LIQID

Case Study:
Building One of the World's Fastest Off-the-Shelf GPU Supercomputers

orange™

POWERED BY DELL Technologies

Overview

GPU supercomputing solutions provide near unprecedented computational performance to deliver the aggregated power of multiple traditional CPU based systems designed for data-intensive AI applications, with some systems boasting 2 petaFLOPS of computing power. In testing scenarios, the top GPU-based systems achieve over 200x the performance over other CPU-based solutions on the market.

In practical terms, these bespoke systems deliver the most powerful GPU supercomputing solutions commercially available today. The power of these systems can be demonstrated using benchmarks such as Resnet, a testing tool used to measure performance in data training and inference operations commonly associated with GPU computing for artificial intelligence workflows. Resnet measures the number of images per second a compute system is capable of recognizing, delivering the performance numbers listed below (Page 6) for both data training operations and data inference.

Powerful GPU supercomputers offer incredible performance for AI and computationally intensive work loads. Liqid sought to create a GPU supercomputer solution that provides compelling GPU performance for visually intensive workloads by leveraging the broader NVIDIA® Quadro RTX™ family of products. The Liqid ThinkTank LQD8360 composable GPU system was designed to unleash the power Nvidia RTX 8000 GPUs in a performance centric solution for the most demanding visual workloads.

LIQID

orange™

— POWERED BY —
DELL Technologies

Solution

To examine additional architectural approaches to GPU supercomputing, Liqid partnered with telecom provider Orange Silicon Valley, a business innovation subsidiary of global telecommunications operator Orange, and Dell Technologies OEM | Embedded & Edge Solutions to design the Liqid ThinkTank, a dynamic GPU powered supercomputer capable of delivering massive performance at a reasonable cost by leveraging standard commodity hardware, low latency PCIe fabrics, and intelligent software.

To build the Liqid ThinkTank, Liqid's PCIe composable fabric is deployed with off-the-shelf Dell EMC PowerEdge R640 server nodes and then composed with up to 20x NVIDIA® Quadro RTX™ 8000 GPUs in an extension chassis capable of housing these GPUs in a separate physical enclosure, or JBOG (Just a Bunch of GPUs). Liqid's Command Center software paired with an intelligent ultra-low latency PCIe-based fabric enables GPUs to be dynamically configured with the Dell EMC PowerEdge R640 nodes at the bare-metal level.

With bare-metal composability, the Liqid ThinkTank composable GPU system permits any number of GPUs to be physically assigned to any node on the fabric without any physical chassis redesign efforts. So far, this collaboration between Liqid, Dell Technologies OEM | Embedded & Edge Solutions, and Orange Silicon Valley has resulted in:

1. The world's highest-capacity expansion chassis (JBOG) capable of housing 20 full-size data center class GPUs with managed power and cooling;
2. Composability enables up to 20x GPUs connected to a node via PCIe fabric and Liqid Command Center software, making the Liqid ThinkTank composable GPU system one of the world's fastest single-node, or multi-node, production-ready AI supercomputers;
3. Equipped with up to 20x NVIDIA Quadro RTX 8000 with 48GB of VRAM each, the Liqid ThinkTank composable GPU system delivers one of the highest GPU memory capacities with 960GB of VRAM;
4. An optimized Dell BIOS enables the Dell EMC PowerEdge R640 to support up to 20x NVIDIA Quadro RTX 8000 GPUs on a single compute node with full GPU peer-to-peer capability enabled;
5. Liqid Command Center enables GPUs to be dynamically reallocated to various nodes as workloads are completed, preventing valuable resources from sitting idle.

LIQID

orange™

— POWERED BY —
DELL Technologies

Results

This approach confirms that Liqid Command Center and a composable PCIe fabric were able to support orchestration of up to 20x NVIDIA Quadro RTX 8000 GPUs to a Dell EMC PowerEdge R640 node as well as enabling fundamental features like NVIDIA GPUDirect™ peer-to-peer, which allows high-speed direct memory access (DMA) transfers between the memory regions of each GPU located on the fabric, directly storing and loading data between the memories of two GPUs.

Below are the test results for the GPU peer-to-peer bandwidth and latency testing. These results show benefit of enabling direct DMA transfers between multiple GPUs.

