

SCALING OUT FUTURE-READY AI INFRASTRUCTURE:

The Complete Guide to the
AMD PENSANDO™ POLLARA 400 AI NIC

THE AMD PENSANDO™ POLLARA 400 AI NIC: PURPOSE BUILT FOR HIGH-THROUGHPUT, LOW-LATENCY AI NETWORKING

A clear-eyed assessment of traditional networks' ability to handle AI workloads points to an inexorable conclusion: What got us here won't get us there.

What got us here – and continues to excel for general-purpose computing – is conventional Ethernet. Today, Ethernet is still the networking technology of choice in AI clusters, largely due to its ubiquity, cost-effectiveness, and familiarity to networking teams. But its limitations in handling the demands of generative AI and large language models (LLMs) are increasingly conspicuous.

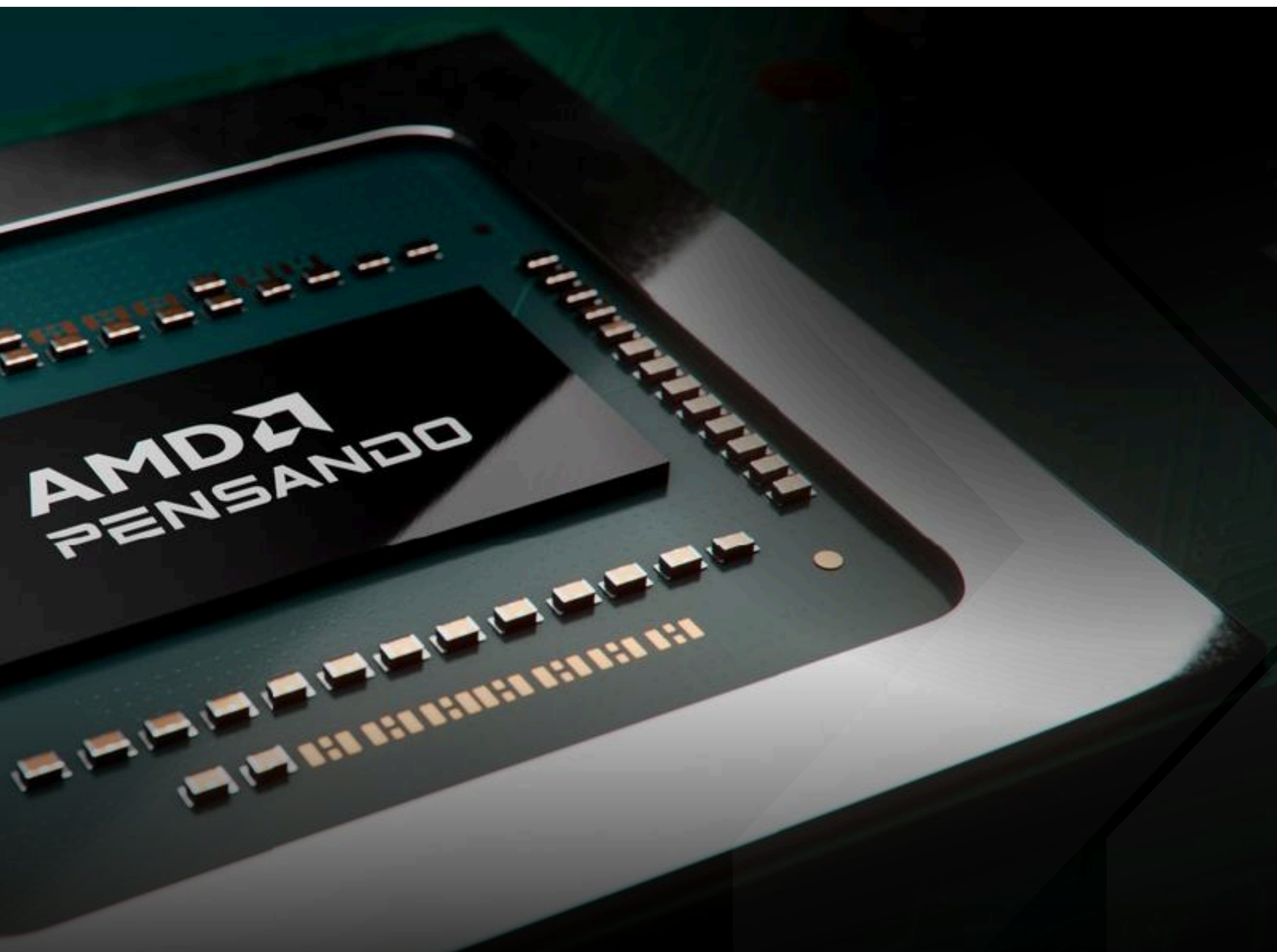
From chatbots and real-time language translation to medical imaging and beyond, AI is driving massive compute demand, and networks are showing the strain. Advanced AI/ML models have unique networking requirements, including the need for tightly coupled parallel processing, rapid data transfers, and low-latency communication.

AI networks are, out of necessity, unforgiving – they have none of the tolerance for dropped or lost data packets that characterizes the “lossy” Ethernet protocol. Ethernet was not designed to handle such stringent demands, resulting in challenges of performance, scalability, and reliability, among others

So, if not conventional ethernet, what will get us there?

