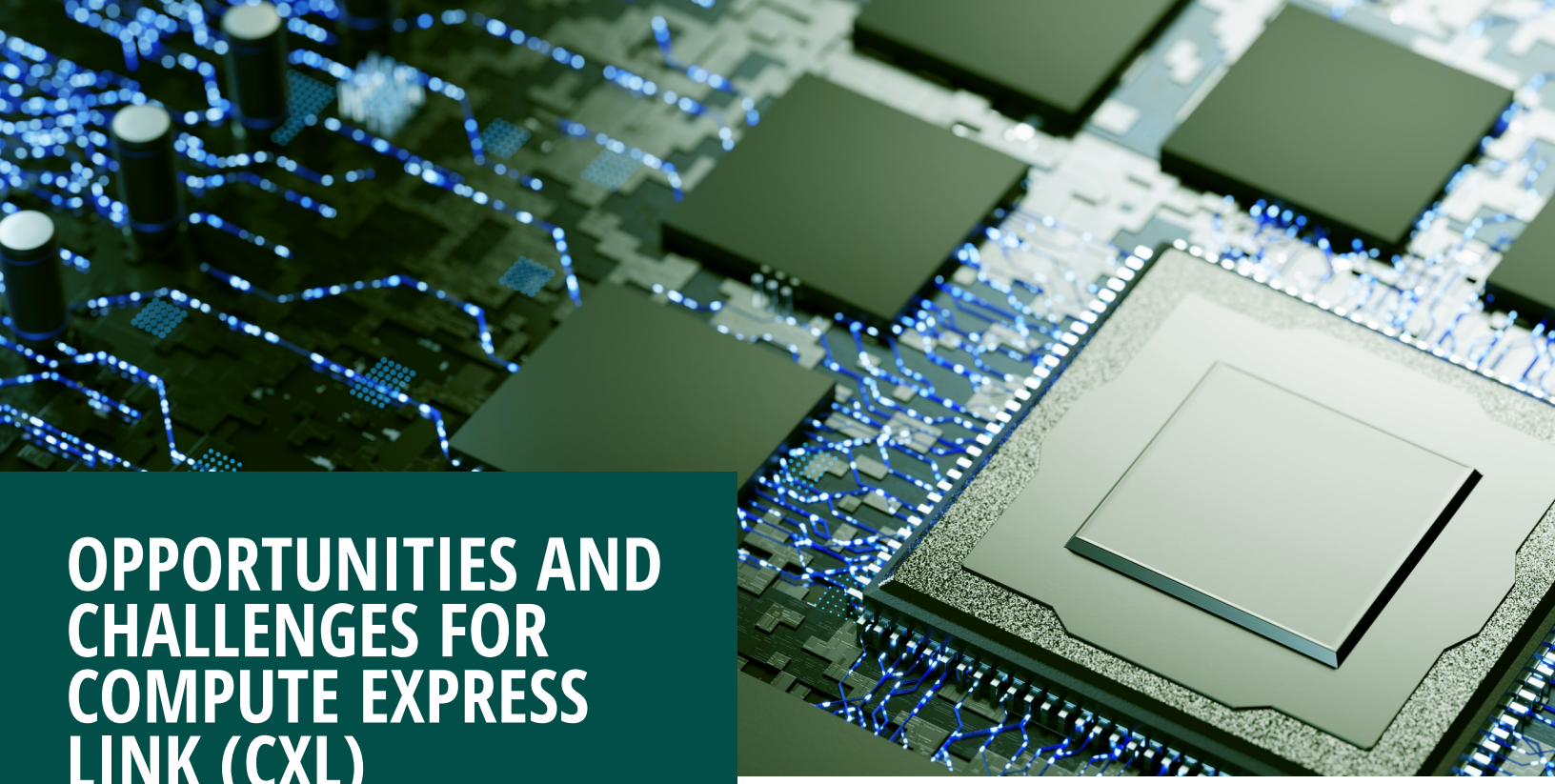


OPPORTUNITIES AND CHALLENGES FOR COMPUTE EXPRESS LINK (CXL)

Reece Hayden, Principal Analyst
Paul Schell, Industry Analyst

ABIresearch THE TECH INTELLIGENCE EXPERTS™

CXL Compute Express Link

CONTENTS

EXECUTIVE SUMMARY

- Data centers are facing a technical and commercial challenge resulting from the widening performance gap between compute and memory.
- Compute Express Link (CXL) offers an economical cache-coherent interconnect solution to solve memory bottlenecks and stranded memory for a range of memory-intensive and memory-elastic workloads.
- CXL enhances both Generative Artificial Intelligence (Gen AI) and non-Gen AI/ Machine Learning (ML) workloads by providing expanded memory access, crucial for tasks like caching inference of Large Language Models (LLMs).
- The CXL market remains nascent, but broad support by "market makers" and Total Cost of Ownership (TCO) considerations in data centers are driving adoption of the standard among processor, accelerator (e.g., FPGA), and memory vendors. More than 20 CXL 1.1/2.0 devices are on the official compliance list with more in the pipeline.
- While widespread commercial deployments are ongoing, the industry is actively addressing technical and commercial challenges, leading the way to new usages and much broader CXL adoption, expected by 2027.
- Sustained momentum in software optimization, ecosystem alignment, and market education will be key to achieving the full potential of a novel technology like CXL.

INTRODUCTION

An arising technical and economic challenge for data centers is the widening gap between the performance of processors and the scalability of memory. This is especially relevant given the rapid growth in memory-intensive AI/ML workloads, which are driving demand for cost-effective computing systems with consistent and coherent access to memory. Maintaining efficient, consistent, and coherent memory access across multiple processing units is vital. Meanwhile, the cost of memory is taking an increasing share of data center expenditures. CXL provides an industry-standard, cache-coherent interconnect for processors, accelerators, and memory, addressing the challenge of scaling memory bandwidth and capacity for data-intensive applications in a cost-effective way.

Although this market remains nascent, CXL is strategically positioned to address future memory challenges driven by AI advancements in data centers (e.g., with near-memory accelerators and heterogeneous computing), potentially revolutionizing architectures, enhancing performance, increasing energy efficiency, and, crucially, lowering costs. Nonetheless, CXL has yet to reach widespread commercial availability or deployment, primarily because wide-spread deployment in data center requires extensive testing cycles. Software optimization, ecosystem alignment, and market education are three key areas gaining momentum, as is typical with a new technology like CXL, supported by an ecosystem that hopes to reap significant benefits from this promising technology.

CXL OVERVIEW

WHAT IS CXL?