Peer-to-Peer Performance Testing Summary:

	Bandwidth(GB/s)	Latency (µs)
Peer-To-Peer Disabled	8.59	33.65
Peer-To-Peer Enabled	25.01	3.1
Delta	291%	1085%

LIQID

orange™

— POWERED BY —
DELL Technologies

Peer-to-Peer Latency Testing:

P2P Enabled (Latency)

P2P=Enabled Latency (P2P Writes) Matrix (us)

GPU	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	3.21	3.26	3.04	3.05	3.02	3.05	3.16	3.15	3.16	2.86	2.96	3.30	2.94	3.09	2.98	2.97
1	3.10	3.38	2.89	3.09	3.05	2.94	3.09	2.99	2.98	3.08	3.20	3.02	3.14	3.00	3.20	2.87
2	3.07	3.16	2.97	3.27	3.18	3.16	3.17	3.30	2.77	2.98	3.12	3.01	3.06	3.07	3.10	2.91
3	3.00	3.02	3.04	3.50	2.99	2.83	2.98	3.18	3.04	2.92	2.95	2.90	2.97	2.82	2.96	3.21
4	3.16	3.10	3.30	3.03	3.15	3.13	3.12	2.74	2.88	3.15	3.00	3.00	2.91	3.01	3.12	3.05
5	2.89	3.05	3.08	3.03	3.08	2.91	3.14	3.11	3.12	3.14	3.07	3.14	3.06	2.94	3.12	2.93
6	2.84	2.89	2.91	3.07	3.10	3.01	3.20	3.09	2.92	2.92	3.17	3.13	2.95	2.91	2.96	3.12
7	3.09	3.00	3.20	3.17	3.22	2.93	3.07	3.03	3.21	3.14	2.97	2.91	3.07	3.05	3.04	3.03
8	2.96	3.21	2.97	3.12	3.01	2.92	2.80	3.02	3.39	2.99	3.10	3.17	3.14	2.76	3.06	3.28
9	2.99	2.90	2.95	3.13	2.84	3.02	3.09	3.30	3.08	3.20	3.20	3.19	2.95	2.92	2.96	3.14
10	3.19	3.10	3.06	3.15	2.87	2.84	3.02	3.01	3.07	3.02	3.22	2.90	3.12	3.05	2.93	2.90
11	3.08	3.08	2.98	3.21	2.83	3.02	3.21	3.23	3.02	2.90	3.09	3.03	3.05	3.01	3.23	2.97
12	3.00	3.13	2.98	3.08	3.25	3.13	3.17	3.00	3.01	3.21	3.03	3.02	3.14	2.98	3.11	2.91
13	3.19	3.10	3.18	3.00	3.27	3.18	2.98	3.20	3.25	3.02	3.01	3.01	3.07	3.22	3.00	3.05
14	2.92	2.98	3.20	2.85	3.20	3.29	3.13	3.00	3.07	2.90	3.08	3.05	2.96	3.07	2.97	3.10
15	2.88	2.92	2.99	3.34	3.11	3.17	3.06	2.97	3.14	3.06	2.97	2.92	2.98	2.95	3.35	3.14

P2P Disabled (Latency)

P2P=Disabled Latency Matrix (us)

GPU	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	3.04	33.25	35.05	34.30	33.36	33.48	34.04	33.99	33.19	34.26	33.60	34.21	33.52	32.80	33.25	31.98
1	33.65	3.21	32.78	34.13	33.94	34.48	32.21	32.63	33.65	34.48	33.31	32.78	32.62	32.34	32.77	30.87
2	34.00	33.99	2.95	33.50	33.39	32.63	33.12	33.17	33.38	33.58	32.40	33.60	32.00	32.31	32.75	31.33
3	33.94	34.98	33.60	3.48	32.96	33.33	33.79	33.38	34.27	33.95	33.57	35.12	32.91	32.98	32.37	32.18
4	33.38	34.17	34.60	34.40	3.02	34.33	34.23	33.27	34.08	34.38	33.72	33.96	33.04	31.57	32.85	32.40
5	32.67	32.93	32.83	32.35	33.58	3.05	33.26	32.72	34.07	33.25	34.36	33.37	33.53	32.07	32.66	30.70
6	34.19	33.20	33.33	32.18	33.84	32.69	3.14	33.82	34.28	34.15	32.39	34.43	32.74	32.52	31.96	32.16
7	32.78	33.34	33.57	32.40	33.48	34.15	33.92	3.11	33.42	33.31	33.39	33.43	34.04	33.63	33.67	31.69
8	33.57	33.49	33.85	32.66	34.82	34.39	33.22	33.99	3.38	33.63	34.34	34.31	34.58	31.63	34.69	31.95
9	33.76	32.99	33.48	32.76	33.68	34.16	32.78	34.05	34.62	3.08	33.64	32.78	33.25	31.17	33.72	33.69
10	33.59	32.47	34.88	34.21	33.21	34.71	33.66	32.95	34.18	33.26	3.36	33.69	33.73	33.04	32.82	31.84
11	34.72	33.05	34.70	33.34	33.43	34.32	33.07	32.50	34.22	34.25	32.52	3.25	34.36	32.52	32.84	32.64
12	33.57	32.82	33.03	35.18	34.48	32.87	34.76	32.63	33.76	32.97	32.86	33.83	2.90	33.04	32.47	31.81
13	32.87	32.49	31.19	33.14	32.54	33.60	34.05	33.02	32.94	33.42	32.46	32.18	33.06	2.99	31.90	30.22
14	33.58	32.62	32.71	33.34	35.09	33.51	33.20	33.12	33.81	32.05	33.09	34.57	32.48	31.80	3.13	30.47
15	31.98	31.91	32.73	33.29	32.81	31.94	32.62	32.23	32.03	33.29	33.00	33.88	32.25	32.40	32.61	2.95

