type 2 device

Type 3 Devices: Devices that contain memory which is independent of the host main memory can be attached via CXL to provide additional memory bandwidth and capacity to host processors.

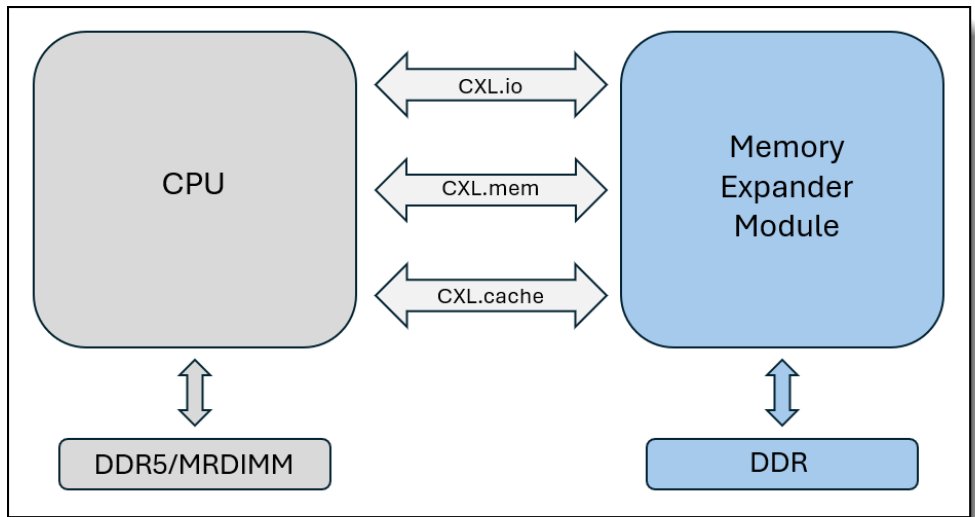


Figure 4. Type 3 device

Supported CXL protocols

The CXL standard supports three protocols: the first of which is largely equivalent to the PCIe protocol, the second of which allows accelerators to access and cache host memory, and the third of which allows a host to access device-attached memory.

These protocols enable the cache-coherent sharing of memory between devices, e.g., a CPU host and an AI accelerator. This simplifies device programming by allowing communication through shared memory.

PCIe basis of CXL

CXL builds upon the physical and electrical interfaces of PCIe, and leverages a feature that allows alternate protocols to use the physical PCIe layer. CXL transaction protocols are activated only if both connected devices support CXL, otherwise they operate as standard PCIe devices.

CXL 1.1 and 2.0 use the PCIe 5.0 physical layer, and transfer data at up to 32 GT/s, or up to 64 GB/s, in each direction over a 16-lane link. DDR5-5600, by comparison, has a bandwidth of 69.21 GB/s – very close to CXL 2.0 speed.

CXL 2.0 differences

CXL 2.0 expands on CXL 1.1 and introduces new features, among which are switching, device partitioning, memory pooling, and EDSFF support. We'll look at these in turn.

CXL 2.0 Switching

CXL 2.0 supports switching to enable memory pooling and access to the pool by multiple hosts. With a CXL 2.0 switch, a host can access one or more devices from the pool. Hosts must be CXL 2.0-enabled to leverage this capability, but memory devices can be a mix of CXL 1.0, 1.1, and 2.0 hardware.

Switching is incorporated into memory devices via the CXL memory controller chip.

CXL 2.0 Device partitioning

A CXL 1.0/1.1 memory device behaves as a single logical device and can be accessed by only one host at a time, as shown in the diagram below.

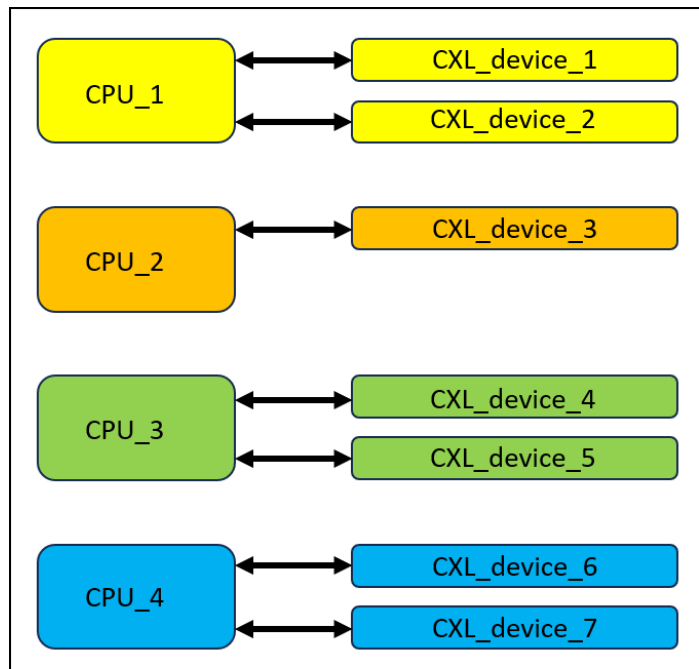


Figure 5. CXL 1.0/1.1 device access

Memory on a CXL 2.0 device can be partitioned into multiple logical devices, and up to 16 hosts can simultaneously access different portions of the memory space. Devices can use DDR5 DRAM to provide expansion of host main memory.

