

AI Resilience: Delivering Mission-Critical AI Services with ZPE's Out-of-Band Management Platform

Maximizing the Value of AI Infrastructure

Discover how Gruve, a global AI services leader, built a resilient, multi-tenant AI infrastructure that stays online around the clock. By leveraging ZPE Systems' Nodegrid platform, Gruve gained complete visibility, fast response times, and simplified compliance, while cutting the cost and effort of managing mission-critical workloads.

Here's the story of how their innovative approach keeps high-stakes AI services running without missing a beat.

Unlike many service organizations that focus on technology first, Gruve works closely with each customer to identify business, financial, and technical objectives, then tailors solutions to deliver measurable outcomes, such as reducing costs, increasing market value, or improving operational efficiency. To accelerate solution development, Gruve has invested heavily in state-of-the-art infrastructure, including GPU clusters, high-speed cluster networks, and multiple flash storage platforms.

In addition to AI platforms, Gruve operates internally, mission-critical data center services that their operations center depends on 24/7. With strict Service Level Agreements (SLAs) in place for production customers, even minutes of downtime can have a big impact on business.

To meet these demands, Gruve's operations and IT teams require fast, reliable out-of-band (OOB) management to make rapid changes with minimal human intervention, ensuring continuous availability for AI workloads and essential services alike.

About Gruve.ai

Gruve is a global AI services company headquartered in Redwood City, California, with offices in New Jersey, Texas, India, South Korea, Japan, Singapore, and Dubai. They help enterprises design, build, and manage AI solutions through an outcome-based engagement model. Their portfolio spans AI Software and Data Sciences, Cybersecurity, AI Networks, Virtualization, and Customer Experience platforms.

CHALLENGES

- Preventing GPU failures & disruptions
- Protecting AI training workloads
- Maintaining compliance

SOLUTIONS

- Full out-of-band remote access
- Holistic cloud management
- Logical & physical isolation

OUTCOMES

- **Infrastructure Protection** – Real-time monitoring & alerts prevent GPU damage & failures.
- **Fast Response** – ZPE Cloud enables instant visibility & problem resolution.
- **Simple Compliance** – Full isolation ensures ISO 27001 & SOC 2 compliance.

The Challenge

Designing and operating multi-tenant AI Clusters at global scale comes with unique demands. Gruve requires a platform that can protect their high-value GPU investments, maintain uninterrupted service for customers, and simplify compliance across multiple industries, all while keeping operational costs in check.

Gruve's initial design phase uncovered several challenges:

- **Investment Protection:** High-performance GPUs are expensive assets. Overheating or throttling can reduce a GPU's lifespan by up to 30% or cause severe, unexpected shutdowns. Normal solutions can take hours to deliver critical power data, but Gruve requires proactive environmental monitoring that can detect temperature spikes in real time and notify teams before there's any service impact.
- **Remote Connectivity:** At 800G network speeds, port flapping or link failures can disrupt AI training and inference workloads. Out-of-band remote access can address this, but traditional solutions are limited by latency and network congestion, forcing engineers to go on-site. Gruve requires full OOB access that allows them to isolate, diagnose, and fix problems before they can cause training disruptions.
- **Service Isolation:** Supporting multiple customer workloads means having strong tenancy controls to separate infrastructure. But with normal solutions or data center services, the only way to achieve this is through physical separation, which costs more and requires more hardware. Gruve requires full infrastructure isolation and visibility, including network access and lights-out management (LOM) connections, to ensure each serviced owner maintains governance over their data and maintenance requirements.
- **Fragmented Management:** Gruve's AI Clusters are set up in colocation data centers that need to be managed remotely. However, traditional solutions require multiple management portals, which creates inconsistent operations and can prolong remediation. Gruve requires a single, secure portal accessible via single sign-on (SSO), that gives them a holistic view for managing compute, storage, and network resources across all sites.

