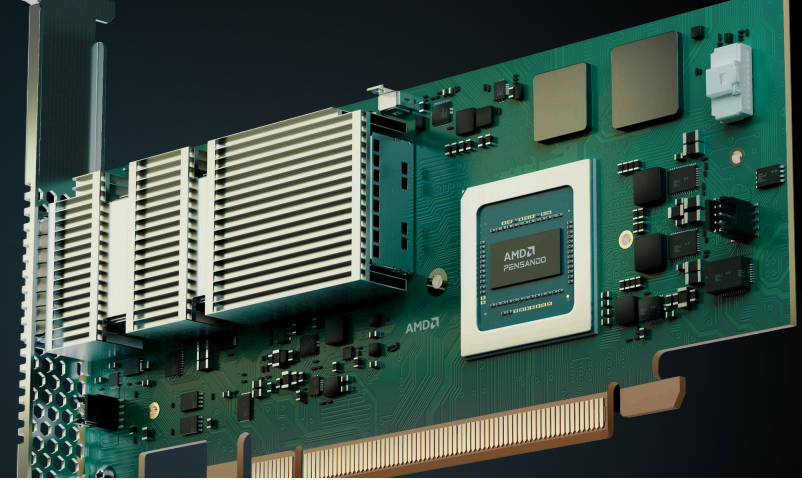
AMD
together we advance_

AMD PENSANDO™ VULCANO 800 AI NIC: SCALE OUT WITH CONFIDENCE

AMD PENSANDO™ VULCANO 800 AI NIC: SCALE OUT WITH CONFIDENCE



INTRO

The AMD Pensando™ Vulcano 800 AI NIC is a leading-edge solution that enables high-performance GPU-GPU communication, delivering Ethernet speeds of up to 800 Gigabits per second (Gbps).

As the only AI NIC on the market delivering up to 2.4 Tbps scale-out bandwidth per GPU[1], the AMD Pensando Vulcano 800 AI NIC is specifically designed to optimize performance and scalability, while delivering reliability, availability, and serviceability.

With full hardware and software programmablility and being backed by an open ecosystem, the AMD Pensando Vulcano 800 AI NIC sets a new standard in networking technology. Built on the AMD Pensando P4 programmable engine, it provides a foundation for building resilient, scalable, and fault tolerant networks for AI systems, helping maximize throughput and helping reduce downtime.

The AMD AI NIC™ fully programmable hardware pipeline allows for customers to utilize industry-standard solutions such MRC or UEC transport protocols, while also providing the flexibility to adopt custom transport protocols. It preserves customer choice while maintaining an open ecosystem for scale-out networking fabric without compromising on performance, enabling data centers to remain future-ready.

PRODUCT HIGHLIGHTS

Scale Out with Confidence

Maximize the value from AI infrastructure investments with the AMD Pensando™ Vulcano 800 AI NIC. Exceptional performance, with the flexibility required to scale out now, and in the future.

50% MORE GPU BANDWIDTH[1]

The AMD Pensando™ Vulcano 800 AI NIC is the only NIC delivering 2.4 Tbps of aggregate scale-out bandwidth per GPU via its unique 3 NICs-per-GPU configuration. With traffic distributed across three independent network planes simultaneously, keep GPUs productive at scale.

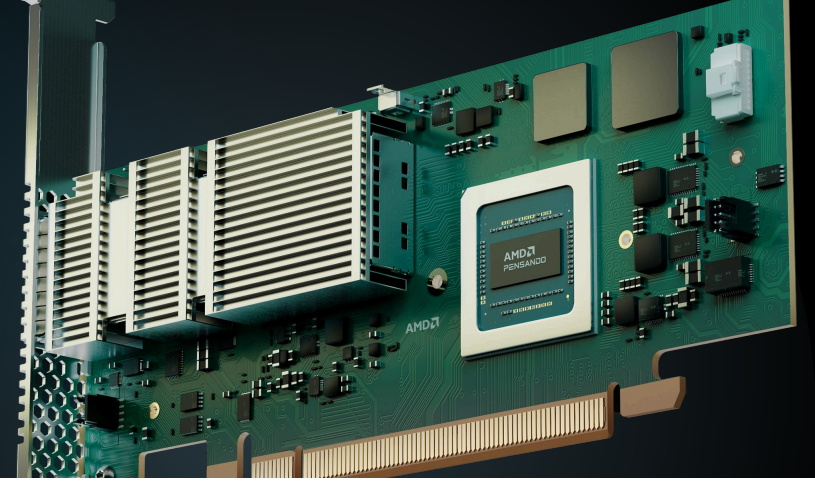
UP TO 13% FASTER AI JOB COMPLETION TIMES[2]

With unprecedented networking bandwidth and features including intelligent network load balancing, fast failover and loss recovery, the AMD Pensando™ Vulcano 800 AI NIC helps accelerate workloads while maximizing AI investments.