CXL is an open industry standard for high-speed, high-capacity, and efficient interconnect from Central Processing Unit (CPU)-to-device and CPU-to-memory connections. The first iteration of this standard was released in 2019 and is now in its third generation, with specification 3.1 released in November 2023. CXL is mostly targeting data centers, but will also impact the enterprise server market, where different board and chassis designs are offered by the same server vendors. This interconnect aims to tackle the memory access bottleneck that is increasingly limiting the rate at which devices can retrieve instructions and data from the system's memory. CXL solves this by expanding the available memory, increasing bandwidth through the Peripheral Component Interconnect Express (PCIe) physical layer, and reducing system memory swap occurrences between Double Data Rate (DDR) and Solid-State Drive (SSD for workload processing, which improves overall memory transfer latency and, therefore, application performance. This is achieved through a series of protocols that enable coherent, inter-device memory:

- **CXL.io:** A foundational communication protocol functioning the same as PCIe 5.0 (PCIe 6.0 after CXL 3.0) and essential to CXL. It is used for link initialization, device discovery, and connections to devices. CXL.io is a required protocol to manage and operate CXL devices, providing the backbone for the operation of other CXL protocols.
- **CXL.cache:** Provides coherent access from the attached devices to the processor's memory. This enables CXL hosts to coherently share their memory with devices, reducing the need for less performant software-managed coherency.

- **CXL.mem:** Allows a host (and devices) to access device-attached memory coherently. It is a transactional protocol enabling an interface with any type of memory, e.g., DDR, Optane, or High-Bandwidth Memory (HBM). This supports sharing and pooling, dynamic allocation of memory based on workload needs, and improved overall system efficiency, which enhances utilization. This is especially important for supporting heterogeneous (and AI) workloads running across several processing architectures.

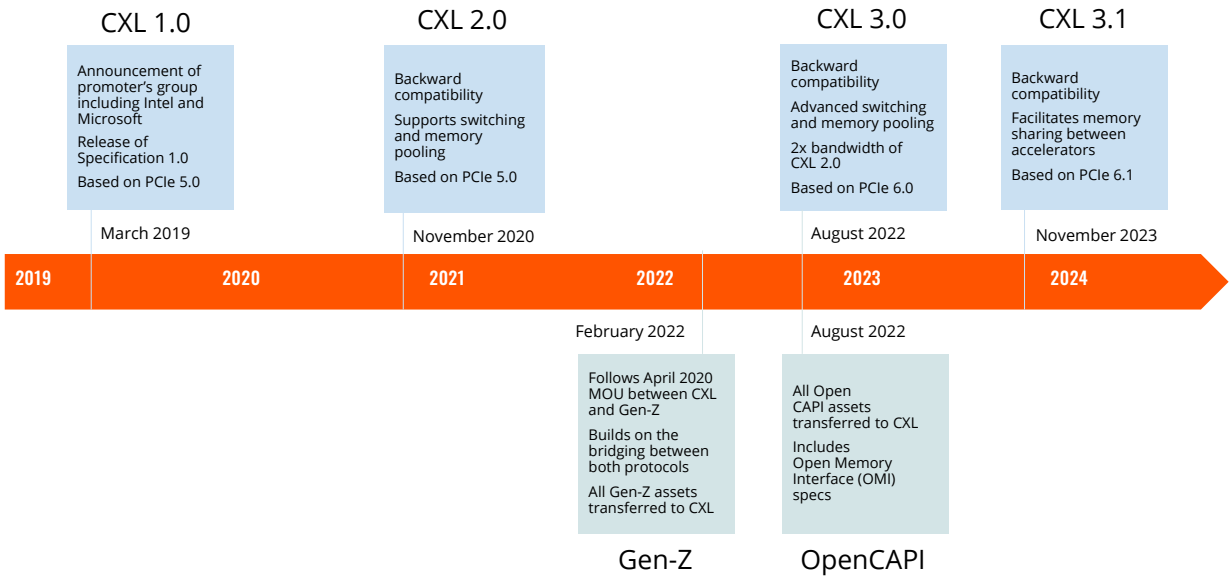
The protocols above support three broad types of CXL devices:

- **Type 1:** Accelerators (SmartNICs) without attached memory (using CXL.io + CXL.cache). As these devices commonly lack local memory, by using CXL, they can communicate with the host processor's DDR memory. Their task is to offload and accelerate specific computational tasks from the main CPU, mainly related to memory management.
- **Type 2:** Accelerators that include Graphics Processing Units (GPUs), Application-Specific Integrated Circuits (ASICs), and FPGAs with DDR or HBM using locally available host memory (CXL.io + CXL.cache + CXL.mem). The accelerator's memory is also made locally available to the CPU.
- **Type 3:** Devices with their own memory, aiming to connect to host processors (using CXL.io and CXL.mem). These devices are designed to provide memory expansion capabilities, i.e., they act as memory devices that can be accessed by the CPU via PCIe, while using CXL protocols for memory operations.

Since CXL's inception in 2019, it has been developed quickly with different specifications providing support for various use cases. Figure 1 shows how CXL has progressed, including the transfer of competing and overlapping standards Gen Z and Open CAPI to CXL. Key recent developments consist of CXL 3.0 and 3.1 alignment with PCIe 6.0 and 6.1, respectively, which bring sufficient bandwidth to enable more memory-intensive AI/ML applications. In fact, CXL 3.1, the specification released to the public in November 2023, uses PCIe 6.1 to support up to 128 Gigabits per Second (Gbps) bi-directional data transfer.

Figure 1: Timeline of CXL Ratification and Significant Expansion

(Source: ABI Research)



CXL does not have a direct competitor, as many open and proprietary solutions that preceded it (e.g., OpenCAPI, Gen-Z, and CCIX) were incorporated into the CXL specification by consortium members as it has grown.

CXL FEATURES

CXL's rapid development since its incorporation in 2019 is demonstrated by Table 1, which shows the features included in each specification and enables significant memory-centric use cases.

Table 1: CXL Features

(Source: ABI Research)

CXL SPECIFICATION	FEATURE	EXPLANATION
CXL 1.0/1.1 – March/June 2019 (based on PCIe 5.0)	Memory sharing	Allows multiple devices to access common memory.
	Memory expansion	Enables devices to directly access and utilize host memory. This expands memory for processors, making access to the high-latency SSD almost redundant.
	Security	Allows proprietary security measures as defined by devices and hosts.
CXL 2.0 – November 2020 (based on PCIe 5.0)	Switching	Enables multiple hosts and devices to connect to each other and share memory between them.
	Memory pooling	Aggregating memory resources from multiple devices into a shared memory pool that can be dynamically allocated. Overcoming memory stranding by communicating excess memory capacity through the CXL interconnect.
	Device integrity	Protocols to ensure secure functioning of devices connected through the CXL interface.
	Security	Integrity and Data Encryption (IDE) protocols: hardware-based, end-to-end encryption that ensures data are secured as they travel between processor and connected devices. Enables encryption on the CXL link for all three foundational protocols.
CXL 3.0 – August 2022 (based on PCIe 6.0)	Multi-level switching	Use of multiple CXL switches in multi-tiered configuration to connect numerous devices and hosts within the system. This enables more scalable processing systems.
	Peer-to-peer direct memory access	Enables devices connected to a CXL fabric to directly communicate and transfer data between each other without involving the CPU host. This reduces latency and frees up CPU resources. This alleviates memory stranding challenges.
	Fabric-attached memory	Enables connectivity between memory modules creating a fabric that enables shared and flexible memory resources across processors, accelerators, and other devices.
CXL 3.1 – November 2023 (based on PCIe 6.1)	Memory sharing	Facilitates memory sharing between accelerators without going through a host, which improves the interoperability of heterogeneous memory and computing systems. Inter-host communication is enabled using Global Integrated Memory (GIM)
	Security	Trusted Execution Environment (TEE) enables confidential compute use cases to be directly attached to CXL memory expander devices.
	Fabric capability	The fabric enhancement enables a port-based routed fabric capability with up to 4,000 nodes on the fabric.