Peer-to-Peer Bandwidth Testing:

P2P Enabled (Bandwidth)

Bidirectional P2P=Enabled Bandwidth Matrix (GB/s)

D\D	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	540.28	25.03	25.02	25.03	25.00	24.31	24.34	24.33	24.33	24.16	24.27	24.28	24.29	24.30	24.24	24.34
1	25.01	539.54	25.02	25.02	25.01	24.33	24.34	24.35	24.34	24.16	24.28	24.30	24.29	24.23	24.21	24.34
2	25.03	25.02	535.85	25.03	25.03	24.33	24.34	24.33	24.33	24.18	24.29	24.31	24.31	24.30	24.32	24.35
3	25.02	25.03	25.02	538.67	25.01	24.34	24.35	24.35	24.35	24.20	24.29	24.31	24.29	24.30	24.28	24.34
4	25.02	25.03	25.02	24.99	544.78	24.35	24.35	24.29	24.34	24.38	24.24	24.28	24.33	24.23	24.30	24.23
5	24.33	24.32	24.25	24.34	24.18	539.54	25.02	24.98	25.02	25.02	24.34	24.34	24.37	24.27	24.20	24.30
6	24.32	24.34	24.35	24.27	24.20	25.03	543.79	25.03	25.03	25.03	24.35	24.34	24.37	24.36	24.26	24.30
7	24.35	24.27	24.36	24.35	24.27	25.03	25.02	535.13	25.01	25.02	24.36	24.33	24.37	24.34	24.21	24.29
8	24.33	24.35	24.35	24.34	24.35	25.02	25.01	25.02	535.97	25.02	24.35	24.34	24.34	24.33	24.20	24.19
9	24.31	24.21	24.21	24.21	24.22	25.00	25.01	25.03	25.01	531.73	24.33	24.23	24.31	24.23	24.19	24.27
10	24.28	24.30	24.30	24.30	24.20	24.33	24.35	24.26	24.36	24.26	529.77	25.02	25.01	25.02	25.01	24.33
11	24.28	24.30	24.30	24.30	24.15	24.32	24.34	24.34	24.35	24.24	25.00	537.31	25.03	25.03	25.02	24.32
12	24.30	24.30	24.31	24.30	24.20	24.36	24.34	24.35	24.35	24.39	25.03	25.02	525.94	25.02	25.02	24.33
13	24.29	24.30	24.31	24.30	24.25	24.35	24.35	24.36	24.35	24.20	25.03	25.03	25.02	530.22	25.00	24.35
14	24.24	24.17	24.30	24.30	24.33	24.23	24.23	24.30	24.35	24.39	25.03	25.02	25.01	25.00	530.77	24.27
15	24.35	24.32	24.33	24.33	24.24	24.29	24.29	24.27	24.30	24.25	24.33	24.32	24.32	24.35	24.20	532.96

P2P Disabled (Bandwidth)

Bidirectional P2P=Disabled Bandwidth Matrix (GB/s)

