

Market Study



#### **Table of Contents**

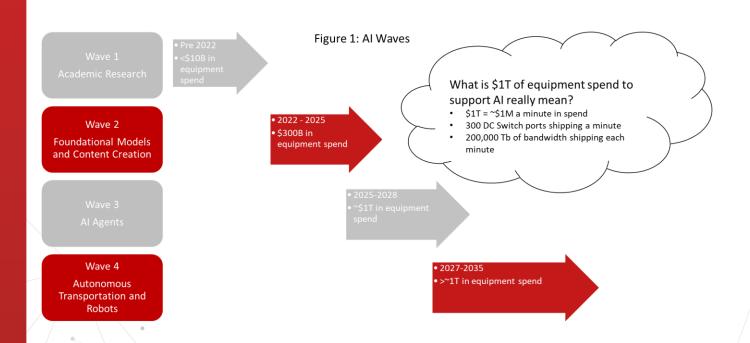
Introduction	1
Enterprise Trends	2
Considerations for Enterprises in the World of Al	2
2026+ Predictions for Enterprise Al Strategy for Data Centers	3
Ethernet Already the Number One Choice for Enterprises	3
Al Networking Vision for the Enterprise	4
Solution and Advantages for Al Networking	4
New Standards Like UEC Enhance Al Networking	5
Unified and Interoperable (Brownfield and Greenfield)	6
Value of Real-time Protection - Zero Down Time or Disruption	6
Backend Infrastructure for Al Networking	6
Frontend Infrastructure for AI Networking	7
Multi-layer Security Integration for all Al Networking Components	7
Best Designs, and Longevity with Investment Protection	8
Partnership Advantage and Uniqueness	9
Conclusion	g



#### Introduction

In today's rapidly evolving digital landscape, artificial intelligence (AI) has become a cornerstone of enterprise innovation, driving efficiencies, enhancing decision-making, and unlocking new revenue streams across industries. Enterprises are increasingly adopting AI to process vast amounts of data, automate complex tasks, and deliver personalized experiences. Still, this shift demands robust AI networking capable of handling unprecedented computational loads. As AI models grow in complexity from foundational large language models to advanced agentic systems, the need for scalable, secure, and efficient data centers has never been more critical. However, many organizations face challenges such as network bottlenecks, security vulnerabilities, and integration hurdles with legacy systems, which can hinder AI deployment and ROI.

This whitepaper explores key trends, considerations, and predictions for enterprise Al strategies, with a focus on Al networking in data centers. By addressing these elements, enterprises can build future-proof infrastructures that support Al-driven growth while minimizing risks and costs.





#### **Enterprise Trends**

Enterprises are witnessing a surge in Al adoption, with Al workloads shifting from experimental pilots to core business operations. Key trends include the proliferation of generative Al applications, edge computing for real-time processing, and hybrid cloud environments that blend on-premises and cloud resources. Data centers are evolving to support distributed Al architectures, where east-west traffic dominates due to inter-GPU communications in training and inference tasks. Additionally, sustainability concerns are rising, as Al's energy demands drive the need for efficient designs that reduce power consumption and carbon footprints.

Another prominent trend is the convergence of networking technologies, with Ethernet emerging as a preferred fabric for its cost-effectiveness and scalability compared to InfiniBand. Organizations are prioritizing interoperability to avoid vendor lock-in, enabling seamless integration of diverse hardware ecosystems. Ethernet also offers the benefit of operational consistency and maturity across frontend and backend networks. This trend gains in importance when an enterprise sits in 2028 and sees multiple GPUs in its installed base and more accelerators in the future. Each Al build cannot be a disparate installation, and the networking fabric plays a key role in bringing multiple Al offerings into the enterprise's disposal.

Call to Action: To align with these trends, enterprises should assess their current data center capabilities through a comprehensive AI readiness audit, identifying gaps in scalability and efficiency to accelerate their AI journey.

### Considerations for Enterprises in the World of Al

In the AI era, enterprises must prioritize low-latency, high-bandwidth networking to support massive data flows between compute nodes, ensuring minimal job completion times for AI tasks. Security is paramount, as AI models handle sensitive data, requiring robust protections against threats like data breaches or model poisoning. Cost management is another key consideration, balancing upfront investments in specialized hardware with long-term operational savings through automation and energy-efficient designs.