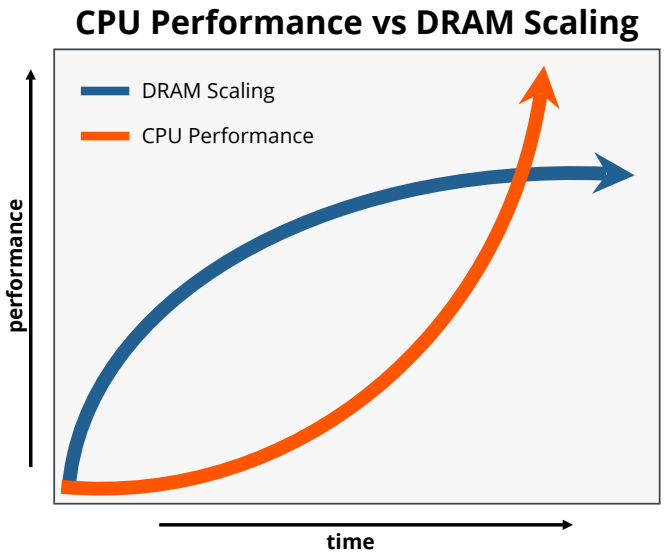
A noteworthy advantage of CXL 3.0's features is the enablement of composable server infrastructure. This relies on a pool of virtualized resources that can be intelligently shared and provisioned between devices to support real-time resource needs. CXL 3.0 enables server disaggregation and composability by coherently sharing memory across resources. Composable server infrastructure brings significant benefits: improved resource utilization by connecting isolated resource pools across any server; dynamic resource allocation to optimize application performance, which reduces necessary spending on resource overprovisioning; and the introduction of "Memory-as-a-Service."

KEY MARKET DRIVERS

Data centers face mounting challenges as workload demands surge due to the increasing adoption of AI. In this environment, memory has become a critical bottleneck, with Dynamic Random Access Memory (DRAM) costs now representing a significant portion of a server's Bill of Materials (BOM). For CPUs, an alternative to accessing the attached DRAM involves switching to SSDs; however, a substantial latency gap exists between DRAM and SSDs, impacting application performance and TCO. This disparity leads to underutilization of compute resources, further exacerbating efficiency challenges in data centers. In addition, with every new generation of CPUs, compute cores and performance are outpacing DRAM density growth, creating memory access bottlenecks and a gap between processor performance and memory bandwidth. This issue is exacerbated by the rapid advancements in processing capabilities driven by parallel and matrix processing architectures like accelerators and GPUs, with performance progress that has outpaced DRAM development. Figure 2 shows the divergent scaling characteristics between processor performance and DRAM memory bandwidth.

Figure 2: Understanding the Growing Gap between Processor Performance and Memory Bandwidth

(Source: ABI Research)



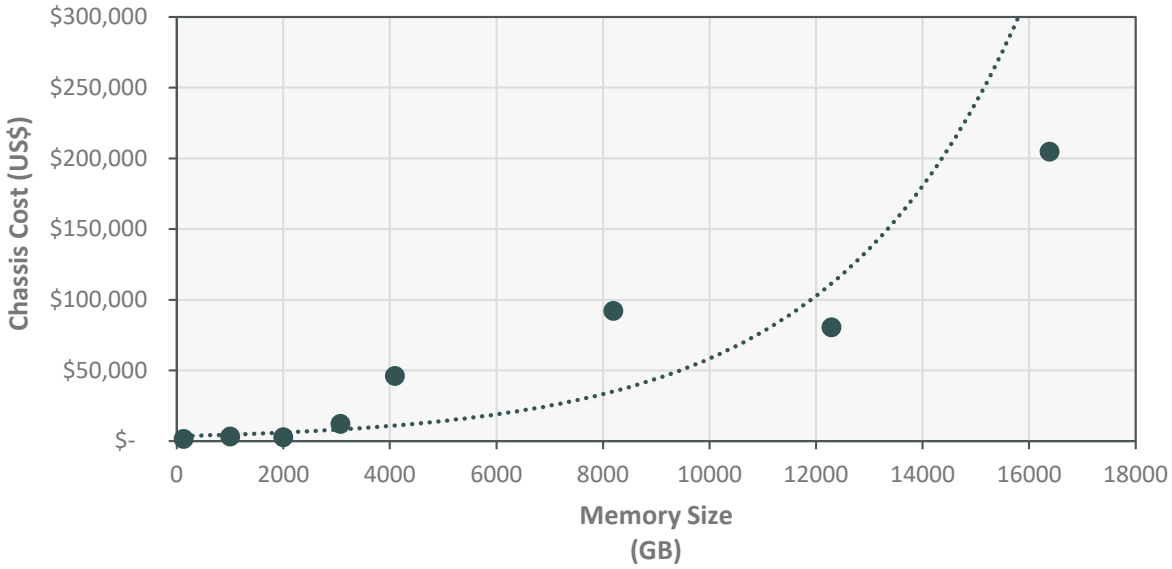
Bridging memory to computing engines and optimizing memory management for peak performance in a cost-effective way creates substantial challenges for the computing market. The following analysis further explores the key factors driving demand for a cost-effective unified interface standard, aiming to address memory barriers that impede data center efficiency across critical applications.

- **Adding More Memory Brings Exponential Increases in Server TCO:** As the computing ecosystem grapples with a widening performance gap between processors and memory, increasing memory capacity drives up total system costs,

primarily due to rapidly escalating chassis prices. The performance gap manifests in several ways, including processing speeds outpacing memory access times, limited memory bandwidth constraining data flow, and increased latency in data retrieval from memory (resulting from a more hierarchical memory structure, limitations to the memory controller in its management of channels, and/or more frequent and longer refresh cycles). Chart 1 illustrates the exponential increase in chassis prices as additional memory is incorporated, highlighting the economic implications of this apparent solution. This cost growth results from the increased manufacturing cost of chassis with higher memory density.

