SOLUTION BRIEF

Fortinet High-Performance Security for AI-Driven Data Centers

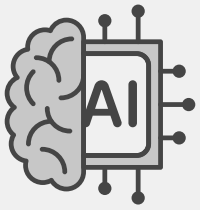
Securing the Performance, Trust, and Resilience of Modern AI Infrastructure

Executive Summary

Data centers are undergoing a rapid architectural shift due to the rapid adoption of AI. Traditional data center environments were built to manage predictable north-south traffic. Today's GPU clusters, training pipelines, and LLM-driven applications, however, generate massive east-west flows that move sensitive data continuously. These workloads require high throughput, low latency, and accurate security inspection.

Legacy data center firewalls struggle to keep pace with the unprecedented scale and velocity of east-west traffic generated by GPU fabrics, the high percentage of encrypted internal flows, and the unique characteristics of AI workloads that traditional inspection engines cannot reliably classify or analyze. As a result, they cannot consistently detect or prevent threats targeting AI and LLM systems.

Fortinet's AI Data Center Firewall Solution addresses this challenge with ASIC-accelerated inspection, LLM-aware threat protection, and unified governance across every layer of the AI stack. Anchored by FortiAIGate and integrated with the Fortinet Security Fabric, it protects runtime AI traffic, enforces zero-trust access, and maintains performance across dense GPU fabrics. The result is an AI infrastructure that remains fast, compliant, and secure.



"Generative AI training and inference place unprecedented pressure on data center fabrics, driving the need for higher-bandwidth, lower-latency interconnects and redesigned east-west architectures."[1]

AI Data Centers Face a New Security Reality

The shift to high-value, high-velocity AI workloads

Modern data centers are rapidly evolving into AI compute fabrics. LLM training, fine-tuning, vector search, and inference pipelines produce enormous lateral traffic as models exchange embeddings, parameters, and data across interconnected GPU nodes. Unlike traditional applications, these workloads constantly interact with dynamic and sensitive information.

New risks emerge inside the data center

This traffic pattern introduces threats that perimeter controls cannot reach. Sensitive data can move laterally without oversight, internal APIs can be abused or manipulated, and AI models themselves can become targets for prompt injection or poisoning. In addition, new compliance frameworks, such as the EU AI Act, are also increasing requirements for visibility, auditability, and data governance.

Traditional security tools fall short

Conventional firewalls were not designed for how AI data centers operate today. The volume and speed of east-west traffic generated by GPU clusters far exceed what traditional inspection pipelines can process without introducing latency or breaking workload performance. And much of this traffic is encrypted, which further limits visibility unless the firewall has specialized acceleration that can decrypt, analyze, and re-encrypt flows at scale.

Legacy devices also lack awareness of how LLMs and AI agents behave. They cannot distinguish normal model-to-model exchanges from manipulation attempts, poisoning, or data leakage in inference responses. As a result, they provide only partial protection for the systems on which AI workloads depend.

Today's AI data centers require a purpose-built approach—one that can secure internal communication at terabit speeds, apply zero-trust controls across microservices and GPU fabrics, and analyze AI-specific behaviors in real time. This is the foundation needed to keep modern AI environments resilient and compliant.

Fortinet AI Data Center Firewall Solution

Fortinet delivers the first fully unified, AI-aware security architecture designed for the scale and speed of today's data center AI workloads. At the center is FortiAIGate, an AI-powered solution designed to inspect, validate, and govern LLM traffic. Combined with FortiGate, FortiWeb, FortiPAM, FortiDLP, and FortiData, and embedded across the Fortinet Security Fabric, the FortiAIGate solution forms an integrated mesh of critical network, application, and data protections.

Inline security for GPU-dense environments

Fortinet security processing units (SPUs) provide ASIC-accelerated performance, enabling inline inspection at terabit scale with negligible latency. This ensures AI workload protection without interrupting training or inference pipelines.

