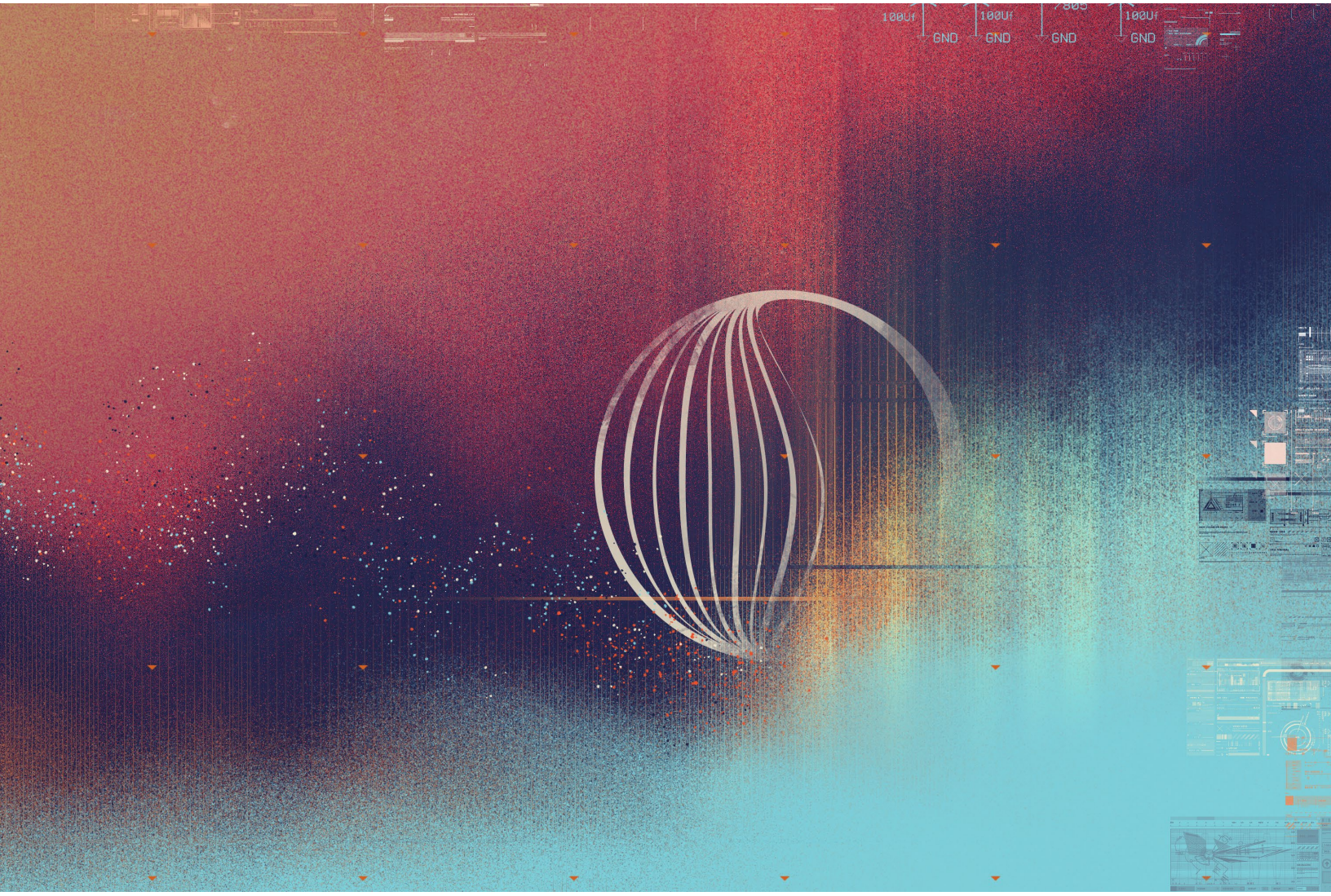
signal65



TOPS/OPS PERFORMANCE ANALYSIS

Dell PowerEdge XE9680 H200 Cluster with Dell 400GbE Networking

Executive Summary

AUTHOR

Brian Martin
AI and Data Center Lead | Signal65

IN PARTNERSHIP WITH

DELL Technologies

SEPTEMBER 2025

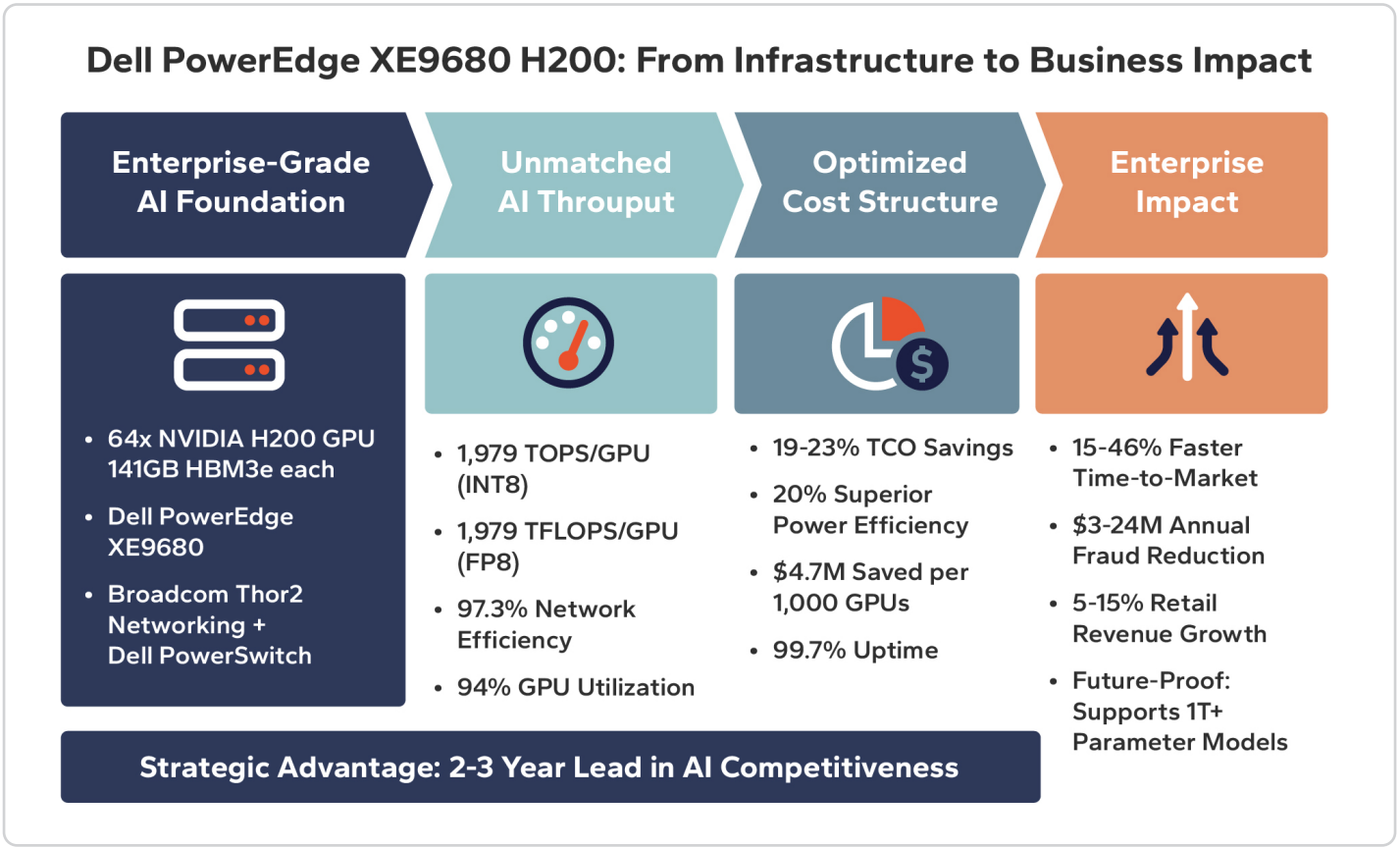
Executive Summary

Performance at Scale for Enterprise AI

The convergence of Dell's enterprise-grade infrastructure with NVIDIA H200 GPUs establishes a solid standard for AI performance at scale. The Dell PowerEdge XE9680 platform, with 64 NVIDIA H200 GPUs across 8 nodes and interconnected through Broadcom Thor2 networking and Dell PowerSwitch fabric, redefines efficiency, scalability, and total cost economics for large-scale machine learning.