D\D	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	542.53	8.60	8.61	8.68	8.64	8.56	8.66	8.63	8.64	8.66	8.61	8.59	8.62	8.63	8.65	8.66
1	8.59	541.16	8.57	8.58	8.61	8.63	8.68	8.57	8.67	8.67	8.61	8.61	8.65	8.63	8.65	8.66
2	8.69	8.67	544.80	8.59	8.50	8.55	8.56	8.61	8.65	8.63	8.60	8.65	8.56	8.65	8.64	8.68
3	8.67	8.64	8.71	545.77	8.61	8.68	8.64	8.63	8.62	8.60	8.58	8.61	8.61	8.66	8.57	8.56
4	8.66	8.65	8.59	8.63	531.46	8.68	8.58	8.61	8.66	8.67	8.66	8.61	8.64	8.58	8.55	8.64
5	8.61	8.53	8.65	8.62	8.66	527.06	8.62	8.61	8.60	8.69	8.62	8.63	8.63	8.65	8.62	8.55
6	8.53	8.50	8.56	8.63	8.61	8.62	532.19	8.62	8.64	8.67	8.66	8.61	8.64	8.69	8.61	8.58
7	8.58	8.58	8.66	8.61	8.61	8.59	8.70	534.35	8.66	8.54	8.54	8.72	8.59	8.64	8.60	8.51
8	8.65	8.60	8.54	8.54	8.56	8.59	8.59	8.60	530.81	8.56	8.57	8.64	8.57	8.60	8.58	8.59
9	8.61	8.51	8.54	8.59	8.52	8.59	8.57	8.51	8.56	532.91	8.65	8.57	8.64	8.57	8.64	8.59
10	8.64	8.50	8.54	8.55	8.64	8.63	8.61	8.68	8.63	8.61	524.43	6.87	6.85	6.84	6.85	8.64
11	8.56	8.59	8.56	8.60	8.54	8.64	8.59	8.62	8.63	8.69	6.84	530.74	6.86	6.89	6.93	8.63
12	8.58	8.65	8.62	8.67	8.67	8.61	8.64	8.63	8.65	8.65	6.85	6.84	529.54	6.86	6.89	8.62
13	8.61	8.60	8.67	8.58	8.61	8.63	8.66	8.67	8.68	8.65	6.90	6.85	6.87	530.82	6.88	8.63
14	8.54	8.60	8.67	8.63	8.62	8.64	8.64	8.66	8.62	8.63	6.82	6.87	6.84	6.86	533.03	8.63
15	8.55	8.58	8.62	8.68	8.67	8.67	8.66	8.61	8.64	8.75	8.64	8.62	8.60	8.66	8.58	532.03

LIQID

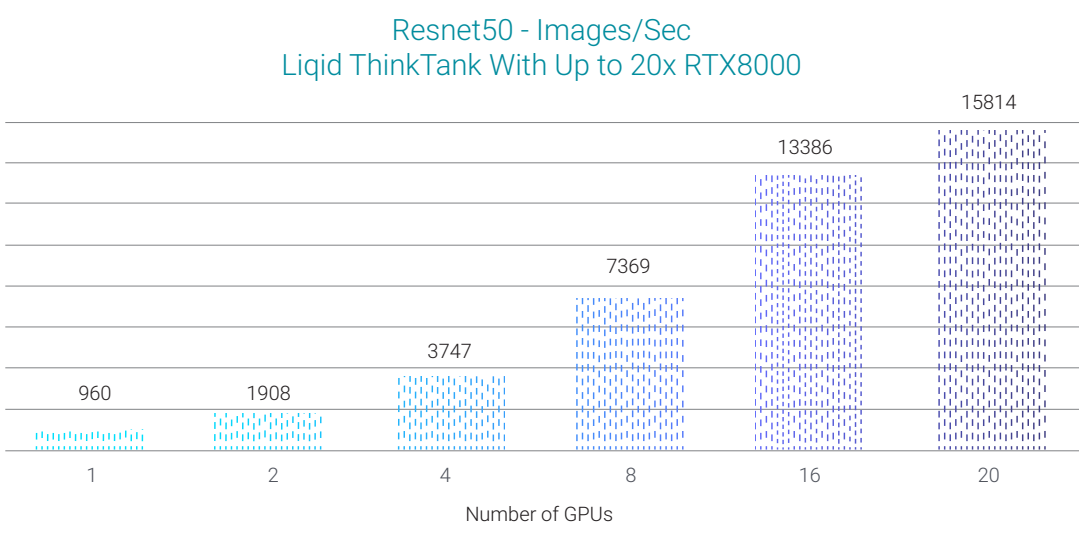
orange™

— POWERED BY —
DELL Technologies

Resnet Performance Testing Summary:

The Resnet50 benchmark is often used to measure performance capabilities of GPU-based systems. The below data was measured on a TensorFlow Resnet50 benchmark using up to 20x NVIDIA Quadro RTX 8000 GPUs connected to a Dell EMC PowerEdge R640 node using Liqid Command Center and a PCIe fabric. For testing, a TensorFlow batch size of 1024 was used and up to 20x NVIDIA RTX 8000 GPUs were composed and the Liqid ThinkTank system was able to achieve an image training throughput of over 15K images per second.