The Gruve operations team must deploy compute, network, and storage infrastructure while ensuring full management connectivity and thorough documentation. With teams launching new services daily, they need complete visibility into every component, clear ownership across GPU clusters, and detailed audit logs to meet ISO27001 and SOC 2 compliance. This requires mapping every device, testing each connection, and validating standard operating procedures to maintain consistent, secure operations.

Solution

Gruve evaluated several out-of-band (OOB) management platforms, covering everything from performance and scalability, to cloud management, logging, and security. Gruve identified ZPE Systems' Nodegrid as the only solution capable of meeting all operational, compliance, and scalability requirements, while also streamlining everyday management tasks.

ZPE's cloud manageability stood out as a major differentiator. With ZPE Cloud, the operations team could securely manage all of their AI Clusters through a single portal, with single sign-on (SSO) at both the cloud and device levels to enforce security standards and maintain centralized control. This eliminated the need for multiple management systems and site-specific policies that often slowed response times.

Gruve's solution included:

Nodegrid Net Services Routers (NSR) – Provided full OOB access to servers, storage, and network management ports, ensuring remote remediation was possible without depending on the production network.

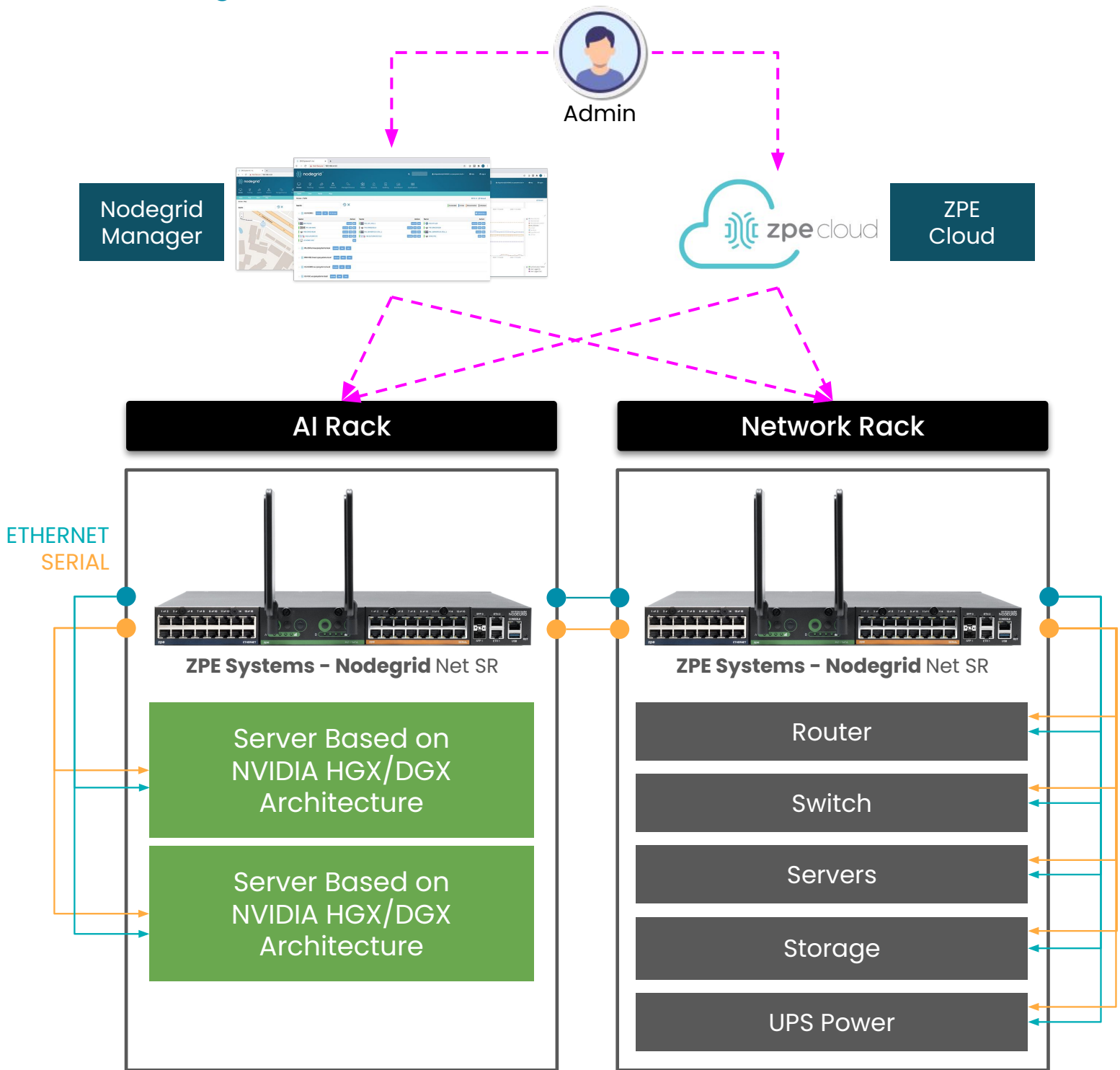
Nodegrid Gate SR (GSR) appliances – Offered cellular failover connectivity that guaranteed access even if primary links were unavailable.

