



Accelerating Sovereign AI Infrastructure with Secure Neo-Cloud Connectivity

The combined Gcore and Graphiant architecture enables enterprises to operationalize AI infrastructure faster while simplifying networking, security, and scalability.

Executive Summary

As enterprises move AI from experimentation into production, they face a new infrastructure challenge. AI applications, models, and data are no longer confined to a single cloud or data center. Instead, they are distributed across multiple public clouds, on-premises environments, edge locations, and geographic regions.

This distributed architecture creates significant operational challenges. Organizations are left managing fragmented infrastructure and disconnected networks, while incurring the high cost of traditional public cloud resources. At the same time, they often lack visibility and control over where enterprise data travels, making it difficult to satisfy data sovereignty requirements, maintain regulatory compliance, and protect sensitive AI workloads.

Graphiant and Gcore provide a unified approach to solving these challenges by delivering an enterprise-ready AI infrastructure that combines high-performance AI computing with secure, intelligent networking. Together, the solution provides a more cost-effective and operationally simpler foundation for production AI than traditional public cloud-only architectures. It enables organizations to connect existing data sources across public clouds, private data centers, and edge environments while maintaining complete control over how and where enterprise data moves, helping organizations meet compliance requirements and build truly sovereign AI infrastructure.

Gcore delivers scalable GPU infrastructure and AI deployment platforms designed to simplify enterprise AI operations through high-performance AI cloud services and distributed serverless inference capabilities with AI Grid. Graphiant provides secure, cloud-native network connectivity that unifies public cloud, neo-cloud, edge, and enterprise infrastructure into a single operational fabric, eliminating fragmented connectivity and securely connecting enterprise data sources to AI workloads regardless of where they reside.

Together, Graphiant and Gcore enable enterprises to securely connect, move, process, and operationalize AI workloads globally while maintaining performance, governance, data sovereignty, and operational simplicity.

The AI Infrastructure Problem

AI is creating an entirely new class of infrastructure requirements. As organizations move AI into production, they must support increasingly distributed applications, data, and compute resources while delivering the performance, security, and governance expected of enterprise environments.

Modern AI infrastructure requires:

- **High-Performance GPU Compute and Real-Time Inference**
AI workloads demand access to scalable GPU infrastructure capable of supporting both large-scale model training and low-latency inference for real-time AI applications.
- **High-Throughput, Low-Latency Data Connectivity**
AI models depend on moving large volumes of data quickly and efficiently between enterprise data sources, cloud environments, and AI infrastructure without introducing performance bottlenecks.
- **Multi-Cloud Interoperability**
Enterprises need the flexibility to leverage multiple public clouds, neo-cloud providers, private infrastructure, and edge environments without creating isolated operational silos.
- **Compliance and Sovereign Data Controls**
Organizations must maintain control over where sensitive data is stored and transmitted to satisfy regulatory requirements, meet data sovereignty mandates, and protect AI workloads from unauthorized access.

However, enterprises often struggle with fragmented infrastructure architectures that require separate networking models, provisioning processes, and security frameworks for each cloud or AI provider. At the same time, GPU demand continues to outpace supply, driving organizations toward neo-cloud providers optimized specifically for AI workloads and GPU availability. This creates several major challenges:

- **Complex Connectivity Between Enterprise Environments and AI Clouds**
Connecting data across multiple public clouds, neo-cloud providers, on-premises environments, and edge locations often requires multiple networking technologies and manual integration, increasing complexity and slowing deployments.
- **High Cost of Data Movement and Global AI Scale**
Moving large AI datasets between cloud providers, geographic regions, and enterprise environments generates significant networking and cloud egress costs. As AI deployments expand globally, these costs continue to grow while organizations face increasing complexity in extending AI infrastructure across regions and providers.
- **Security, Compliance, and Data Sovereignty Risks**
Distributed AI architectures make it more difficult to enforce consistent security policies, maintain visibility into data flows, satisfy regulatory requirements, and ensure sensitive enterprise data remains within approved geographic boundaries.
- **Operational Overhead from Managing Multiple Environments**
Separate networking, security, and operational models for each cloud and AI platform increase administrative effort, create inconsistent processes, and reduce operational efficiency.

Why Neo-Clouds Matter

Neo-cloud providers are emerging as a critical layer in the AI ecosystem.

Unlike traditional hyperscalers, neo-clouds are purpose-built for AI workloads and optimized for GPU access, AI model training, and inference operations.

Gcore's AI infrastructure portfolio includes:

- GPU Cloud with H100/H200/B300/GB300 GPU infrastructure
- AI Cloud Stack for AI service providers
- Everywhere AI for 3-click AI deployment and inference
- CDN-integrated AI Grid routing, supporting latency, cache awareness, and GPU Load balancing
- Hybrid and air-gapped deployment models

Gcore supports organizations requiring:

- Faster AI Deployment
- Lower GPU infrastructure costs
- Real-time inference
- Data sovereignty
- Secure AI environments
- Global AI scalability

The Connectivity Gap in AI Infrastructure

While AI infrastructure innovation has accelerated, networking architectures have lagged. Enterprises adopting AI quickly discover that deploying AI infrastructure extends well beyond provisioning GPU resources. As AI environments become increasingly distributed, organizations face new infrastructure challenges. Provisioning can be slow, and networking across multiple clouds and AI providers adds significant complexity. Many deployments rely on the public internet, creating security and performance concerns. At the same time, maintaining consistent security policies and integrating neo-cloud providers into existing enterprise networks becomes increasingly difficult.