Interoperability remains a challenge, particularly in brownfield environments where legacy systems coexist with new Al infrastructures. Enterprises need solutions that offer complete



visibility and telemetry to monitor performance in real-time, preventing disruptions that could derail Al initiatives. Finally, regulatory compliance and ethical Al practices, such as data privacy, must be embedded into infrastructure planning to mitigate risks.

**Call to Action**: Enterprises should engage with AI networking experts to conduct vulnerability assessments and develop tailored roadmaps that integrate security and interoperability, supporting a seamless transition to AI-centric operations.

#### 2026+ Predictions for Enterprise Al Strategy for Data Centers

By 2026, enterprise AI strategies will focus on agentic AI systems that autonomously interact with environments, demanding data centers with ultra-low latency and adaptive networking. There is a shift toward Physical AI, integrating intelligence into real-world applications like robotics, requiring data centers to support exponential data growth and real-time collaboration. Ethernet will dominate as the backbone, offering scalability for thousands of GPUs while reducing costs compared to alternatives.

Sustainability will drive the adoption of power-efficient technologies, including advanced cooling methods and modular designs for longevity. Data centers will increasingly leverage Al-driven management tools for predictive maintenance and anomaly detection, minimizing downtime. Partnerships between networking and compute vendors will become standard, fostering ecosystems that accelerate innovation and investment protection.

**Call to Action:** Forward-thinking enterprises should invest in scalable Ethernet fabrics and partner with ecosystem leaders to future-proof their data centers, ensuring readiness for evolving AI demands.

### Ethernet Already the Number One Choice for Enterprises

InfiniBand has historically been favored for its ultra-low latency and lossless design, making it suitable for high-performance computing. However, Ethernet surpasses it in enterprise settings due to its cost-effectiveness, open standards, and widespread interoperability, which avoid vendor lock-in and enable seamless integration with existing infrastructure. Advancements in effective congestion management with RoCEv2 and the Ultra Ethernet Consortium (UEC) have narrowed performance gaps, delivering comparable bandwidth and latency while supporting



dynamic, scalable environments for Al workloads.

A key challenge with InfiniBand is the lack of skilled human resources, as it requires specialized expertise that is scarce compared to the abundant Ethernet talent pool. Managing multiple networks increases complexity, operational costs, and downtime risks. Ethernet's unified fabric simplifies operations, reduces TCO, and supports brownfield upgrades, making it ideal for enterprises seeking flexibility and longevity in Al data centers.

Call to Action: Enterprises should train their employees to expand their skillset around Ethernet-based solutions to prepare for additional Al workload demand. Unifying their Al infrastructure to overcome HR shortages and multi-network challenges is key.

#### Al Networking Vision for the Enterprise

The evolving vision for AI networking emphasizes a unified, intelligent fabric that enables enterprises to leverage AI while maintaining high performance and embedded security fully. This approach focuses on creating agile, scalable data centers optimized for diverse AI workloads, including model training, inference, and emerging agentic applications. While many vendors will pitch such solutions, only vendors with scale and robust portfolios can keep pace with a three- to five-year journey. With advanced silicon and sophisticated software, networks can dynamically adjust to varying traffic demands, providing lossless data transfer and reduced latency across distributed systems. Even if enterprises are not doing training or delaying the AI journey, it is important to build a high-speed network that can quickly adapt, as it is unlikely that any network built today won't have AI-based traffic during its lifecycle.

At its core, this vision integrates comprehensive visibility, telemetry, and automation to support proactive oversight and sustainable operations. It accommodates both new builds and existing infrastructure upgrades, helping enterprises navigate the Al landscape with lower complexity and improved return on investment from project one today, to the nth project in 2030.

Call to Action: Enterprises starting their Al journey should seek demonstrations of unified management platforms, uncovering ways to simplify Al fabric management and operations not just on today's deployments, but with a robust roadmap for future expansion.