LLM-aware security controls

FortiAIGate analyzes and filters AI-specific content in real time, including:

- Detection of prompt injection and model manipulation
- Prevention of model poisoning and jailbreak attempts
- Sanitization of inputs and outputs
- Guardrail enforcement for compliance and governance

Unified protection across AI infrastructure

The solution also integrates key components of the Fortinet Security Fabric:

- FortiAIGate for AI and LLM traffic inspection
- FortiGate for perimeter and internal segmentation
- FortiWeb for API and application security
- FortiPAM for privileged access and session auditing
- FortiDLP and FortiData for data protection and policy enforcement

Continuously updated AI threat intelligence

FortiGuard AI-Powered Security Services ingests global telemetry from hundreds of millions of sensors and applies advanced machine learning models to identify emerging AI-focused attacks. This includes new forms of prompt manipulation, model evasion techniques, data extraction attempts, and exploitation patterns targeting LLM-driven applications. Threat insights are continuously refined and automatically propagated across the Security Fabric, providing organizations with real-time protection without manual tuning or an added operational burden. This ensures AI workloads remain protected as adversaries adapt their tactics and new forms of AI-generated threats emerge.

Outcomes for AI Data Centers

Fortinet's unique architecture enables operators to protect both the infrastructure and the intelligence running on top of it.

Stronger protection of proprietary AI assets

Integrated zero-trust controls ensure that only authorized users, systems, and services can interact with high-value models and data sets. This prevents unauthorized access and protects intellectual property.

Reduced east-west exposure

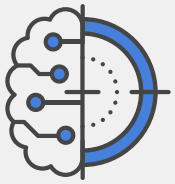
Inline segmentation stops malicious lateral movement across GPU nodes, microservices, and training pipelines without impacting performance.

Reliable and trustworthy model operations

LLM-aware inspection prevents prompt manipulation, poisoning attempts, and data leakage, helping maintain the integrity and stability of AI systems.

Simplified compliance and governance

Built-in data classification, audit logs, and usage guardrails support alignment with frameworks such as the EU AI Act, NIST AI RMF, GDPR, and HIPAA.



"More than 80% of AI-related security incidents involve unauthorized access, data leakage, or misuse of training datasets."[2]

Operational Impact for AI Data Centers

Stronger protection of proprietary AI models and datasets: Organizations can apply strict zero-trust access controls across GPU clusters, API endpoints, and internal services to ensure only authorized entities can interact with high-value AI assets. Inline inspection and AI-aware threat detection prevent manipulation attempts, data harvesting, and unauthorized access to models. This keeps sensitive training data and proprietary model logic secure throughout the AI life cycle.

Performance-preserving inspection for GPU fabrics: ASIC-accelerated security processing enables real-time inspection of lateral AI traffic without introducing latency that could disrupt training or inference workloads. And GPU clusters maintain optimal throughput even under full security enforcement. This allows operators to protect east-west flows at scale without sacrificing runtime performance.

End-to-end visibility across hybrid and cloud AI workloads: Unified telemetry from on-premises data centers, private clouds, and public cloud AI services provides a complete view of how models, data, and APIs communicate. Security teams can trace activity across containers, microservices, and distributed pipelines with consistent policies. This reduces blind spots and supports continuous monitoring across the full AI ecosystem.

Simplified policy management through the Security Fabric: A single management plane allows administrators to create, deploy, and enforce policies across all AI infrastructure components. Integrated visibility and shared context eliminate the need to maintain fragmented rule sets across multiple point products. This reduces operational complexity and enables more consistent security outcomes.

Faster time-to-compliance with integrated governance tools: Automated reporting and real-time monitoring reduce the efforts required to demonstrate compliance during audits. This accelerates readiness while lowering the administrative overhead normally required for AI governance.

A Security Model Built for AI

AI workloads require a security model built for high throughput, low latency, and LLM-specific risks. Fortinet's AI Data Center Firewall Solution provides an integrated, ASIC-accelerated approach that protects east-west traffic, secures runtime model interactions, and unifies governance across the entire data lifecycle. It gives AI data center operators the confidence to scale innovation without compromising performance or trust.

[1] AvidThink. NGI 2025: Data-Center Networking in the Era of AI and Cloud. AvidThink, 2025. https://arrcus-admin.prod.unomena.io/media/documents/AvidThnik_NGI_2025_DataCenter_Networking_AI_Cloud.pdf

[2] European Union Agency for Cybersecurity (ENISA). AI Threat Landscape 2024: Security Risks and Challenges. ENISA, 2024. https://www.enisa.europa.eu/publications/ai-threat-landscape-2024/

December 6, 2025 12:44 AM / MKTG-2530-0-0-EN