Below are the test results of the benchmark Resnet50 test as measured on the Liqid ThinkTank composable system with up to 20x NVIDIA Quadro RTX 8000 GPU.



Resnet50/ImageNet

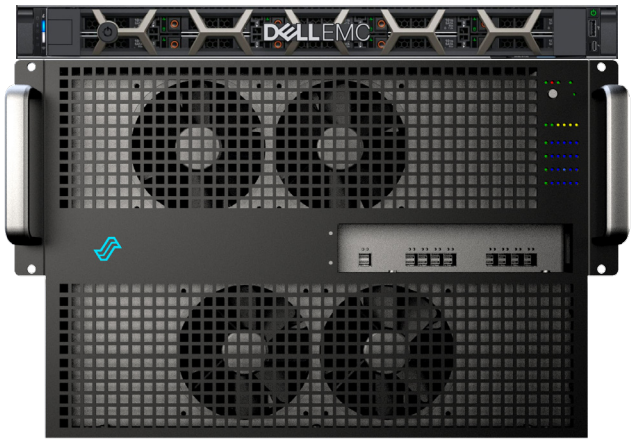
# of GPUs	Images/second
1	960
2	1908
4	3747
8	7369
16	13386
20	15814

Testing Conclusions:

The NVIDIA Quadro RTX 8000 GPUs performed very well in the Liqid ThinkTank composable GPU system. At the time of this writing, this is one of the fastest recorded sets of results for the NVIDIA Quadro RTX 8000 GPU. From a GPU compute perspective, this is a very capable system for training large image classification models. Peer-to-peer functionality was validated with expected performance results. Additionally, no significant storage optimization was performed which could, theoretically, improve the tested Resnet50 benchmark output numbers slightly higher.

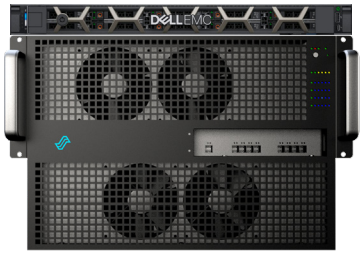
Test Set-up:

Liqid ThinkTank LQD8360 Composable GPU System

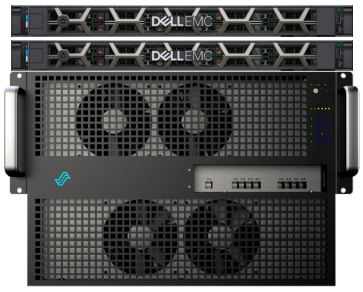


Configuration Overview	
Solution	GPU Supercompute
Compute	1x Nodes Dell EMC PowerEdge R640
GPU	Up to 20x NVIDIA Quadro RTX 8000
Storage	60TB NVMe SSD
Networking	4x 100Gb/s (IB or Eth)
Video	Yes – Video Out Available
Architecture	Disaggregated Composable
Multi-Node	Yes

Multi-Node Configurations Optional:



Up to 20xGPU – 1xCPU
L8360-001GSP-B30 (single node)



Up to 20xGPU – 2xCPU
L8360-002GSP-B30 (dual node)

Up to 20xGPU – 4xCPU
L8360-004GSP-B30 (quad node)

