Special Analysis

UALink Group Formed to Develop Standard High Speed Low Latency Interconnects for Accelerated HPC and AI Systems

Mark Nossokoff and Tom Sorensen
July 2024

HYPERION RESEARCH OPINION

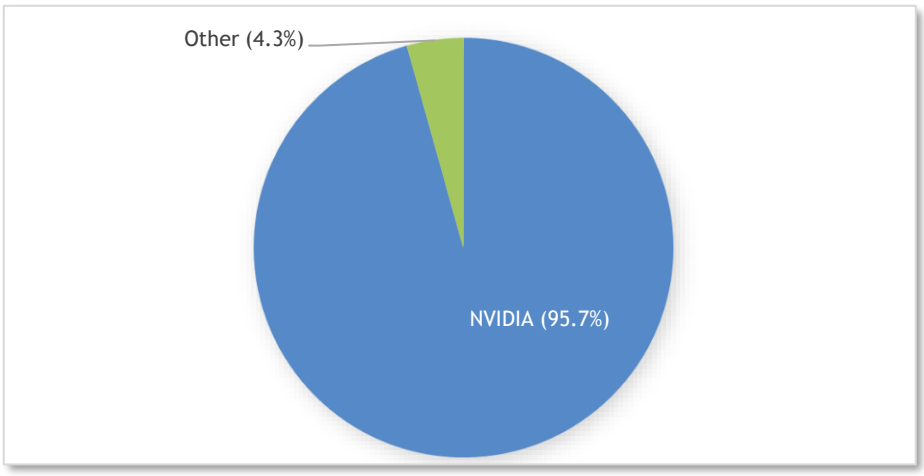
Openness and collaboration are seen by many as key elements for the long-term success for any technology developed to span a diverse and dispersed population of users and suppliers.

- Users want to know that the product they're using is stable, reliable, well supported, has a long life with continued investment in new features and support, and is available from more than one source.
- Suppliers want to provide their products to the widest audience as possible with minimal variations and customizations to optimize their investments.

Industry standards defined and supported by a large set of constituents within an ecosystem is the method by which to demonstrate and achieve openness. Standards also enable the promotion of competition in an industry that is dominated by a single vendor. The UALink group has been formed with the stated goal of developing a new industry standard dedicated to advancing high speed and low latency communication for scale-up AI systems. An unstated but possibly intended consequence of the group is to promote increased competition in the accelerated computing market.

FIGURE 1

2023 Accelerator Market Share (Units) for HPC and AI Advanced Technical Computing



Source: Hyperion Research, 2024

SITUATION OVERVIEW

Complex Hierarchy

The emergence of modern AI workloads is stressing the capability limits of most advanced technical computing architecture elements, including system interconnects. Today's complex HPC and AI architectures are comprised of several hierarchical interconnects, as summarized in Table 1.

TABLE 1

Hierarchy of Interconnects

Use Case	Definition	Interconnect
Data center to outside world	Connection between geographically dispersed data centers, and between remote users and the data center	Ethernet
Rack to rack	Connection between racks within a data center	Ethernet InfiniBand OmniPath Slingshot
Node to node (aka, server to server)	Connection between compute elements either between modules in a rack or within a shelf; typically carries traffic associated with Interprocessor communication such as MPI In many cases this interconnect is also used for rack to rack	Ethernet InfiniBand OmniPath Slingshot
Node to storage (aka, server to storage)	Connection between compute elements and storage elements either between modules in a rack or within a shelf	Ethernet InfiniBand OmniPath PCIe Slingshot
xPU to xPU xPU to memory	Connection between compute elements within a node	InfinityLink NVLink PCIe UCIe
Chiplet to chiplet in-package	Connection between functional elements within a package	InfinityLink NuLink NVLink UCIe

Source: Hyperion Research, 2024

The primary use cases of interest with respect to this research note are the node and xPU use cases. The node use cases are primarily focused on connectivity between servers within a data center, whether in the same rack or between racks. Traffic consists either of interprocessor communications (node-node) or data (node-storage). The physical network is typically Ethernet, InfiniBand, OmniPath, or Slingshot. Key performance metrics are bandwidth and throughput.