### Solution and Advantages for Al Networking

Advanced Al networking solutions deliver end-to-end platforms for enterprise data centers, featuring high-performance switching and intuitive dashboards. These platforms tackle the



rigorous demands of AI through durable hardware, innovative software, and broad ecosystem compatibility, ensuring efficiency and dependability. Capabilities such as adaptive load balancing, congestion control, and intelligent buffering optimize resource use, shortening job times and boosting GPU efficiency. The network and the automation tools around it must scale faster than humans to keep pace and to allow staff to move to higher-value tasks.

Benefits include operational cost reductions and long-term value via adaptable, forward-compatible structures. These solutions scale effortlessly for varying workload sizes, from small clusters to expansive GPU factories, all while upholding security and oversight throughout the environment. While mixing GPUs in the same rack is unlikely, the enterprise journey will span multiple GPUs and multiple suppliers.

Call to Action: To capitalize on these benefits, enterprises should pursue proof-of-concept trials with Al-optimized Ethernet switches, confirming efficiency improvements tailored to their Al setups.

#### New Standards Like UEC Enhance Al Networking

Pioneering efforts in Ethernet AI networking, bolstered by contributions to the Ultra Ethernet Consortium (UEC), advance standards for robust, expandable fabrics tailored to AI and ML demands. Benchmarking showcases capabilities like near-line-rate RoCEv2 performance at high speeds and strong outcomes in industry tests, often rivaling or surpassing traditional alternatives.

Highlights include techniques for balanced traffic flow and significant efficiency boosts over standard methods, enabled by innovative chip designs that enhance congestion handling. This establishes Ethernet as a reliable, high-performing option for AI environments. Leading Ethernet vendors will be founding members in standards, like UEC, to ensure that Ethernet continues to evolve and support rapid changes in AI infrastructure.

**Call to Action:** Enterprises should examine detailed benchmarking studies and participate in UEC-inspired trials to evaluate their Al clusters against top standards.



#### Unified and Interoperable (Brownfield and Greenfield)

Effective solutions provide cohesive integration across switching and management, offering complete oversight, automation, and insight for Al networks. This encompasses ready-to-use configurations for fabric setup and protocol enhancements, streamlining interactions among compute, storage, and connectivity elements.

At the same time, they prioritize interoperability, fitting into existing setups or new deployments via open designs and collaborations with major compute providers. Compatible optics facilitate smooth transitions across systems without interruptions.

Call to Action: To ensure compatibility, enterprises should collaborate with specialists to align legacy assets with integrated Al frameworks, enabling a smooth evolution.

#### Value of Real-time Protection - Zero Down Time or Disruption

Superior value comes from real-time safeguards that address vulnerabilities on the fly without halting operations, applying fixes immediately upon threat identification. This is enabled by enterprise-grade OS capabilities such as in-service upgrades, live patching, and integrated telemetry.

Robust operating systems bolster this with consolidated control over diverse environments, incorporating fabric tools for durability. Elements like intelligent fault handling and irregularity spotting ensure continuous availability, protecting AI processes from disruptions or risks.

Call to Action: To secure non-stop AI functionality, enterprises should adopt real-time protection mechanisms, starting with vulnerability assessments to pinpoint security vulnerabilities.

### Backend Infrastructure for Al Networking

Backend infrastructure forms the core of AI networking, utilizing high-density switches to enable non-blocking, lossless communication between GPUs and other compute elements. This setup supports advanced load-balancing techniques to minimize latency and ensure efficient data flows in large-scale AI clusters, where east-west traffic predominates. Security is intertwined through built-in encryption for data in transit, anomaly detection to identify potential breaches in real-time, and access controls that prevent unauthorized inter-node interactions, safeguarding



sensitive AI model training data from threats.

By integrating telemetry for continuous monitoring, backend solutions provide insights into performance metrics while embedding security layers that adapt to emerging risks. This holistic backend design not only optimizes computational efficiency but also ensures resilience against cyber threats, reducing the risk of data compromise.

Call to Action: Enterprises should evaluate backend switching options with integrated security features, scheduling assessments to enhance latency reduction and threat protection in their Al clusters.

#### Frontend Infrastructure for Al Networking

Frontend infrastructure focuses on user-facing and management aspects, delivering intuitive interfaces and tools for orchestration, visibility, and policy enforcement in Al data centers. This includes topology-aware monitoring, enabling administrators to oversee operations with full contextual awareness and make real-time optimizations. Security is embedded via endpoint protections, such as encryption for management traffic and role-based access controls, ensuring that frontend interactions remain secure against external intrusions or insider threats.