Chart 1: Chassis Cost Grows Quickly as You Add Memory

(Sources: ABI Research; MemVerge)

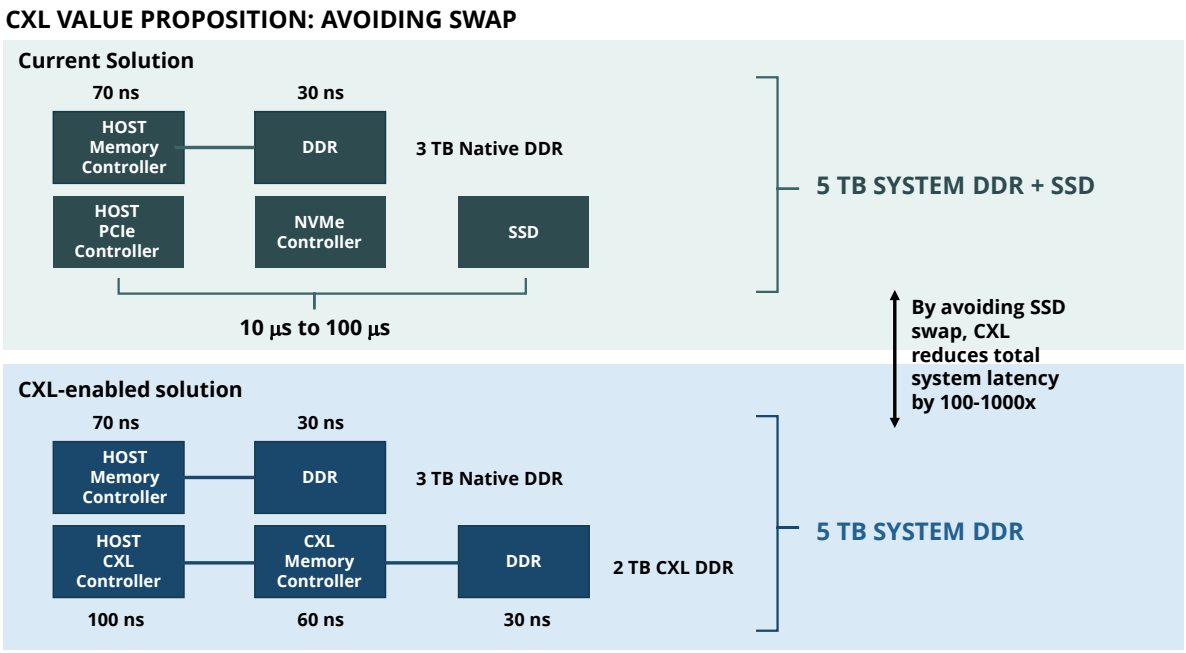


- **Memory Linked to Specific Devices Creating Underutilization and Stranding:** Server-dedicated memory often faces underutilization issues, particularly when the server's processing cores are not fully active. This inefficiency is especially challenging for Cloud Service Providers (CSPs), resulting in two major issues: under-utilized memory allocation and locked memory capacity within individual servers. Improving resource efficiency requires a new approach such as allowing memory to be shared across different servers or computing devices within the data center, allocating it dynamically based on computational needs. This approach has the potential to Increase the overall memory utilization across data center clusters, reduce the TCO for CSPs through more efficient hardware use, enhance application performance via flexible memory allocation, and improve adaptability to varying workload demands in real time.

- **Democratization of Memory and Latency-Sensitive Applications:** AI/ML is just one application that relies on HBM and low-latency memory to support efficient processing. Other workloads like In-Memory Database (IMDB), High-Performance Computing (HPC), financial modeling, and Electronic Design Automation (EDA) also have demanding memory and latency requirements. Advanced interconnect technologies capable of leveraging advances in DDR memory standards have become crucial in these modern computing environments. The need for advanced interconnect minimizes the need for using SSDs for additional memory, which can increase latency by a factor somewhere between 100X and 1,000X.

Figure 3: CXL-Based Memory Expansion Application Performance Improvement

(Sources: ABI Research; Timothy Pezarro, Microchip)



- **Heterogeneous, Commodity Computing Devices Implemented Widely across Deployments:** Commodity hardware is usually cheaper to deploy, maintain, develop, and refresh, so it is increasingly becoming an area of interest for data center vendors. Interconnects must be compatible with a variety of standardized off-the-shelf servers and devices. As of October 2024, there are 23 CXL 1.1 devices in the official CXL integrators list, with Type-1 and Type-2 coherency benefiting Fintech and VM migration use cases, in particular. There has also been significant software development: Linux Kernel 5.15 offers full support for Type-3 devices.

CXL'S VALUE PROPOSITION

CXL is being developed to solve major memory challenges in the data center, spurred by the key market drivers outlined above. This includes the performance difference between attached memory and SSD storage; underutilization of stranded resources; a growing memory wall created by the widening performance gap between processors and memory channels (DRAM density); and the need for more coherent memory across processors. Solving these challenges brings substantial compute performance improvements and TCO savings, which support efficient deployment of memory-intensive workloads like IMDBs, HPC, and AI/ML. Below, ABI Research breaks down CXL's value proposition across technical and commercial factors:

- **Coherency:** CXL maintains a coherent memory system where multiple components can share and access the same memory space in real time without data inconsistency. This enables resource sharing and the addition of more system memory across various devices connected by CXL.

- **Built on PCIe:** PCIe's ubiquity across various devices (including CPUs, GPUs, storage devices, and network interfaces) ensures compatibility with most devices and systems. In addition, it reduces development costs for the physical interface, while allowing CXL to benefit from PCIe's key features and improvements around bandwidth, power management, error detection, and scalability.
- **Memory Coherency:** Alignment with PCIe provides a foundation for implementing coherency mechanisms more efficiently and ensures greater system data consistency, which reduces the need for constant, often redundant, data transfers.
- **Switching/Routing & Workload Management:** CXL switches aide communication between various heterogeneous devices, enabling workload and memory management. This supports the disaggregation of memory from compute resources via shared memory (i.e., system disaggregation), enabling more scalable solutions and the transition toward software-defined virtualization by distributing the memory through the Virtual Machine (VM) network, enabling more flexible and dynamic resource access based on demands.
- **Open Standards & Cross Ecosystem Support:** Bringing new memory solutions to commercial scale takes between 2 and 3 years. Open standards with ecosystem support can reduce Time to Market (TTM) by accelerating verification, device compatibility, middleware development, and software development. The prominent role of CPU market makers (Intel, AMD) also brings stability to the market, especially by highlighting their CXL support across generations. This ensures that Research and Development (R&D) investment remains future-proofed.
- **Device Agnostic:** The seamless integration with widely deployed memory technologies like DDR5 and HBM demonstrates the flexibility of the fabric. CXL has been designed to support any device within a server—this means that it supports a heterogeneous model that is becoming increasingly common given its applicability to growing AI/ML workloads. In addition, device agnosticism enables the deployment and enhancement of commodity hardware. This has huge value given that most servers are upgraded using a modular approach, which is easier and more cost effective.
- **Security:** Over time, more advanced security features have been added to the CXL specification to protect data transiting a CXL link. This includes implementing a hardware root of trust, which can provide the basis for security and support requirements for both secure boot and secure firmware downloads. In addition, all three CXL protocols are secured via Integrity and Data Encryption (IDE), which provides confidentiality, integrity, and replay protection—vital for AI/ML and HPC computing applications.