A key innovation in the quest to meet the evolving demands of AI is the **AMD Pensando Pollara 400 AI NIC** – a network accelerator that's purpose-built to optimize large-scale AI workloads over Ethernet-based networks.

Serving as a bridge between Ethernet's broad compatibility and the specialized demands of AI workloads, the AMD Pensando Pollara 400 AI NIC combines the best of both worlds. In addition to its widespread adoption and cost-effectiveness, Ethernet offers deep operational familiarity, a broad ecosystem of hardware and tools, and established automation and observability frameworks, benefits that remain essential as organizations transition to AI-optimized networking. As the first fully software and hardware programmable AI NIC on the market, the AMD Pensando Pollara 400 AI NIC allows hyperscalers, enterprises, cloud and neocloud providers, and sovereign AI alike to harness the potential of their AI workloads, boosting performance, scalability, and reliability without sacrificing the benefits of Ethernet infrastructure.



- > **BROAD ETHERNET COMPATIBILITY**
- > **SPECIALIZED DEMANDS OF AI WORKLOADS**
- > **HIGH-PERFORMANCE, LOW-LATENCY AI NETWORKING**

OPERATIONAL AND PERFORMANCE GAINS FOR AI WORKLOADS

As organizations push increasingly muscular AI workloads into production, they face an array of pain points, such as performance bottlenecks, energy- and cost-efficiency pressures, and the demand for continuous uptime for mission-critical applications.

The AMD Pensando™ Pollara 400 AI NIC addresses these and other challenges through the following key measurable business and technical benefits.

1. HIGHER PERFORMANCE

Of the many imperatives that networking teams seek to address in their AI infrastructure clusters, high performance and high resilience top the list. Performance demands in AI networking are exceptionally rigorous due to the cost of training AI models and the extreme sensitivity to efficient movement of data between cluster nodes.

However, conventional Ethernet does not prioritize the high bandwidth, low latency, and efficient communication patterns demanded by modern AI workloads. To meet these demands, networks need to be able to support distributed computing over multiple GPU nodes, with low jitter and the power to exchange data in high bandwidth memory at lightning speeds. through Remote Direct Memory Access (RDMA).

Scale-out networks for AI infrastructure typically employ one of two networking technologies: InfiniBand or “lossless” Ethernet – a flavor of Ethernet that eliminates packet loss under heavy traffic, also known as RDMA over Converged Ethernet (RoCE). Ethernet-based technologies are increasingly adopted in large-scale AI clusters due to their ecosystem maturity, cost profile, and operational flexibility.

The **AMD Pensando Pollara 400 AI NIC** offers high throughput across all available links, along with congestion avoidance, low tail latency, scalable performance, and fast job completion times – all while allowing customers to leverage familiar Ethernet-based fabrics.

The AMD Pensando Pollara 400 AI NIC is the first AI NIC on the market in its class to offer transport protocols designed to bring Ethernet into the era of AI, enabled by its unique programmability. These protocols are based on standards set by the **Ultra Ethernet Consortium (UEC)**, of which AMD was a founding member in 2023, and include the consortium’s UEC-ready RDMA. Through UEC-ready RDMA, the AI NIC achieves up to 20% higher performance and cluster uptime at scale for AI workloads than a comparable NIC¹.

It also accelerates both RDMA and AMD ROCm™ Communication Collectives Library (RCCL) communications that underpin AI infrastructure operations (RCCL is a distributed communication library for coordinating data movement across multiple GPUs and nodes in ROCm software-based systems).

As organizations push increasingly muscular AI workloads into production, they face an array of pain points, such as:

- Performance bottlenecks
- Energy- and cost-efficiency pressures
- The demand for continuous uptime for mission-critical applications

KEY BENEFITS:

400

Gigabits per second

Optimizes data transfer within back-end AI networks, delivering up to 400 Gigabits per second (Gbps) for GPU-to-GPU communication across scale-out AI infrastructure clusters.

~50%

Higher performance

Supports both lossless and lossy fabrics, achieving ~50% higher cluster-level performance in measured configurations².

20X

Greater scaling

Enables up to 20x greater scaling efficiency than InfiniBand for AI workloads³.

10%

Better RCCL performance

Achieves up to 10% better RCCL performance in measured configurations, helping overcome GPU-CPU bandwidth bottlenecks and reduce network latency⁴.

25%

Performance boost

Unlocks an average 25% performance boost when running UEC-ready RDMA vs. RoCEv2 workloads (Allreduce)⁵.

2. COST EFFICIENCY AND FLEXIBILITY

Cost efficiency and flexibility are key considerations in AI networking. In addition to its technical benefits, the AMD Pensando™ Pollara 400 AI NIC can help on the business side of the equation by slashing the CapEx of network infrastructure in half⁶. These savings stem from factors such as the solution's support for multiplane fabric architecture.

