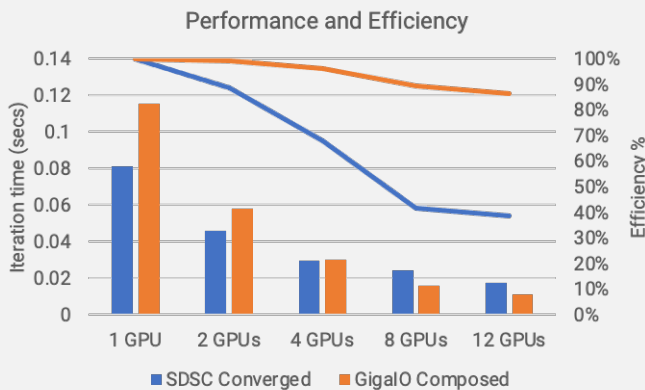
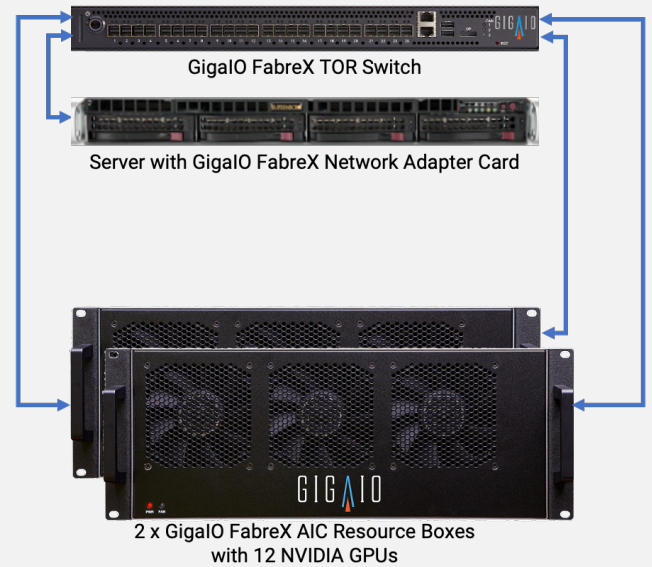


Composed Architecture Demonstrates Superior Utilization and Performance Versus Converged Architecture for Large Scale Simulation GPU Workload



Configuration



Lower Cost, Better Performance, Higher Scale, Increased Efficiency, and Easier Management of GPU Workloads

Converged configuration – The SDSC Comet is a converged configuration: all GPU resources are locked inside the servers and cannot be shared. The Comet test configuration used three 4U nodes interconnected with InfiniBand. Each node is a dual processor with 4 NVIDIA P100 GPUs for a total of 12 GPUs.

Composed configuration – GigaIO composed architecture is a single 1U server connected to the FabreX™ PCIe Network fabric. The FabreX system uses 12 NVIDIA 1080Ti GPUs distributed over two Add-in-Card resource boxes. GPU to GPU communication uses FabreX at PCIe latency and bandwidth.

Results Summary – Test results demonstrate the dramatic impact of increasing GPU utilization rate, even with less performant GPUs. While a single \$6,500 server achieved better performance than a \$1,000 GPU, when using four GPUs, a step function improvement in utilization (98% vs. 62%) drives improved performance. Because the GPUs have to communicate over the UPI bridge in Comet over InfiniBand to link GPUs located in different servers, vs. peer to peer in FabreX, performance in Comet drops to 39% with 12 GPUs vs. 82% with FabreX.

Composability delivers faster time to results, even with less performant GPUs. Users can reduce CapEx or OpEx, or get faster time to results with the same investment.

GigaIO FabreX is a Rack-Scale composable infrastructure solution that delivers the unlimited flexibility and agility of the cloud, at a fraction of the cost. Benefits include:

Improved system agility by disaggregating system resources on the fly and creating shared resource pools that can then be dynamically composed in real-time.

Slashed Total Cost of Ownership by enabling device sharing which increases resource utilization and eliminates over provisioning, resulting in reduced CapEx and OpEx.

Simplified and automated system set-up, administration and serviceability with freedom of choice for management tools from powerful CLI and Redfish APIs to ready-to-run, off-the-shelf enterprise-class orchestration software.

Seamless support for any PCIe-compliant device including servers, CPUs, memory, 3D-XPoint, storage, GPUs, FPGAs, specialty ASICs and NICs.

Blazing system performance with industry leading PCIe latency and bandwidth throughout the rack and beyond. As PCIe resources are added they immediately benefit from the native PCIe performance as all data transfers and buffers are completely eliminated.

Visit www.gigaio.com to discover more about GigaIO and FabreX, the industry's only pure PCIe Network Fabric.