LIQID

orange™

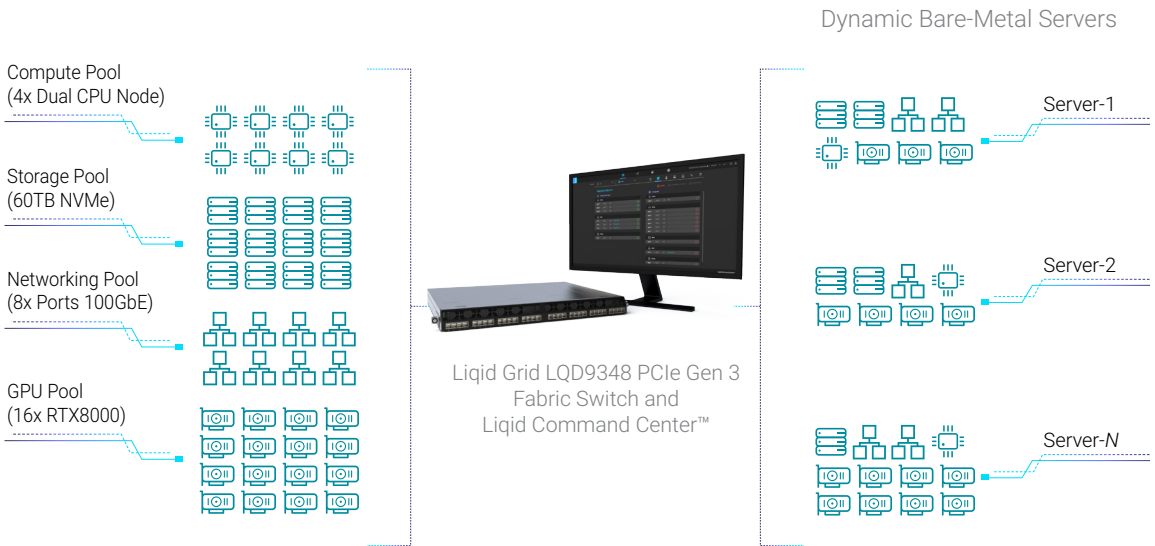
— POWERED BY —
DELL Technologies

Composability:

The Liqid ThinkTank is a fully dynamic and composable system. The users can utilize a system which dynamically scales GPUs and compute nodes as required by the workload and application. Liqid's composable solutions reduce the cost of deployment by optimizing the ratio of GPUs to CPUs and dynamically changing such ratios as needed, significantly improving the total cost of ownership (TCO) for high-density computing environments. The composable model enables GPUs to be incorporated into compute nodes on the fly to take maximum advantage of these powerful compute accelerators through software-defined technology.

The Liqid ThinkTank composable GPU system comes in multiple-node configuration options: Single Node, Dual Node, or Quad Node. Pooled GPUs can be composed to any of the nodes deployed on the fabric. Users can change the ratio between GPUs and CPUs to best meet the application requirements. Liqid bare-metal composable solutions enable the Liqid ThinkTank system to connect up to 20x NVIDIA Quadro RTX 8000 GPUs to a standard 1U Dell EMC PowerEdge R640 compute node, which is otherwise impossible.

Any-to-Any Composability



LIQID

orange™

— POWERED BY —
DELL Technologies

Summary:

The Liqid ThinkTank composable GPU system is able to deliver massive performance for visually intensive workloads with multiple NVIDIA Quadro RTX 8000 GPUs deployed with off-the-shelf server components. This is due to Liqid's ability to leverage an ultra- low latency PCIe fabric, in conjunction with Liqid Command Center software, and compose large numbers of GPUs to a standard high performance 1U compute node, delivering industry-leading GPU compute performance.

In addition to the ultra-performance capabilities of the Liqid ThinkTank composable GPU system, Liqid's software allows for GPUs to be composed in the quantities appropriate for the workload at hand. This means a task that requires only four GPUs can be assigned exactly four GPUs, and the remainder of the devices can be deployed for other applications via Liqid Command Center software.

As artificial intelligence permeates more and more industries at an increasingly granular level, systems such as the Liqid ThinkTank composable GPU system provide a way for organizations to benefit from improved compute performance and data analytics. By better democratizing GPU deployment, companies and research entities can take better advantage of advancements in computing to accelerate and automate processes previously impossible to deploy without complex systems and specialized methods.

While an off-the-shelf, disaggregated solution may not satisfy the needs of all organizations, the price-to-performance ratio aligns for most users who want to take advantage of GPU-driven computing for next-generation applications.

The Liqid ThinkTank composable GPU system is production ready. Get a quote today and incorporate the powerful composable GPU platform into your data system center.

LIQID

orange™

— POWERED BY —
DELL Technologies

About Liqid

Liqid provides the world's most-comprehensive software-defined composable infrastructure platform. The Liqid Composable platform empowers users to manage, scale, and configure physical, bare-metal server systems in seconds and then reallocate core data center devices on-demand as workflows and business needs evolve. Liqid Command Center software enables users to dynamically right size their IT resources on the fly. For more information, contact info@liqid.com or visit