Benchmarks confirm the transformative capability of the H200. Each GPU delivers the full **1,979 TOPS for INT8 inference** and **1,979 TFLOPS for FP8 training workloads**, enabled by the Transformer Engine and backed by 141GB of HBM3e memory. Combined with Broadcom BCM57608 Thor2 NICs and Dell PowerSwitch Z9864F switches, the cluster sustains **97.3% network efficiency under peak workloads**, supporting near-linear scaling across all 64 GPUs.

This translates directly into business impact. This platform accelerates AI from pilot projects to production systems, enabling faster innovation cycles, predictable economics, and a strategic foundation for long-term competitiveness.



Business Impact

From Infrastructure to Competitive Differentiation

Accelerated Revenue Growth

Production deployments achieve **15–46% faster time-to-market** than industry benchmarks, cutting product launch timelines by 8–12 weeks. Real-world cases include fraud detection improvements valued at **$3M–$24M annually** for financial services and **5–15% incremental revenue growth** in retail through improved recommendation and demand forecasting systems.

Operational Efficiency

With **94% GPU utilization** and **93% scaling efficiency**, enterprises maximize return on infrastructure investments. Dell proactive monitoring and predictive maintenance enable **99.7% uptime**, while eliminating an average **$2.3M annual cost** of specialized AI infrastructure expertise.

Cost Optimization

Three-year total cost of ownership (TCO) analysis demonstrates **19–23% savings** compared to alternative solutions. Gains come from optimized acquisition, **20% superior power efficiency**, and streamlined operations. Unlike cloud deployments, cost structures remain predictable and improve over time, yielding **$4.7M savings per 1,000 GPUs**.

Innovation Velocity

Unified management tools simplify operations, allowing AI teams to focus on model development and deployment instead of infrastructure. Customers report faster iteration cycles and the ability to deliver breakthrough applications.

Future-Proof Investment

With architectural flexibility, support for **1T+ parameter models**, and 141GB of HBM3e per GPU, organizations are positioned to lead through successive AI technology cycles without disruptive infrastructure migrations.

Strategic Value

Catalyze Business Transformation

By combining Dell's proven enterprise reliability, NVIDIA's next-generation accelerators, and Broadcom's high-efficiency networking, the Dell PowerEdge XE9680 H200 platform provides more than just computational capacity. **It offers a scalable foundation for enterprise-wide AI transformation**—from early pilot deployments to global production systems—while mitigating risk through supply chain stability, robust support, and predictable economics.

Organizations adopting this platform secure a **two- to three-year competitive advantage**, leverage infrastructure capable of training frontier-scale models, and reduce execution risk that often derails experimental AI projects. The result is a strategic investment that accelerates revenue, optimizes cost, and positions enterprises for long-term AI leadership.

Important Information About this Report

CONTRIBUTORS
Brian Martin
AI and Data Center Lead | Signal65

PUBLISHER
Ryan Shrout
President and GM | Signal65

INQUIRIES
Contact us if you would like to discuss this report and Signal65 will respond promptly.

CITATIONS
This paper can be cited by accredited press and analysts, but must be cited in-context, displaying author's name, author's title, and "Signal65." Non-press and non-analysts must receive prior written permission by Signal65 for any citations.

DISCLOSURES
Signal65 provides research, analysis, advising, and consulting to many high-tech companies, including those mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

IN PARTNERSHIP WITH

DELL Technologies

ABOUT SIGNAL65
Signal65 is an independent research, analysis, and advisory firm, focused on digital innovation and market-disrupting technologies and trends. Every day our analysts, researchers, and advisors help business leaders from around the world anticipate tectonic shifts in their industries and leverage disruptive innovation to either gain or maintain a competitive advantage in their markets.

signal65

This white paper was developed in partnership with Dell Technologies and represents comprehensive TOPS/OPS performance analysis conducted by Signal65's AI performance engineering team. All benchmark results and recommendations are based on rigorous testing methodologies and real-world deployment scenarios optimized for enterprise AI infrastructure excellence.

Performance Disclaimers: Results based on synthetic benchmarks and controlled testing environments; actual performance may vary based on specific workload characteristics, configuration variations, and operational conditions. Benchmark data referenced from NVIDIA H200 Datasheet (August 2025), MLPerf Training v5.0 results, and Broadcom Thor2 specifications. For production planning, conduct proof-of-concept testing with organization-specific AI workloads.

Read the full report at: https://signal65.com/research/dell-poweredge-xe9680-h200-cluster-with-dell-400gbe-networking/

signal65

CONTACT INFORMATION
Signal65 I signal65.com