

FabreX Hyper-Performance Network



The enormous increase in data is driving fundamental change to data center architecture

The remarkable increase in the amount of data being collected, and that must be analyzed and stored, is driving the rapid adoption of advanced data analytics and Artificial Intelligence (AI) that is challenging the fundamental architectures of today's data centers, in a way not seen since the 1990s.

AI, and the associated Machine Learning (ML) and Deep Learning (DL) applications are fueling demand for fundamental change in the creation of compute and storage clusters. Faster and larger storage arrays and a rapid proliferation of specialized compute accelerators, like GPUs, FPGAs and custom ASICs, are creating bottlenecks and configuration problems for the interconnect systems, as the traditional networks were never designed to handle the performance requirements of these workloads and devices.

Further, the rapid pace of change in acceleration technology and AI software fuels the necessity for flexible and easy to upgrade architectures, capable of incorporating new technology without demanding forklift upgrades to expensive equipment. This means disaggregating elements of the traditional server into separate pieces that can be easily shared. But, in order to effectively disaggregate storage and accelerators, the interconnects must support both an exceptionally low latency AND high bandwidth.

And, of course, data center managers want to drive high utilization of expensive new storage and acceleration products to keep both Capex and Opex costs down. Add all these up and these Advanced Scale Computing, Enterprise, Cloud and Edge data centers need both scale-up and then scale-out resources across the cluster and require a network technology that will grow in both directions.

The Solution: GigaIO FabreX


GigaIO FabreX is a fundamentally new network architecture that integrates computing, storage and other communication I/O into a single-system cluster fabric, using industry standard PCI Express technology. PCIe is used internally to power virtually every server on the market today. GigaIO enables true host-to-host communication across PCIe and makes true cluster scale networking possible, with direct memory access by an individual server to system memories of all other servers in the cluster fabric, for the industry's first in-memory network.

This new architecture enables a hyper-performance network with a unified, software-defined, composable infrastructure.

With its exceptional low latency and high bandwidth, FabreX makes the disaggregation of storage and accelerators possible. And by offering standard Redfish APIs, leading third-party orchestration and composition software can easily run on FabreX, making a true software defined infrastructure that dynamically assigns resources to match changing workloads. Gone are the days of having to make massive and costly changes to large numbers of servers to incorporate new technology.

Upgrade or add compute, storage and application accelerators at the component level that plug-n-play with your environment. Every major subsystem can now operate on its own upgrade cycle. And the total cost of the system is optimized as FabreX drives much higher utilization of the resources.

FabreX is built to be an open platform and offers robust, industry standard APIs at all levels of the software making the integration of FabreX into your existing data center management systems easy and risk free.



FabreX is 100% compliant with the industry's leading standard, PCI Express, insuring high levels of support and compatibility from an enormous selection of vendors and technologies. Every new storage, acceleration and compute technology fully support the latest PCIe standards ensuring full access to the latest and most capable technologies.

And as IT experts begin to use new architectures for moving compute to the data at the edge of your networks, FabreX is a natural choice, offering the greatest depth of support in storage, compute and acceleration technology, the highest density, lowest power and best performance.

Performance

FabreX delivers the industry's lowest latency AND the highest effective bandwidth. Latency from system memory of one server to system memory of any other is less than 200ns – true PCIe performance across the entire cluster. The current Gen 3 implementation delivers 256Gbits/sec bandwidth, soon to scale up to 512Gbits/sec with PCIe Gen 4.

Flexibility

FabreX can unite an unprecedented variety of resources, connecting accelerators of all types including GPUs, TPUs, FPGAs and SoCs to other compute elements or storage devices, such as NVMe, PCIe native storage, and other I/O resources. FabreX can span multiple servers and multiple racks to scale up single-host systems and scale out multi-host systems, all unified via the FabreX software.

Efficiency

Featuring 100% PCI-SIG compliance, the FabreX switch can integrate heterogenous computing, storage and accelerators into one symmetrical system-area cluster fabric, so you can do more with less. Patented GigaIO technology strips away unnecessary conversion, software layers and overheads that add latency to legacy interconnects.

Memory-Centrix Fabric

FabreX is the next generation, memory-centric fabric for a changing compute world. Effortlessly connect new memory / storage products, the multitude of new accelerators and your choice of processors either directly attached or via server configs like NVMe-oF.

Open Platform, Standards-Based

FabreX is built on, and 100% compliant with, the industry's most widely adopted standard, PCI Express, insuring low risk, easy integration and long life.

Further, the FabreX operating system easily integrates third-party applications with its open-source design, including the DMTF open-source Redfish® APIs to provide unprecedented integration with a range of third-

party applications for fabric automation, orchestration, resource allocation and job management.

The result is lower Capex and Opex through less hardware, higher utilization of resources, lower power consumption, and less cooling. Avoid overprovisioning and add just the elements you need. Maximize utilization of the footprint of your data center and contribute to your bottom line.

The FabreX Building Blocks:

FabreX Operating System

The FabreX Operating System is the central building block for the FabreX network and the software engine that drives the performance and dynamic composability of GigaIO software-defined infrastructure (SDI). This Linux-based, resource-efficient software layers onto FabreX hardware for easy-to-use composing of computing clusters on-the-fly.

The FabreX OS is designed to be an open environment with robust APIs to encourage third party development to enhance the usability, flexibility and performance of the network.

The FabreX Switch

A classic Top-Of-Rack network switch, the FabreX Switch communicates with FabreX host drivers to identify and coordinate resources required by the hosts.

FabreX Network Adapter Card

The FabreX Network Adapter card is the high-performance, cabled interface to cluster subsystems across the FabreX hyper-performance network. The card includes both host and target (for PCIe I/O) modes and is FPGA-powered for configuration flexibility. Applications can access remote PCIe devices as if they were attached to the local system.

The FabreX Network Adapter Card is designed for maximum cable length at full PCI Express Gen 3 speed. Copper cables can connect up to 5 meters or use active optic cables to extend the distance up to 100 meters.

The flexibility of the FabreX interconnect breaks the constraints of old architectures, opening up possibilities for new configurations that maximize utilization of all elements within your Advanced Scale Computing, Enterprise, Cloud or Edge data centers. At the same time, FabreX allows for far greater utilization of these next generation technologies, keeping both your capital expenditures and operating costs down.

FabreX: The World's Highest Performing Network

Through an all-new architecture, GigaIO introduces a hyper-performance network with extreme flexibility, latency and bandwidth that enables a true unified, software-driven composable infrastructure.