As a result, AI infrastructure is no longer simply a compute challenge. It is fundamentally a data movement and connectivity challenge. The ability to securely connect applications, data, users, and AI services across diverse environments has become critical to operationalizing AI at enterprise scale.

Graphiant addresses these challenges with a private network fabric that unifies enterprise infrastructure and AI environments. Organizations can connect AI infrastructure to existing data sources without redesigning their networks, they can also do so without deploying additional networking hardware. This creates a secure connectivity layer that scales as AI deployments grow enabling AI workloads and enterprise data to communicate seamlessly across cloud, on-premises, and edge environments.

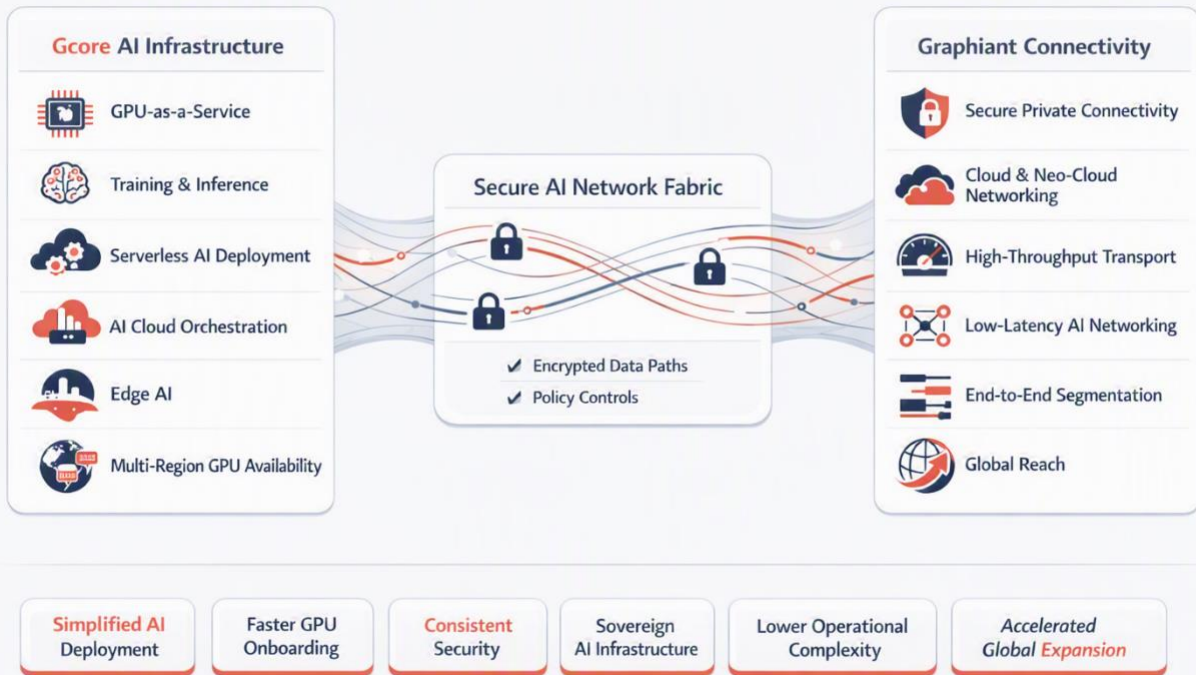
Joint Solution Overview: Graphiant + Gcore

The combined Graphiant and Gcore architecture enables enterprises to operationalize AI infrastructure faster while simplifying networking, security, and scalability.

By combining AI compute with intelligent connectivity, the joint solution enables organizations to build AI infrastructure that is simpler to deploy, more cost-effective to operate, and easier to scale than traditional public cloud-only architectures. Enterprises can securely connect existing data sources to AI workloads while maintaining policy-driven control over how and where enterprise data moves. This helps organizations satisfy compliance requirements, enforce data sovereignty policies, and operationalize AI with confidence.

The joint solution can be understood as a secure AI infrastructure architecture that combines Gcore's distributed AI compute capabilities with Graphiant's private connectivity fabric and policy-driven control.

Joint Solution Overview: Graphiant + Gcore



Use Case: SaaS AI Workloads

Modern AI-driven creative applications require massive GPU resources, rapid data movement, and globally distributed inference environments.