Environmental Monitoring Sensors – Monitored temperature and power consumption in real time to prevent GPU overheating, avoid power-related failures, and extend the life of high-value hardware assets.

Integrated Connectivity Modules – Allowed Gruve to expand the NSR's serial and Ethernet management connections without purchasing additional devices, which reduced hardware sprawl and simplified rack space planning.

This architecture enabled the Gruve operations team to rapidly install and configure compute, network, and storage infrastructure, while keeping detailed audit logs for ISO 27001 and SOC 2 compliance. By mapping every device to its corresponding NSR, testing all connections, and automating SOP adherence, Gruve achieved fast deployments, consistent operations, and strong resilience across their AI service portfolio.

Network Diagram



With **ZPE Systems' Nodegrid Net SR**, users can utilize both **SERIAL** and/or **ETHERNET** connectivity to access, manage, and troubleshoot AI Clusters, all from a single pane of glass. Access devices through **Nodegrid Manager (On-prem)** or via **ZPE Cloud**.

Results and Benefits

By implementing ZPE Systems' Nodegrid platform, Gruve reduced operational overhead while increasing responsiveness, compliance, and system resilience of their mission-critical AI Infrastructure Clusters.

Key outcomes include:

Faster Response SLAs – Engineers can securely log in through ZPE Cloud to diagnose and resolve connectivity or hardware issues in minutes, without traveling on-site. This has significantly reduced mean time to resolution and avoided SLA penalties.

Simplified Compliance & Maintenance – Logical and physical tenant separation, combined with centralized policy enforcement, ensures ISO 27001 and SOC 2 requirements are met while minimizing service disruption during maintenance.

Optimized Resource Allocation – Remote upgrades, reconfigurations, and troubleshooting allow IT staff to focus on revenue-generating AI initiatives instead of routine maintenance visits.

Improved Infrastructure Protection – Real-time monitoring of temperature, power, and device health helps protect high-value GPU assets from performance degradation and premature failure.

"We rely on ZPE Systems' Nodegrid to help us leverage the value of our AI Cluster investments. The Nodegrid platform gives us full visibility and adaptability as we build new AI solutions for customers and partners."



–Matt Robinson, CTO, Gruve

ZPE Systems' Nodegrid platform is a cornerstone of Gruve's service delivery model, bringing together unified management, proactive infrastructure protection, and the flexibility to scale across a global footprint.

With full visibility and control from anywhere, Gruve can meet aggressive SLA commitments, protect high-value GPU assets, and keep mission-critical AI workloads running without interruption.

By reducing the time and cost of managing complex infrastructure, Nodegrid enables Gruve's engineers and architects to focus on what matters most: delivering innovative AI solutions that drive measurable business outcomes for their customers.

See how ZPE Systems can help your organization achieve the same operational agility and resilience.

Contact us today or request a demo to explore the full capabilities of Nodegrid for your AI and mission-critical workloads. Email Sales@zpesystems.com for more information.