OPPORTUNITIES & CHALLENGES FOR CXL

Given the key market drivers and CXL's value proposition outlined above, CXL technology has the potential to significantly improve the way data centers and enterprise markets address their server memory challenges, offering innovative solutions to overcome both technical and commercial hurdles. CXL technology opens a range of new opportunities, and this section dives into the most notable, as well as the key challenges it faces as it reshapes the future of data center computing architectures.

OPPORTUNITIES

- **Emerging Applications Like IMDB and Other High Memory Applications Will See Huge Benefits from Economical Memory Extension:** One of the earliest use cases that CXL is enabling is memory expansion. A number of high-value, data-heavy applications leverage CPUs, but are facing major challenges around memory limitations. Given CXL's economical and technical advantages over other solutions (e.g., swapping to SSD), it has plenty of commercial opportunities as data centers seek more economical solutions to address the issue of scaling memory to keep up with computational advances, while maintaining low-latency performance.

- **Many "Traditional AI" Use Cases Will Still Run on CPUs:** Although leading-edge AI use cases will rely on GPUs (with HBM) and ASIC accelerators, many commercially scaled "traditional AI" use cases will rely on CPUs. Therefore, given the standard's current compatibility with CPUs from AMD and Intel, the near term will see benefits in AI/ML from CXL deployment. For example, the built-in AI accelerator, AMX, found in Intel's 4th Gen Xeon Scalable processor family can utilize CXL to improve productivity. The widespread adoption of such AI-enabled CPU platforms presents a substantial market opportunity for CXL in its current form in optimizing AI/ML workloads.

- **AI/ML Growth Will Contribute to Greenfield Data Center Construction:** Technology refresh cycles will play a large part in scaled deployments of CXL due to cost considerations. But growth in enterprise AI/ML workloads is creating higher demand for new data center capacity. ABI Research expects this to lead to significant greenfield development over the next 2 years. CXL will benefit from this compute deployment, even if it only supports CPU-memory interconnect.

- **Early Partner-Led Solutions May Not Be Commercially Viable, but They Will Help Demonstrate and Mature Addressable Use Cases:** While CXL commercial solutions are not yet widely available, bringing industry partners together across the ecosystem is yielding promising prototypes that integrate diverse hardware and software components, demonstrating the potential of CXL deployment across various use cases. These early Proof of Concept (PoC) systems are crucial in demonstrating CXL's versatility and effectiveness. They validate potential performance improvements in real-world commercial products and enable stakeholders to identify and address potential integration challenges in refining CXL implementations.

- **Increasing Activity from Software Players Will Help Improve CXL-Related Application Performance:** New memory architectures (without application optimization) will negatively impact application performance due to cross-Non-Uniform Memory Access (NUMA) problems. However, partnering with VMware, MemVerge, SAP, Red Hat, or other Independent Software Vendors (ISVs) can solve these performance issues through middleware development, which helps optimize data placements across NUMA nodes and enable continuous monitoring of application memory patterns. Linux 6.10 support for CXL will reduce software barriers by enabling applications to manage NUMA nodes to ensure memory usage is optimized for CXL architectures.

- **Simplifying and Maturing CXL System Design:** The Joint Electron Device Engineering Council (JEDEC) announced a new standard to support CXL memory module implementation. The open body of CPU and DRAM vendors—as well as memory and CXL controller vendors—that defines standards applicable to their devices defined the specification of interface parameters, signaling protocols, environmental requirements, packaging, and other features as reference for specific target implementations of CXL-attached memory modules. This freely available standard will simplify system design and accelerate product development, as well as lending CXL more credibility. Given JEDEC's prominent position in the DRAM market, this is a significant step toward deployment.

- **Potential Opportunities for Direct GPU Support:** The CXL 3.1 specification will push the standard further by extending access to the memory pool currently shared by CPUs to GPUs through switching. Previously, pooling had been singled out around CPUs, single switches, and a limited number of multi-level switches. GPUs require ultra-low-latency memory from HBM (built into the chipset, such as in AMD's MI300X accelerator) to support leading-edge workloads in AI and HPC. However, applications are often limited by this memory and rely on adding more GPUs and splitting up vector calculations to alleviate this constraint. But early testing from Panmnesia and KAIST has shown that CXL IP can provide the necessary double-digit nanosecond latency to support these workloads. This is a significant opportunity to expand the GPU's memory pool given the scale at which GPUs are being deployed to support leading-edge workloads.

- **Key CPU "Market Makers" AMD and Intel Continue Support for CXL:** Although CXL 3.0 extends support to GPUs, the standard was originally developed to support CPUs. This means that investors are reliant on key "market makers"—AMD and Intel—for compliance and verification. CXL 1.0+ and CXL 2.0+ support has been available in AMD and Intel CPUs going back several years, including AMD's EPYC 9004 and 9005 series, and Intel's 4th and 5th Gen Xeon series. This reliance is also true for CXL controllers and memory devices: Astera Labs has continued to advance CXL 1.1 and 2.0 showing cloud-scale interoperability with leading CPU and memory vendors; Rambus has already adopted specification 3.1 in some of its memory controller solutions; and SK hynix announced at CXL DevCon 2024 in May that it will continue to incorporate CXL in its upcoming AI memory solutions. Continued and clearly communicated long-term support for the CXL specification from these key "market makers" is vital to driving industry investment and adoption.

- **Specification Development and Industry Alignment:** The rapid evolution of CXL specifications brings opportunities to unlock novel use cases. While the industry focuses on CXL 2.0 implementation, discussions about CXL 3.0's potential highlight the technology's promising future. This dual focus allows data centers, enterprises, and hyperscalers to strategically plan their adoption, aligning with their refresh cycles. This measured approach ensures optimal investment and paves the way for robust, efficient CXL implementations across the computing landscape, aided by the backward compatibility of all CXL standards, to date.

