CASE STUDY

Cloud Solutions Provider Accelerates GPU Deployment to Meet Surging High-Density Workload Demand

Csquare delivers advanced networking, cooling, and power solutions to support next-gen NVIDIA H100 GPU capacity and regional expansion.

Background

This cloud tech powerhouse specializing in GPU workloads faced mounting pressure to scale infrastructure in response to surging demand for AI and machine learning capabilities. The company needed to expand its regional footprint and deliver high-performance compute for customers without compromising speed, reliability, or the customer experience.

Challenge

Our customer was racing to deploy NVIDIA H100 GPUs to support its customer's training and inference workloads in a specific region. The aggressive timeline required advanced networking, cooling, and power solutions to build a scalable, resilient environment that met the highest demands in AI processing while maintaining operational continuity.

Solution

Csquare's support for this customer's next-gen infrastructure upgrade included:

- **Collaborative deployment model** designed by our engineers and the customer's technical team.
- **Advanced cooling and thermal management** engineered for high-density GPU environments, optimal performance, and system longevity.
- **Infiniband networking architecture** enable ultra-low latency and high-throughput data transfer critical for real-time AI model performance.
- **100% power availability SLA** to ensure uninterrupted operations for mission-critical workloads through a resilient power strategy.
- **High-density, AI-optimized infrastructure** was purpose-built to support environment to support training and inference workloads with NVIDIA H100 GPUs.

Results

Our customer successfully met its customer's needs while also building a foundation for continued long-term growth in the GPU cloud market through:

- **On-time deployment** in a high-demand region, meeting customer SLAs.
- **Revenue acceleration** with more service offerings and GPU capacity.
- **Scalable foundation** for future AI workloads, enabling rapid respond to evolving market needs.
- **Operational excellence** recognized through seamless execution of complex networking and cooling requirements.
- **Customer satisfaction** due to reliable performance and rapid delivery.

SNAPSHOT

Industry
Cloud Computing

Use Cases
- High-performance GPU deployment
- AI training and inference workloads
- Regional infrastructure scaling

Objectives
- Rapidly deploy NVIDIA H100 GPUs to meet AI workloads demand
- Ensure high-reliability, low-latency, scalable infrastructure
- Meet aggressive timelines without compromising performance

Results
- Seamless execution of complex infrastructure
- On-time deployment to support high-density training and inference compute
- Expanded GPU capacity and service offerings
- Scalable foundation for future AI workloads
- Increased customer satisfaction and retention

About Csquare

Csquare is a preeminent colocation provider with a proven track record of providing highly reliable data center services. Prioritizing security, reliability and superior customer service, Csquare offers flexible data center solutions backed by a 100% uptime guarantee for thousands of diverse organizations across all industries.

Learn more at
csquare.com

© 2025 Csquare. All rights reserved.

CSQUARE