



EXECUTIVE SUMMARY

# Dell AMD Instinct Series GPU Cluster with Dell Networking

**AUTHOR**

**Brian Martin**  
AI and Data Center Lead | Signal65

IN PARTNERSHIP WITH

**DELL**Technologies

SEPTEMBER 2025



# Executive Summary

## Enterprise AI Leadership with Infrastructure Innovation

The Dell–AMD–Broadcom solution represents a pivotal shift in enterprise AI infrastructure, delivering hyperscale-class performance in a platform designed for cost efficiency, operational simplicity, and supply chain resilience. By combining Dell’s enterprise-grade platforms, AMD’s Instinct accelerators, and Broadcom’s high-performance networking, organizations gain an integrated solution that accelerates time-to-value and reduces total cost of ownership by 23–28% over a three-year cycle.

## Business Velocity at Scale

This architecture removes the bottlenecks that traditionally slow AI initiatives. With 192GB of high-bandwidth memory per GPU, enterprises can experiment with models 3.4x faster, train 400B+ parameter workloads without complex sharding, and shorten AI development cycles by up to 12 weeks. The result: faster innovation, quicker time-to-market, and the ability to outpace competitors in deploying transformative AI applications.

## Strategic Differentiation and Risk Reduction

A multi-vendor design ensures greater resilience and freedom of choice. Component availability improves by more than 2x, procurement delays are reduced by 73%, and the open-source ROCm ecosystem removes lock-in, enabling organizations to adapt rapidly as AI use cases evolve. Backed by Dell’s enterprise services and Broadcom’s proven networking, companies gain both the agility and confidence required for sustained innovation.

## Competitive Imperative

In today’s AI-driven economy, infrastructure determines innovation velocity. The Dell–AMD–Broadcom solution transforms infrastructure from a constraint into a competitive accelerator. Early adopters are already realizing transformative business outcomes:

- **Financial services:** Cutting fraud detection cycles from weeks to days
- **Pharmaceuticals:** Accelerating drug discovery pipelines by 5x
- **Retail:** Driving double-digit revenue growth through enhanced personalization and analytics

These aren’t just performance wins — they are strategic outcomes that expand market share, reduce costs, and strengthen leadership positions over time.

## Key Takeaways: Performance at Scale



### Faster AI Cycles

3.4x faster experimentation



### Lower TCO

23-28% TCO savings



### Scalable Performance

89.7% scaling efficiency



### Resilient & Flexible

Multi-vendor, open ecosystem

## Accelerated Innovation

- **3.4x faster model experimentation** enables AI teams to validate ideas and pivot quickly — reducing time-to-market by up to 12 weeks.
- **89.7% scaling efficiency** ensures larger models can be trained reliably across 64 GPUs, unlocking next-generation AI use cases.

## Productivity & Cost Advantage

- **192GB HBM3 memory per GPU** supports **50% larger batch sizes**, maximizing infrastructure utilization and lowering costs per experiment.
- **BF16 mixed precision** doubles throughput (2.1x over FP32) without compromising accuracy — cutting compute time and energy spend.

## High-Performance Inference

- **<50ms latency** on 70B models delivers near real-time responsiveness, powering advanced customer experiences and decision automation.
- **900+ tokens/sec/GPU** throughput ensures scalable AI services without runaway infrastructure costs.

## Resilient Networking Backbone

- **96%+ fabric efficiency** means the network is no longer a bottleneck.
- Consistent **18.7 GB/s bandwidth** maintains performance under scale, even at **95% utilization** — enabling predictable operations and business continuity.

# 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.

## LICENSING

This document, including any supporting materials, is owned by Signal65. This publication may not be reproduced, distributed, or shared in any form without the prior written permission of Signal65.

## 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



## 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.



Read the full report at: <https://signal65.com/research/dell-amd-instinct-series-gpu-cluster-with-dell-networking/>



## CONTACT INFORMATION

Signal65 | [signal65.com](https://signal65.com)