CHALLENGES

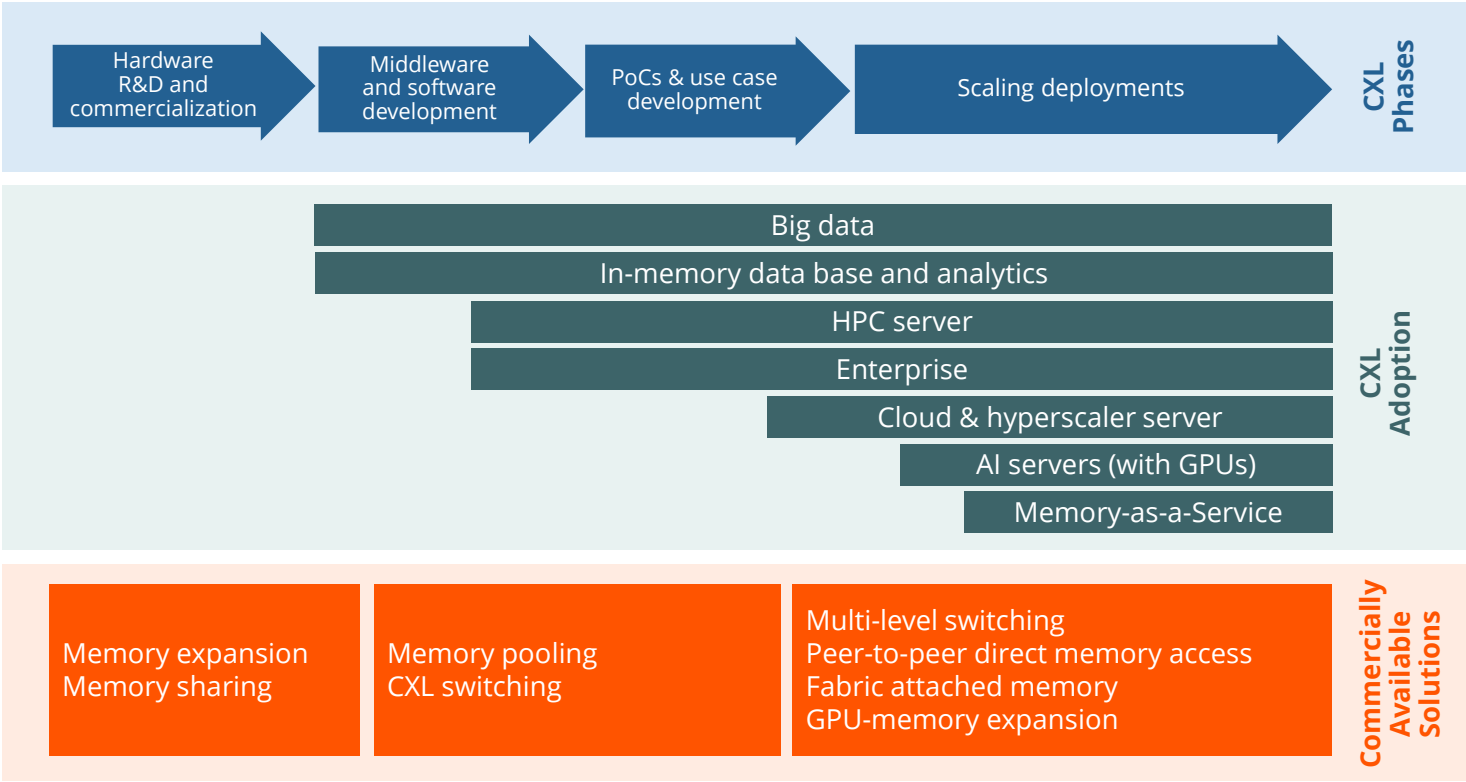
- **Hardware and Software Are Slow to Align, but CXL 2.0 Brings High Expectations:** The transition of CXL 2.0 devices to commercial scale presents an exciting opportunity for software optimization. While this phase may initially slow large-scale deployments, it is a crucial step in maximizing CXL's potential. This period of refinement will drive innovation in application development and system architecture, particularly benefiting hyperscalers. The cautious rollout ensures that when CXL solutions reach mass production, they will be robust, highly optimized, and ready to deliver unprecedented performance. Looking forward, this collaborative optimization phase will ultimately accelerate CXL adoption and its transformative impact on data center architecture.
- **Currently, CXL Interfaces Are Not Being Featured in GPU Roadmaps:** GPUs are content with using HBMs for high-bandwidth data access, while relying on PCIe for bulk data fetch from host node or Non-Volatile Memory Express (NVMe) SSDs. CXL 3.0 enables a fabric-attached memory that will allow GPUs to bypass the host or SSD for faster bulk data access. However, CXL 3.0 devices will not be commercially available in the near term
- **Time, Effort, and Cost Required to Refresh Brownfield Data Centers:** CXL deployment is largely reliant on technology refresh cycles given the time, effort, and cost to implement in brownfield data centers.

NEAR, MEDIUM, AND LONG-TERM EXPECTATIONS FOR CXL

CXL remains in an early stage with limited commercially available solutions mostly built utilizing CXL 1.1 or CXL 2.0. However, strong ecosystem support promises rapid advancements of the standard and wide adoption in the long term. ABI Research projects widespread CXL adoption by 2027 or soon after, allowing time for improving PCIe bandwidth and ecosystem maturation. The formation of the promoters group in 2019 is fairly recent and has attracted key players in the computing ecosystem, including the incorporation of rival interconnects Gen-Z and OpenCAPI into CXL, signaling robust support. The rate of adoption is also encouraging, and members of the consortium with a legacy in memory solutions report positively about their CXL roadmap. Figure 4 provides an overview of ABI Research's expectations for CXL.

Figure 4: CXL Adoption Expectations

(Source: ABI Research)



The following section dives deeper and explores near, medium, and long-term technical and commercial expectations for CXL.

NEAR TERM (NOW TO 3 YEARS)

CXL is on its way to achieving commercial maturity and an adoption inflection point. Most solutions are built on CXL 1.1 or CXL 2.0 with mass production of CXL 2.0 solutions expected in 2025. Commercial deployment will be supported by innovation-led partnerships by all members of the ecosystem:

- ISVs driven by CXL's commercial opportunity will continue to contribute toward software development and application R&D. Meanwhile, key players (VMware and Red Hat) support CXL verification and implementation within cloud environments.
- Key industry bodies and market makers are accepting the importance of CXL and embedding support within their releases, as exemplified by the recent cooperation between CXL and Linux on the latter's Kernel.
- Memory vendors are actively building instruction sets and reference architectures to showcase CXL's value proposition.

Industry support for CXL 2.0 software compatibility will enable steady adoption with a focus on memory and big data use cases.

By the end of the near term, commercial deployments will be coming to market, but this will remain focused on greenfield data centers and enterprise servers supporting memory applications. Although some deployments will support AI/ML, these will be focused on CPU-enabled workloads (i.e., graph-based) and will not support leading-edge AI models, as these will be reliant on GPU processing.

MEDIUM TERM (3 TO 7 YEARS)

The medium term will see an inflection point in CXL adoption. ABI Research expects that, by 2027, solutions leveraging CXL 3.0/3.1 based on PCIe 6.0 will be made commercially available with sufficient software support to accelerate commercial adoption. The higher bandwidth offered by PCIe 6.0 will enable more bandwidth-sensitive use cases, including in AI/ML.

The medium term will also bring maturing ecosystem partnerships that integrate hardware and software R&D. This tighter integration will accelerate product TTM and contribute toward deploying key applications and use cases. Combined R&D will help accelerate TTM, bringing memory pooling and switching to market based on CXL 3.0+. As CXL innovation stabilizes, expect key customers like hyperscalers to move away from "watch and wait" to accelerating commercial deployments for core memory use cases. This will also bring support for non-Gen AI/ML workloads by expanding CPU, FPGA, and other accelerator memory types.

Although ABI Research expects most commercial deployments to focus on core memory pooling and extension, vendor R&D will increasingly look to expand their Total Addressable Market (TAM) by supporting GPU-memory solutions. This R&D will rely on CXL 3.0+, and support and compatibility with GPU market makers NVIDIA, AMD, and Intel.