UP TO 33% LOWER NETWORK SWITCHING COSTS[3]

Drive lower infrastructure costs for scale-out AI networking without sacrificing performance. With support for double the SerDes capacity from 100G to 200G and an open Ethernet ecosystem, operators can scale AI clusters efficiently and flexibly.

AMD PENSANDO™ VULCANO 800 AI NIC: SCALE OUT WITH CONFIDENCE



Enhanced cluster uptime

Built for reliability, availability, and serviceability (RAS), rapid fault detection and fast failover isolate unhealthy paths before workload performance is impacted. Accelerated network convergence then restores full cluster throughput with minimal GPU idle time.

Programmability

Built on the proven AMD Pensando™ P4 programmable architecture, the AI NIC is fully hardware and software programmable, supporting UEC-ready RDMA, Multipath Reliable Connection (MRC), and other intelligent custom network transports.

Operational excellence

AI infrastructure maintains production-readiness with automated deployment, minimized downtime, and cluster-wide visibility.

PROGRAMMABILITY

Fully Programmable Software + Hardware Stack

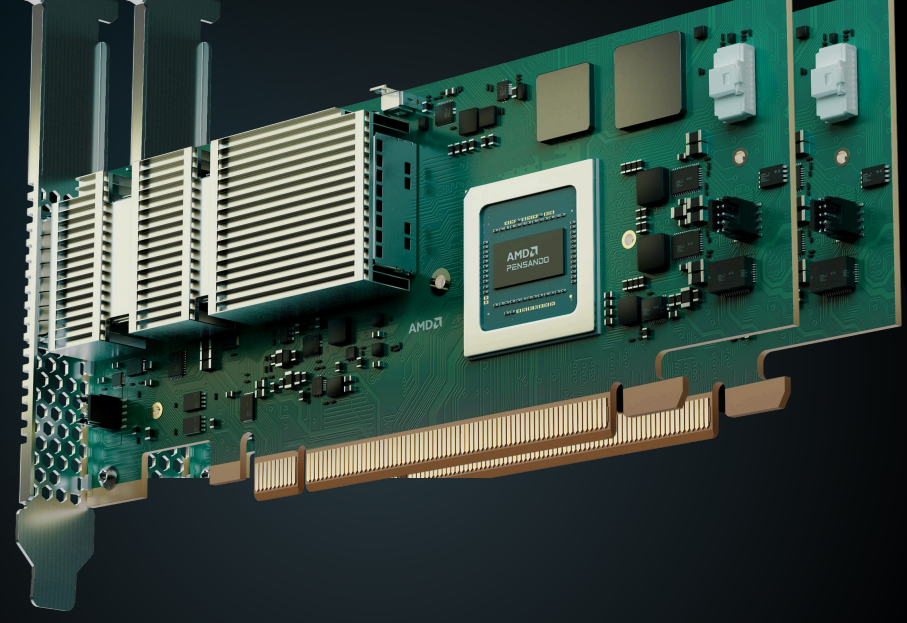
The AMD Pensando™ **Vulcano 800 AI NIC** supports **Multipath Reliable Connection (MRC)**, the newest open-Ethernet protocol codeveloped for OpenAI. Full programmability decouples innovation from silicon refreshes, enabling the AI NIC to meet evolving AI workload demands as they emerge.

The full hardware programmability of the P4 engine also allows AI teams to delve into rich telemetry data. The data dives beyond bandwidth and packet counters, providing job-level telemetry data. This allows teams to understand where the AI job bottlenecks are occurring, even beyond the network.

Built-in traffic and congestion management enable consistent performance as you grow from one rack to an entire data center. The AMD AI NIC™ supports three distinct load balancing architectures for flexibility with workload requirements:

	NIC-Based Packet Spray	NIC-Based Source Routing	Switch-Based Packet Spray
Use Cases	Large fabrics with many equal-cost multi paths.	Application communication patterns are well defined.	Environments where switching infrastructure natively supports per-packet DLB.

AMD PENSANDO™ VULCANO 800 AI NIC: SCALE OUT WITH CONFIDENCE



Are you ready to advance your data center?

2.4 Tbps scale-out bandwidth for where AI takes you next.

Contact your **AMD sales representative** or visit **AMD.com/Pensando** to learn more about our ability to deliver big value for scale-out AI workloads.

ENDNOTES

1. Result of Vulcano board 3-NICs-per-GPU configuration.

2. MI400-018: Based on AMD Engineering silicon modeling and AMD synthetic benchmark simulation, using the MOE_4p5T hi_sparsity_ea9 benchmark test to project time- to- solution speed up in days using the FP8 training datatype on a simulated LLM system modeled with 8000 AMD Instinct MI455X GPUs and (3) versus (2) AMD Pensando Vulcano NICs.
   Results reflect analysis of pure training compute (decode layers only) Configuration evaluated with a global batch size of 4096 and a sequence length of 4K, using SwiGLU activation. Results exclude evaluation and checkpointing overhead. FlashAttention v3 with matmul–softmax overlap is assumed. AllReduce and All2All communication costs are fully accounted for and not hidden via tiled compute–communication overlap. Gradient synchronization and FSDP weight prefetch are evaluated across varying levels of overlap to assess scale-out sensitivity. Results assume ideal and not fully optimized real-world behavior and may vary when actual product(s) are released in market.