These frontend elements complement backend capabilities while maintaining a secure perimeter through features like threat intelligence integration and automated compliance checks. This approach fosters operational agility, allowing enterprises to scale AI initiatives confidently without exposing vulnerabilities at the network's edge.

Call to Action: To strengthen frontend management, enterprises should explore tools offering visibility and security integrations, arranging trials to secure and streamline oversight of Al networks. Observability enables customers to get end-to-end visibility into Al workloads, so making sure all toolsets move in this direction will ease Al adoption.

#### Multi-layer Security Integration for all Al Networking Components

Security integration across AI networking unifies protections throughout backend and frontend layers, creating a cohesive defense strategy that addresses the unique vulnerabilities of AI workloads. This encompasses advanced shielding mechanisms, such as hyper-distributed firewalls and MACsec encryption, applied consistently to prevent data leaks during GPU



communications or management access. Intelligence-driven security refines policies dynamically, using telemetry insights to detect and respond to anomalies, ensuring complete visibility into potential threats without compromising performance.

By intertwining security at every level—from backend data flows to frontend interfaces—enterprises achieve zero-trust architectures that protect against model poisoning, supply chain attacks, and regulatory non-compliance. This end-to-end approach not only mitigates risks but also enhances overall network intelligence, enabling proactive defenses that evolve with Al advancements.

Call to Action: Enterprises should prioritize security-integrated networking solutions, initiating comprehensive audits to embed protections across their entire Al infrastructure.

#### Best Designs, and Longevity with Investment Protection

While Ethernet facilitates easier brownfield retrofitting by leveraging existing infrastructure, it's important to note that such upgrades can often take longer than greenfield deployments due to the need for compatibility assessments, phased migrations, and potential custom configurations to integrate legacy systems. This extended timeline arises from evaluating current assets, ensuring minimal disruption during transitions, and validating performance post-upgrade, which contrasts with the streamlined setup of greenfield environments where designs can be optimized from scratch without historical constraints.

Despite these considerations, Ethernet's adaptability in brownfield scenarios ultimately provides greater long-term benefits, allowing enterprises to evolve their networks incrementally while avoiding the pitfalls of proprietary systems. By choosing Ethernet, organizations can mitigate the delays associated with retrofitting through strategic planning and interoperable tools that accelerate integration.

**Call to Action:** Capitalize on collaborative ecosystems through reference design sessions, ensuring facilities can handle industry-validated liquid cooling solutions, rack weight considerations, and evolving power requirements across multiple hardware generations. This will accelerate Al rollouts by leveraging established synergies.



#### Partnership Advantage and Uniqueness

Strategic alliances between networking specialists and computing innovators merge expertise with cutting-edge tech, yielding refined Al fabrics featuring automated setups and precise balancing for enhanced flow management.

Gains encompass faster processing, fluid compatibility in large-scale clusters, and novel secure validation frameworks. These partnerships excel in promoting scalable Ethernet, allowing enterprises to craft adaptable, top-tier Al systems.

Call to Action: Leverage networking-compute ecosystem partnerships by participating in joint reference architecture workshops. These collaborations combine advanced compute platforms with scalable Ethernet fabrics, delivering validated designs that shorten time-to-deployment while ensuring long-term investment protection. Company partnerships indicate an interest in vendors coming together to help customers move faster, and enterprises should learn from these vendor activities, which can include partnerships beyond networking and computing, such as facilities, cooling, and power.

#### Conclusion

As Al continues to reshape industries, enterprises must prioritize adaptable data centers that support innovation while addressing challenges like scalability, security, and sustainability. The journey toward Al excellence involves embracing technologies that offer performance, efficiency, and future-proofing, ensuring competitive advantage in a data-driven world. By focusing on integrated, intelligent infrastructures, organizations can unlock Al's transformative potential without undue risks.

All businesses will adopt AI, both directly and indirectly, and will undergo significant changes over the next decade. Having the ability to change and adapt is key to trying new AI tools and scaling AI at the right pace. Enterprises that take proactive steps, such as partnering with industry leaders, to build resilient AI ecosystems will drive long-term success.