Hardware and software partnerships, especially those focused on cloud environments (VMware and Red Hat), will accelerate the push toward composable infrastructure and virtualization. The medium term will see PoCs and testing for new solutions like "Memory-as-a-Service" and software-defined memory, which will benefit from hyperscalers taking an active partner-led role in the CXL Consortium. However, it is unlikely that hyperscalers or CSPs will make these solutions commercially available.

LONG TERM (7+ YEARS)

ABI Research expects that, over the long term, CXL will offer scaled deployments across CSPs and enterprise servers with the majority of CPUs leveraging CXL 3.0+. These deployments will support a wide variety of high memory use cases, including AI/ML. In addition, ABI Research expects that some commercially available GPUs will offer compatibility with CXL, which will support some leading-edge AI/ML workloads. Multi-level switching and direct peer-to-peer memory will play a key role in enabling GPU-memory connectivity across an entire system. One of the most significant developments will be hyperscaler adoption of new memory solutions based on virtualization and composable infrastructure. In contrast to the more near-term forecasts, the long term will see the widespread availability of "Memory-as-a-Service," which will enable significant infrastructure cost savings.

CXL COST SAVINGS ANALYSIS

The following assessment explores the cost savings that CXL memory expansion can create for servers. This assessment was developed using DRAM configurations based on either CXL 1.1 or CXL 2.0 Add-in Cards (AICs). Although solutions are not commercially available yet, they are in pre-production, having gone through validation tests and certification by Intel. They are expected to move into mass production by the end of 2024.

Chart 2: CXL Cost Savings for 4,096 GB DRAM Configurations

(Sources: ABI Research; MemVerge)

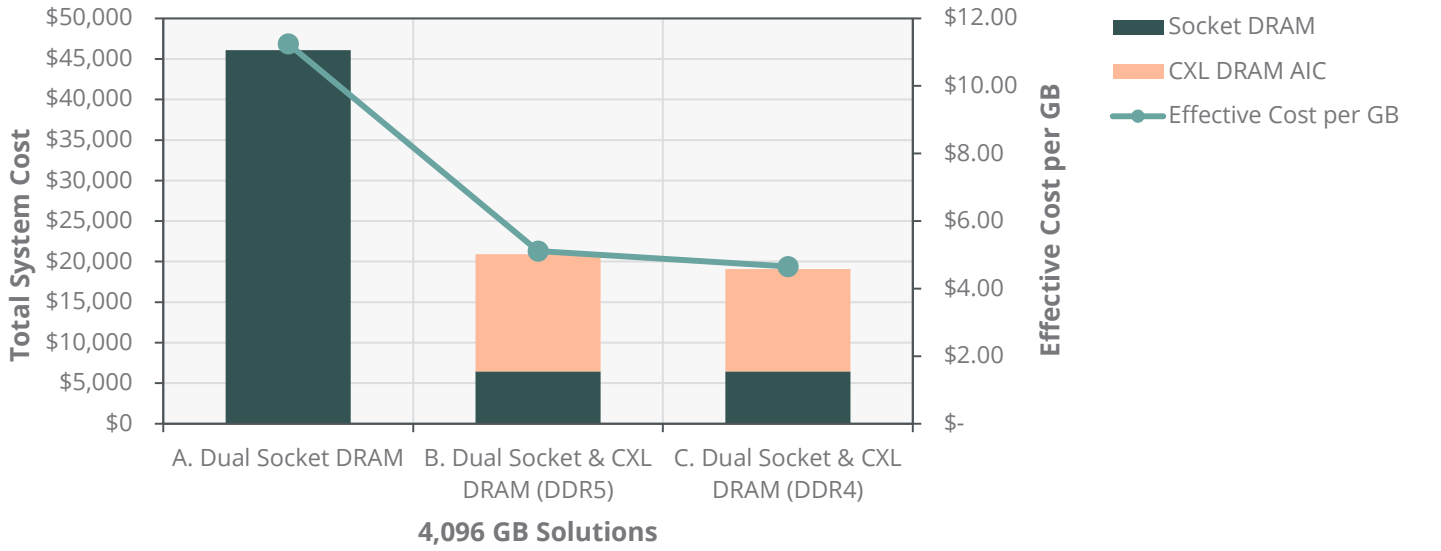
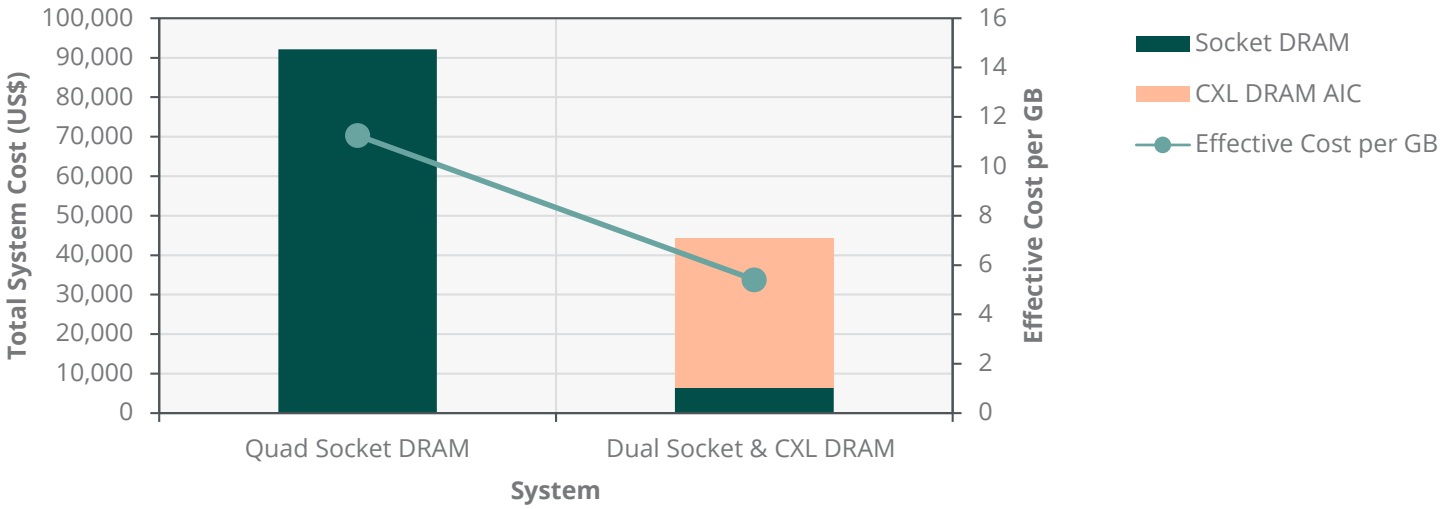


Chart 2 shows that by adding CXL AICs to enable memory expansion, memory cost per GB reduces around 56%. Achieving the same total memory capacity requires a less expensive server chassis.