A deployment model for cloud-native AI-driven platforms (like cloud-based graphic design platforms) may involve:

- AI image generation
- Large-scale inference workloads
- Global user traffic
- Multi-region GPU utilization
- High-throughput asset movement
- Secure processing pipelines

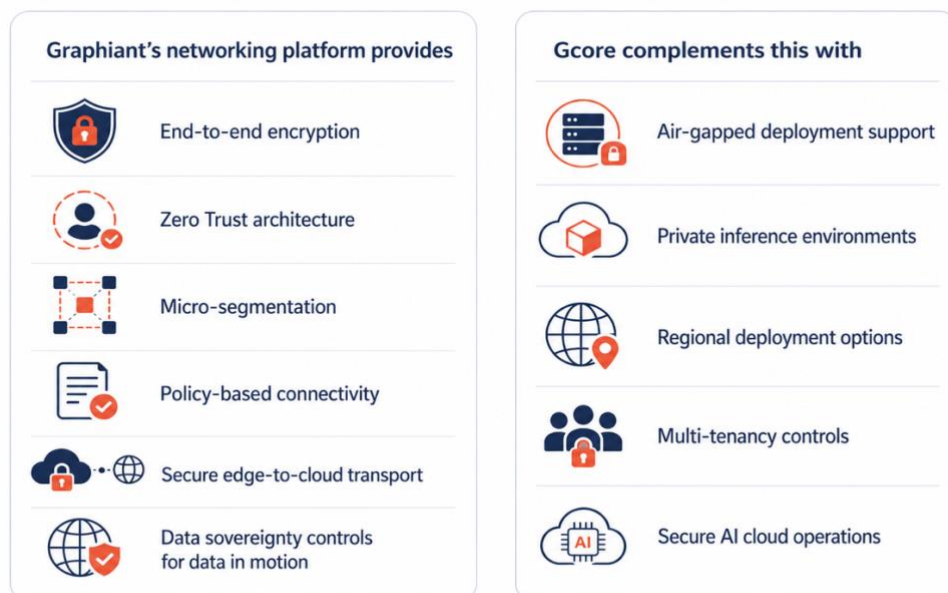
In these environments Gcore provides scalable GPU compute and AI inference capabilities while Graphiant provides secure, performant connectivity between enterprise infrastructure, cloud environments, and neo-cloud GPU services

This allows organizations to:

- Avoid public internet bottlenecks
- Improve AI response times
- Maintain secure segmentation policies
- Scale globally without redesigning infrastructure
- Operate production AI workloads cost-effectively

Security and Sovereignty for AI

AI workloads increasingly involve sensitive enterprise and customer data. For many enterprises and government agencies, controlling where sensitive data is processed and where it travels is becoming a critical requirement. AI workloads supporting regulated, classified, or mission-critical operations often require compute infrastructure that can be deployed within approved jurisdictions, while ensuring enterprise data never leaves authorized sovereign boundaries. This requires a solution that combines sovereign AI compute with policy-driven control over data movement, enabling organizations to meet regulatory and national security requirements without compromising performance or operational simplicity.



Together, the combined architecture enables sovereign AI infrastructure that balances performance, compliance, and operational simplicity.

Key Business Outcomes

Organizations leveraging Graphiant and Gcore together can achieve:

- **Faster AI Deployment.** Rapid onboarding into GPU infrastructure and AI environments without lengthy network provisioning.
 - **Lower Infrastructure Complexity.** Unified connectivity and AI operations reduce operational overhead across environments.
 - **Improved AI Performance.** Low-latency private connectivity supports real-time inference and large-scale data movement.
 - **Enhanced Security.** Consistent segmentation and Zero Trust architecture protect sensitive AI workloads.
 - **Global Scalability.** Expand AI infrastructure across regions without redesigning networking architectures.
 - **Better Cost Efficiency.** Optimize GPU utilization and reduce expensive cloud networking overhead.
 - **Data Sovereignty and Compliance.** Maintain control over where AI workloads are processed and where enterprise data travels to satisfy regulatory, contractual, and sovereign data requirements.
-

Target Industries

The combined Graphiant + Gcore solution is especially valuable for industries with:

- AI-intensive workloads
- Regulatory requirements
- Latency-sensitive applications
- Distributed infrastructure

This includes the following industries:

- Financial Services
- Healthcare
- Public Sector/Government Agencies
- Telecommunications
- Retail
- AI-native SaaS platforms

Conclusion

AI infrastructure is evolving beyond traditional cloud architectures.

Enterprises now require:

- Flexible GPU access
- Secure multi-environment connectivity
- Sovereign AI operations
- Real-time performance
- Simplified scalability

Gcore and Graphiant together help organizations bridge the gap between AI compute infrastructure and enterprise networking.

By combining Gcore's AI cloud and GPU platform capabilities with Graphiant's secure network fabric, enterprises can accelerate AI adoption while maintaining control, security, and operational efficiency.

The future of AI infrastructure will depend not only on compute power but on the ability to securely move and operationalize data across distributed environments at global scale.

That future requires both AI infrastructure and intelligent connectivity working together.



Assured, Agile, Awesome.

Graphiant delivers a single, unified Network-as-a-Service platform built for the AI era. By combining a stateless core architecture with end-to-end encryption, real-time observability, and policy-driven routing, Graphiant helps enterprises and service providers connect users, sites, clouds, and AI workloads with the speed, security, and agility modern business demands.

[Learn more at graphiant.com](https://graphiant.com)

Copyright © 2026 Graphiant Inc. All Rights Reserved