3. PEN-022: AMD comparison and pricing as of May 18, 2026, for network fabric costs to support 32,000 GPUs. Comparison of a Vulcano-based NIC (VULCANO-CUSTOM-2.4T) deployed as part of a Helios rackscale system with a network based on 1.6T Tomahawk 6 switching with 200G SerDes versus using a competitor 800G NIC with 800G Tomahawk 6 switching with 100G SerDes. Both fabrics were fat-tree topologies built on Tomahawk 5 800G switching platforms, with NIC costs considered comparable. The Vulcano-based design is estimated to deliver up to 33% savings in network switching costs by enabling a more cost-effective architecture with fewer switching platforms, more bandwidth per port on the network, and reduced transceiver cables/optics.

   TH6-100G Serdes Fat-Tree (Competition):
   Switching (TH6C BCM78914 - 128x800G):
   - Leaf Units 1,000
   - Spine Units 500
   - Total Switches 1,500
   - TH6 100G Unit Price $79,587
   - Total Switching Cost $60M

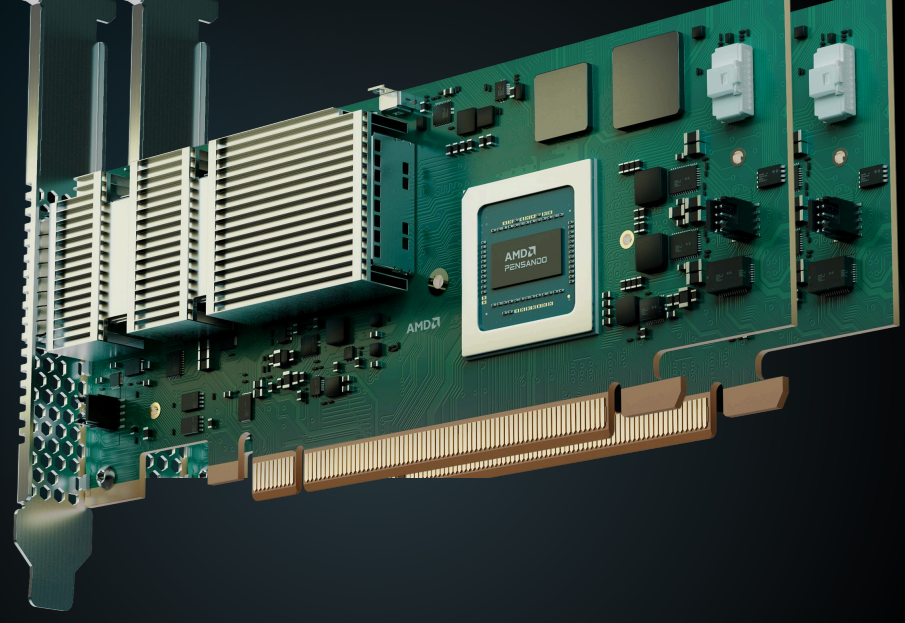
   Cables/Optics:
   - NIC Transceivers (800G-DR8) 64,000 @ $500
   - Leaf/Spine Transceivers (800G-DR8) 192,000 @ $500
   - MPO Cables 256,000 @ $89
   - Total Optics/Cables $64M

   Total Fabric Cost TH6-100G (Switches + Cables/Optics): $124M

   TH6-200G Serdes Fat-Tree (Vulcano-Custom2.4T / AMD Solution):
   Switching (TH6P BCM78910 - 64x1.6T):
   - Leaf Units 1,000
   - Spine Units 500
   - Total Switches 1,500
   - TH6 200G Unit Price $66,336
   - Total Switching Cost $50M

   Cables/Optics:
   NIC Transceivers (1.6T-DR8) 32,000 @ $900 (50% fewer vs. competition)
   Leaf/Spine Transceivers (1.6T-DR8) 64,000 @ $900
   MPO Cables 128,000 @ $89 (50% fewer vs. competition)
   Total Optics/Cables $43M
   Total Fabric Cost TH6-200G (Switches + Cables/Optics): $93M

AMD PENSANDO™ VULCANO 800 AI NIC: SCALE OUT WITH CONFIDENCE



Capex Savings (Fabric only):

Savings $:	$30.7M
Savings %:	33.1%

Pricing sources:

SemiAnalysis Hyperscaler Networking Model data used with permission; full analysis available via SemiAnalysis subscription. Edgecore switch pricing as of May 17, 2026. Results may vary based on system configuration.