Chart 3: CXL Cost Savings for 8,192 GB DRAM Configurations

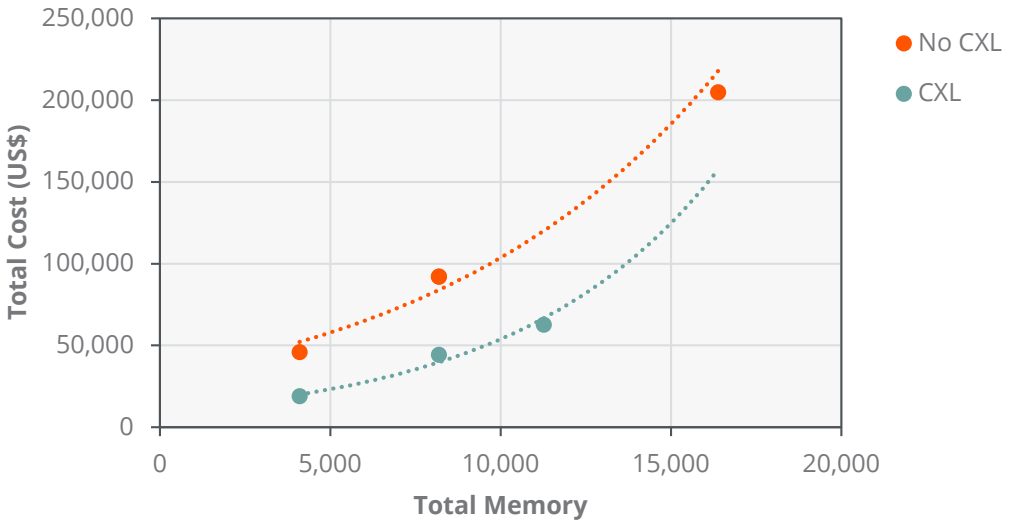
(Sources: ABI Research; MemVerge)



Similar to Chart 2, Chart 3 shows that memory expansion reduces cost per GB by around 52%. Chart 4 shows the impact that deploying CXL with Dual In-Line Memory Module (DIMM) can have on price per GB. With CXL, the memory per GB cost will be much lower for CPUs using memory expansion. On average, these savings will be around 55%.

Chart 4: Comparing Cost-to-Memory Ratio between CXL and No CXL-Based Solutions

(Sources: ABI Research)



Utilizing CXL memory expansion will clearly have a huge impact on CAPEX. It is widely understood that for hyperscalers, memory costs account for around 50% of total CAPEX. This implies that using CXL would have substantial savings for both greenfield and brownfield sites. Memory cost savings are a major demand driver for CXL, but total performance is also a core part of its value proposition. Memory expansion significantly improves total application performance by reducing swap between low-latency DDR and high-latency SSD.

STRATEGIC RECOMMENDATIONS

CXL offers clear technical and commercial benefits for a range of stakeholders in the market—especially CSPs. However, given the pace of development over the last 5 years, hardware and software R&D have become misaligned, which may hinder commercialization. Moving forward, ABI Research recommends that the ecosystem focus on building the logistical foundation for CXL commercial acceleration. Below are the key strategic areas to focus on::

- **Develop and Showcase Applications/Use Cases:** CXL shows significant promise; however, many potential customers are still following a "watch and wait" approach to emerging technology. Combating this mentality requires vendors to develop applications and use cases to showcase the potential performance enhancements and cost savings that CXL deployment can bring across various locations.

- **Deeper Alignment between Device Development and Software:** Application optimization will be a core barrier to CXL adoption at scale. Key prospects like CSPs will not deploy CXL without guarantees on application performance for all customer workloads. Software that can help bridge the application performance gap through auto-tiering, software-defined memory pooling (e.g., Samsung's Scalable Memory Development Kit, Astera Labs' Leo platform, and MemVerge's Memory Machine) will play a pivotal role, especially for CSPs that cannot optimize customer workloads for new memory architectures. Device vendors must look to expand partnerships with middleware vendors, focusing on early and deep alignment during R&D. This process will accelerate the availability of commercially valuable CXL solutions.

- **Expand Ecosystem with Partnerships and Contributions from Hypervisors (e.g., VMware) and Enterprise Software Vendors (e.g., Red Hat):** CXL 2.0+ will start to drive server disaggregation and virtualization through memory pooling and CXL switching. This will open up opportunities for hyperscalers to build software-defined memory solutions. Enabling this development requires deeper partnerships between device vendors and hypervisors like VMware. VMware is playing a role already through providing support for Samsung's CXL-enabled memory solutions. But as these solutions are rolled out, VMware and other hypervisors will be vital to enabling CXL value maximization.

- **Cooperation between Server Original Equipment Manufacturers (OEMs) and Chipset Vendors to Support CXL Validation and Build Logistical Foundation for the Market:** Software vendors must look to fast-track verification with assistance from hardware vendors to bring products to market quicker. Server OEMs like Dell, Lenovo, Hewlett Packard Enterprise (HPE), and Supermicro will have a critical role to play in the rollout of pooled memory chassis products.

- **"Market Makers" Must Continue to Release Long-Term CXL Product Strategies to Encourage the Ecosystem:** Cost and time lag to develop new memory solutions will create significant investment risk, meaning that CPU "market makers" must continue to showcase their long-term support for emerging specifications. ABI Research recommends that a 3+ year timeline for support is necessary to ensure continued interest and development targeting CXL. Beyond CPU giants, Arm could be seen as a potential "market maker." Arm recently announced support for CXL, and given its expanding position in the infrastructure market, this announcement will impact incumbents' decision-making.

- **Explore Opportunities to Collaborate and Verify Solutions with GPU Vendors:** GPUs still use attached HBM to support applications given the double-digit latency requirements, but recent news shows memory vendors like Panmnesia (in collaboration with KAIST) running PoCs that demonstrate CXL's opportunity for GPUs. Although CXL is not yet commercially applicable to GPUs, ecosystem players should start building relationships and integration with key GPU vendors to build long-term opportunities, especially for leading-edge AI development. This relationship must be built on an understanding of GPU vendor pain points and how CXL can be best used to solve challenges.

CONCLUSION

Memory is becoming one of the core commercial and technical challenges facing the data center market. CXL offers the market an economical solution to expand memory, while improving overall performance for memory-intensive applications like IMDBs, HPC, and AI/ML. However, CXL remains in a commercially early stage as the ecosystem aligns product R&D time frames and software development. Ecosystem and consortium members must engage more effectively across the hardware-software divide to help accelerate the deployment of commercially viable memory solutions, especially given the growing TAM of AI/ML workloads.

ACKNOWLEDGEMENTS

We would like to thank the following CXL Consortium members for their contributions to the content of this whitepaper: Ampere Computing, Astera Labs, Intel, MemVerge, MetisX, Microchip, MIPS, Samsung, SK hynix and Teledyne.

Published October 2024
157 Columbus Avenue
New York, NY 10023
Tel: +1 516-624-2500
www.abiresearch.com

We Empower Technology Innovation and Strategic Implementation.

ABI Research is uniquely positioned at the intersection of end-market companies and technology solution providers, serving as the bridge that seamlessly connects these two segments by driving successful technology implementations and delivering strategies that are proven to attract and retain customers.