The xPU use case defines connectivity between CPUs and GPUs, with the physical networks being primarily PCIe for CPUs, and NVLink, InfinityLink, and PCIe for CPUs. Latency also is an important performance metric, in addition to bandwidth and throughput, for the xPU use case.

Accelerator Vendor Landscape

The interconnects mentioned in Table 1 represent both standard (ethernet, PCIe) and proprietary or single-sourced (InfiniBand, OmniPath, Slingshot, NVLink, InfinityLink, etc.) interconnects. Proprietary interconnects afford system vendors a measure of competitive differentiation but also require substantial investments to maintain a roadmap for each element (e.g., ASIC, network interface card, switch, connectors, and cables) of the ecosystem. Adopters of standard interconnects benefit from shared technology investment, choice of vendors for the various ecosystem elements, and interoperable solutions.

Infrastructure for GPU computing is currently dominated by NVIDIA. With multiple generations of GPUs in the field connected by its proven, proprietary NVLink interconnect, NVIDIA captures a large share of system vendor and end user spending for their AI infrastructure needs. While the NVIDIA solutions appear to be meeting most vendors' and users' performance requirements, vendors and users have historically preferred a broader choice in both technology and suppliers relative to primary elements of system architectures.

Competition is emerging in the accelerator market. AMD and Intel are shipping accelerator products and both have announced product and technology roadmaps to evolve their solutions. Each of the major cloud service providers have developed and are expected to continue developing AI accelerators, both for their internal AI needs as well as to offer to users as accelerated compute instance offerings. And there are numerous new accelerators in development or already on the market, many customized for some specific aspect of the overall generative AI ecosystem, spanning large-scale training, fine tuning, end-user training, and inferencing.

While the existence of multiple vendors and innovative alternative accelerator technologies affords some competition, NVIDIA's leadership position effectively establishes NVLink as today's de facto accelerator interconnect standard. Without a clear migration path to move users from NVIDIA's accelerator solutions to anticipated competitive performance and features, as well as stable, interoperable, and cost-effective alternative solutions, new entrants will likely have difficulty taking meaningful market share from NVIDIA.

THE ULTRA ACCELERATOR LINK (UALINK)

Announced in late May 2024, the UALink group aims to develop a new industry standard dedicated to advancing high speed and low latency communication for scale-up AI systems. With the goal of defining and establishing an open standard that will enable AI accelerators to communicate more effectively, the group expects the effort will enable system OEMs, IT professionals, and system integrators to create a pathway for easier integration, greater flexibility, and enhanced scalability of their AI-focused infrastructure.

Enabling the connection of up to 1,024 accelerators within an AI computing cluster and allowing for direct loads and stores between the memory attached to accelerators, such as GPUs, the 1.0 specification is planned to be available to members of the UALink Consortium in 3Q2024. Working to build on key elements of the PCIe roadmap, UALink seeks to decouple itself from full PCIe compatibility to allow optimizations for the extreme low latency needs of accelerated computing and adopt only those PCIe-based items required for an accelerator interconnect.

UALink complements the Ultra Ethernet Consortium (UEC), which was formed in July 2023 to establish an open, interoperable, high-performance architecture leveraging the broad Ethernet ecosystem that envisions a system interconnect solution that will capitalize on the ubiquity and flexibility of Ethernet to address demanding heterogenous HPC and AI performance, functionality, and scalability requirements. Note that there is a substantial overlap in the founding members for both nascent standard groups. Table 2 identifies the founders of both organizations.