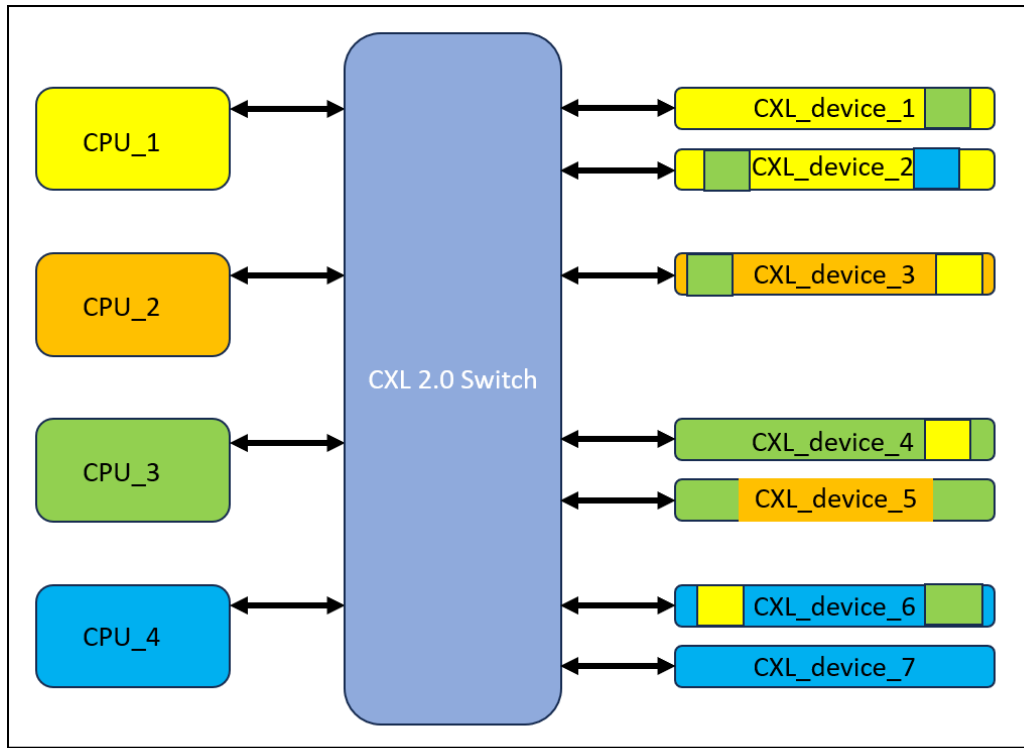


Figure 6. CXL 2.0 device access

As an example, the four CPUs in the diagram above can use a memory device exclusively, as is the case with CPU_4 and CXL_device_7, especially if the device is CXL 1.0 or CXL 1.1. CXL 2.0 devices, such as CXL_device_1 through CXL_device_6, can be shared by all 4 hosts. Hosts can match the memory requirements of workloads to the available capacity in the memory pool.

CXL 2.0 Memory pooling

CXL 2.0 pools all memory, and allocates it to hosts as required, achieving greater memory utilization and efficiency.

CXL 2.0 EDSFF support

CXL 2.0 supports CXL devices in the EDSFF format, meaning that memory can be installed in server disk slots. This improves the versatility of the devices, and also helps customers make better use of available resources like disk slots. Figure 1 shows a CXL expansion module in EDSFF E3.S format.

Customer benefits

Customers who adopt CXL 2.0 will benefit in a number of ways:

- **Memory Expansion**

Host main memory can be accessed in nanoseconds but is limited in size due to memory slot count and cost. Disk access times are around 1000x longer - a massive performance gap. This gap can be filled by CXL memory, which has around 20x to 50x the latency of main memory. The performance of high-capacity workloads such as AI depends on large memory size. Since these are the workload classes most businesses and data-center owners are currently investing in, the advantages of CXL are clear.

- **Memory Pooling**

Memory pooling allows devices such as host CPUs and accelerators like GPUs to share memory. This can make messaging simpler and faster, leading to improved performance. In addition, CXL memory can be partitioned into multiple logical memory spaces, each of which can be assigned to a host or device, or shared among them. More efficient use of memory resources, and decreased TCO, will be the result.

- **Potential for Persistent Memory**

The now-discontinued Intel Persistent Memory (PMem) showed that there is a benefit to having fast, non-volatile memory in certain environments. CXL offers the promise of non-volatile memory which is not attached to the memory bus. The advantage is that it does not limit the amount of main memory attached to the bus, or affect the performance of the bus. This can help improve TCO for customers.

What's next: CXL 3.1

The CXL 3.1 specification was released in November 2023. Devices that meet new specifications typically only become widely available two to three years after the specifications are released.

CXL 3.1 introduces peer-to-peer direct memory access and enhancements to memory pooling, as well as multi-tiered switching which enables the implementation of switch fabrics. The data rate is doubled to the PCIe 6.1 standard of 64 GT/s. Reconfiguring memory without a reboot enhances flexibility.

Summary

CXL 2.0 has many advantages for implementers. It can increase the size and bandwidth of attached memory, and does so without affecting the memory bus. Pooling and partitioning make memory allocation more efficient, while the potential for data persistence should not be ignored.

Author

André Rossouw is a Technical Course Developer in the Lenovo Education team. Since graduating from Pretoria Technikon (now Tshwane University of Technology) in Pretoria, South Africa, André has held engineering and education positions in a number of companies, including EMC. André has been with Lenovo since August 2017, and is based in Morrisville NC, USA.

Related product families

Product families related to this document are the following:

- [Memory](#)

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
8001 Development Drive
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

© Copyright Lenovo 2025. All rights reserved.

This document, LP2146, was created or updated on February 3, 2025.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
<https://lenovopress.lenovo.com/LP2146>
- Send your comments in an e-mail to:
comments@lenovopress.com

This document is available online at <https://lenovopress.lenovo.com/LP2146>.

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at <https://www.lenovo.com/us/en/legal/copytrade/>.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Lenovo®

ThinkSystem®

The following terms are trademarks of other companies:

AMD and AMD EPYC™ are trademarks of Advanced Micro Devices, Inc.

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries.

Other company, product, or service names may be trademarks or service marks of others.