

REFERENCE DESIGN:

GPU Composability Liquid Powered 1/2 Rack

Liquid Powered GPU composability delivers unprecedented rack-scale agility, marking the next evolution in bare-metal data center architecture. Liquid Composable allows users to manage, scale out, and configure physical bare-metal servers in seconds. Liquid Command Center software, delivered with the Liquid Grid PCIe switching fabric, enables true disaggregation of industry standard server components.

The Liquid Powered GPU Composable Rack enables multiple graphics processing (GPU) elements to be deployed natively across PCI-Express (PCIe) and dynamically allocated to any Inspur Node instantly. GPU-to-CPU resource assignments can be configured and reconfigured through policy-based automation in real time, as infrastructure needs change.



Liquid Powered GPU Composable 1/2 Rack (Max Configuration)

Compute

12x Compute Nodes
Dual Intel Xeon Scalable 28-Core Processors per Node
512 GB DDR4 DRAM per Node

Storage

72x U.2 Solid-State Drives (SSD), 6.4 TB per SSD
460 TB of Aggregate Capacity
6x SSD per Node

Networking

12x Network Adapters (NIC), Dual 100 Gb per NIC
2,400 Tb/s (300 TB/s) of Aggregate Throughput
1x NIC per Node

Graphics

24x Graphics Processing Units (GPU)
Supports FHFL Double-wide GPU Elements
Dynamically Assign up to 16x GPUs per Node

Composable Fabric

Liquid Grid (Managed PCIe Gen 3.0 Switching Fabric)
Liquid Command Center (Software)

Top-of-Rack Networking

Mellanox 100/200 Gb Ethernet

Control Plane

48-port 10 Gb / 1 Gb Ethernet

Mini-SAS HD Cables

Gen 3.0 (Passive or Active Copper, Active Optics)

Power

Up to 12Kw

Form Factor

24 Rack Units (RU)



Liquid Command Center (Software)

Extensible composable infrastructure management software that automates, orchestrates, and dynamically composes bare-metal machines from pools of disaggregated bare-metal elements..



Expansion Chassis

PCIe-connected GPU expansion chassis by Inspur that holds 4x double-wide GPU elements in a 2 RU chassis. Flexible topology design supporting multiple GPU types.

Liquid Grid

The ultra-low latency, intelligent, and managed switching fabric that electrically interconnects pools of disaggregated system elements.

Compute Elements (Nodes)

Dual processor Intel Xeon Scalable compute elements with 16x DDR4 DIMMs per node. Includes up to 6x SSDs and 1x OCP NIC per node. Four nodes per 2 RU chassis.

The Liquid Powered GPU Composable Rack allows multiple GPU elements to be dynamically and instantly allocated to any bare-metal server in seconds. GPU resources can be configured and reconfigured on demand as workflow requirements demand. Liquid unleashes the power of GPU scaling to enable the highest levels of resource utilization.