TABLE 2

Ultra Accelerator Link (UALink) and Ultra Ethernet Consortium (UEC) Founders

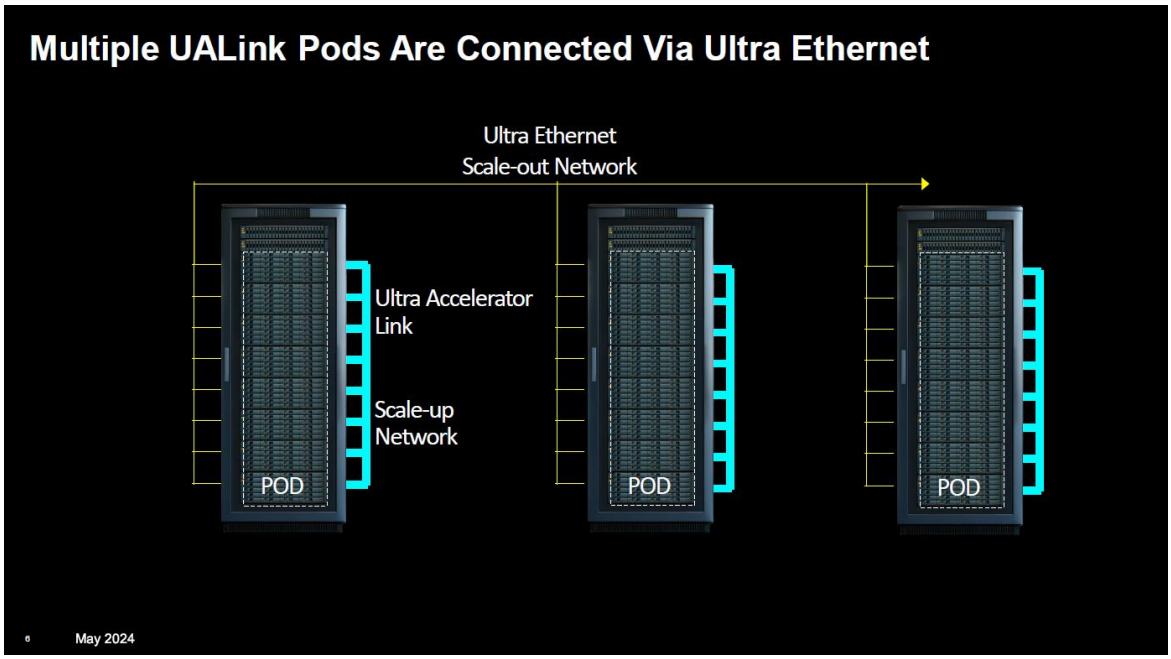
Founder	UALink	UEC
AMD	X	X
Arista		X
Broadcom	X	X
Cisco	X	X
Eviden		X
Google	X	
HPE	X	X
Intel	X	X
Meta	X	X
Microsoft	X	X

Source: Hyperion Research, 2024

Whereas UALink is focusing on the scale-up interconnect within a rack, UEC is initiating its efforts on the scale-out interconnect between racks. See Figure 2.

FIGURE 2

Scale-up and Scale-out Networks



Source: UALink press briefing, May 2024

FUTURE OUTLOOK

Noticeably absent from the UALink group is NVIDIA. NVIDIA's NVLink presents a high bar and challenging target for UALink to achieve and surpass. With multiple generations in production providing a stable, high performance solution today, NVLink isn't likely to stand still and is expected to continue to evolve to address anticipated future system requirements. UALink will need to not only achieve stability with its solution and interoperability between supporting vendors, but also aim ahead of where NVLink is today.

Additionally, while establishing a stable, performant, interoperable physical interconnect available from and supported by multiple vendors is necessary to be competitive, it is not sufficient. NVIDIA's CUDA software development environment is also a formidable barrier to overcome. Users are loathe to port their application codes to new hardware only to achieve equivalent performance and feature capabilities; they would rather focus their investments in solutions to deliver new features, capabilities, and value to their customers.

Despite the challenges identified above, UALink is headed in the right direction. The organization consists of leading vendors from across the advanced technical computing ecosystem (system, switch, cloud, hyperscaler, accelerator, I/O, component design) with a rack record of success in collaborating

to define and bring new standards to market. Aligning itself and collaborating with the UEC is also a wise move as the more encompassing the ecosystem adoption, the higher likelihood of success. And while admittedly outside the scope of UALink, establishing a tie-in with and support of higher-level software efforts to aid users in porting from CUDA to a standard accelerator software environment would further enhance the probability of UALink achieving its goal of widespread adoption.

About Hyperion Research, LLC

Hyperion Research provides data-driven research, analysis and recommendations for technologies, applications, and markets in high performance computing and emerging technology areas to help organizations worldwide make effective decisions and seize growth opportunities. Research includes market sizing and forecasting, share tracking, segmentation, technology, and related trend analysis, and both user & vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). Hyperion Research provides thought leadership and practical guidance for users, vendors, and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue
St. Paul, MN 55102
USA
612.812.5798
www.HyperionResearch.com and www.hpcuserforum.com