A key driver of these savings is the NIC's ability to intelligently manage AI traffic using advanced capabilities that have traditionally lived in switches. One of these capabilities is adaptive packet spray, a technique that intelligently load-balances AI traffic by simultaneously sending individual packets from a single data flow down different paths, often of varying latencies. Because packets can arrive at the destination out of order, the AMD Pensando Pollara 400 AI NIC delivers them to applications' memory region in the correct order.

Together, these features – adaptive packet spray, out-of-order packet handling, and in-order message delivery – help avoid bottlenecks and ensure full utilization of available bandwidth.

Having these capabilities live on the NIC lessens dependency on deep buffer switches. This helps reduce cost because a deep buffer switch costs more per port than simpler switches when being used in multiplane architecture. Thanks to the multiplane architecture, fewer switches are required for the same size of cluster, helping keep CapEx low. In a network comparison buildout of a 128K GPU cluster, the number of switches needed dropped from ~5,000 to just 3,000 with the AMD Pensando Pollara 400 AI NIC⁶.

The flexibility afforded by the AMD Pensando Pollara 400 AI NIC derives in part from its proven support for industry standards and features published by the UEC. As a result of this support, the AMD Pensando Pollara 400 AI NIC allows customers to choose their preferred Ethernet switching vendor and type of Ethernet implementation, whether lossless or lossy.

Significantly, in eliminating the need for a lossless network, the NIC reduces the complexities of configuring and operating networks. This flexibility and efficiency make the AMD Pensando Pollara 400 AI NIC a powerful solution for enhancing the performance of AI workloads and network reliability.

KEY BENEFITS:

~58%

Unmatched Infrastructure Economics

Can achieve up to ~58% lower network infrastructure cost⁶.

40%

Do More with Less

A dramatic reduction (up to 40% in our example) in the number of network switches needed.

1M

1M GPUs

Massive scalability with the ability to connect more than 1 million GPUs over RoCEv2 transport (compared with just 48,000 GPUs over InfiniBand), maximizing GPU cluster scale at lower cost⁷.

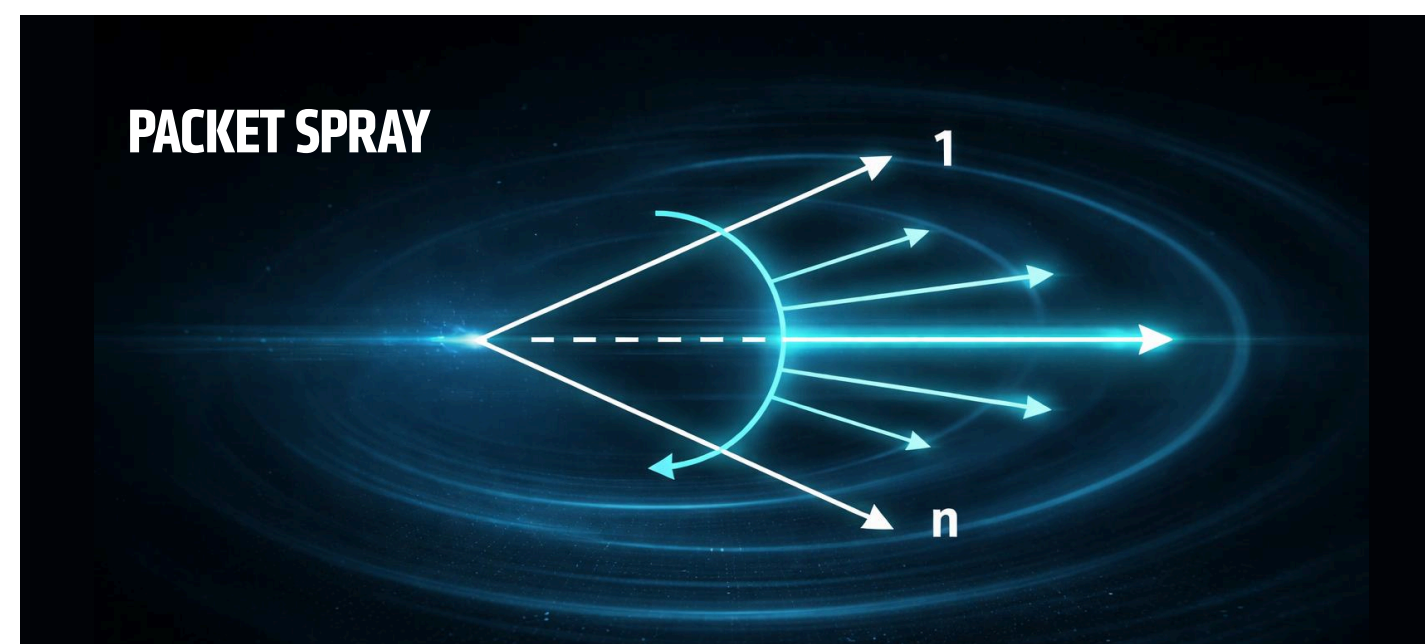
RESILIENT NETWORKING

Up to 40% fewer switches required for large-scale AI clusters when using AMD Pensando Pollara 400 AI NIC.

5000
switches needed

3000
switches needed with the AMD Pensando™ Pollara 400 AI NIC

PACKET SPRAY



3. IMPROVING CLUSTER UPTIME AND RELIABILITY

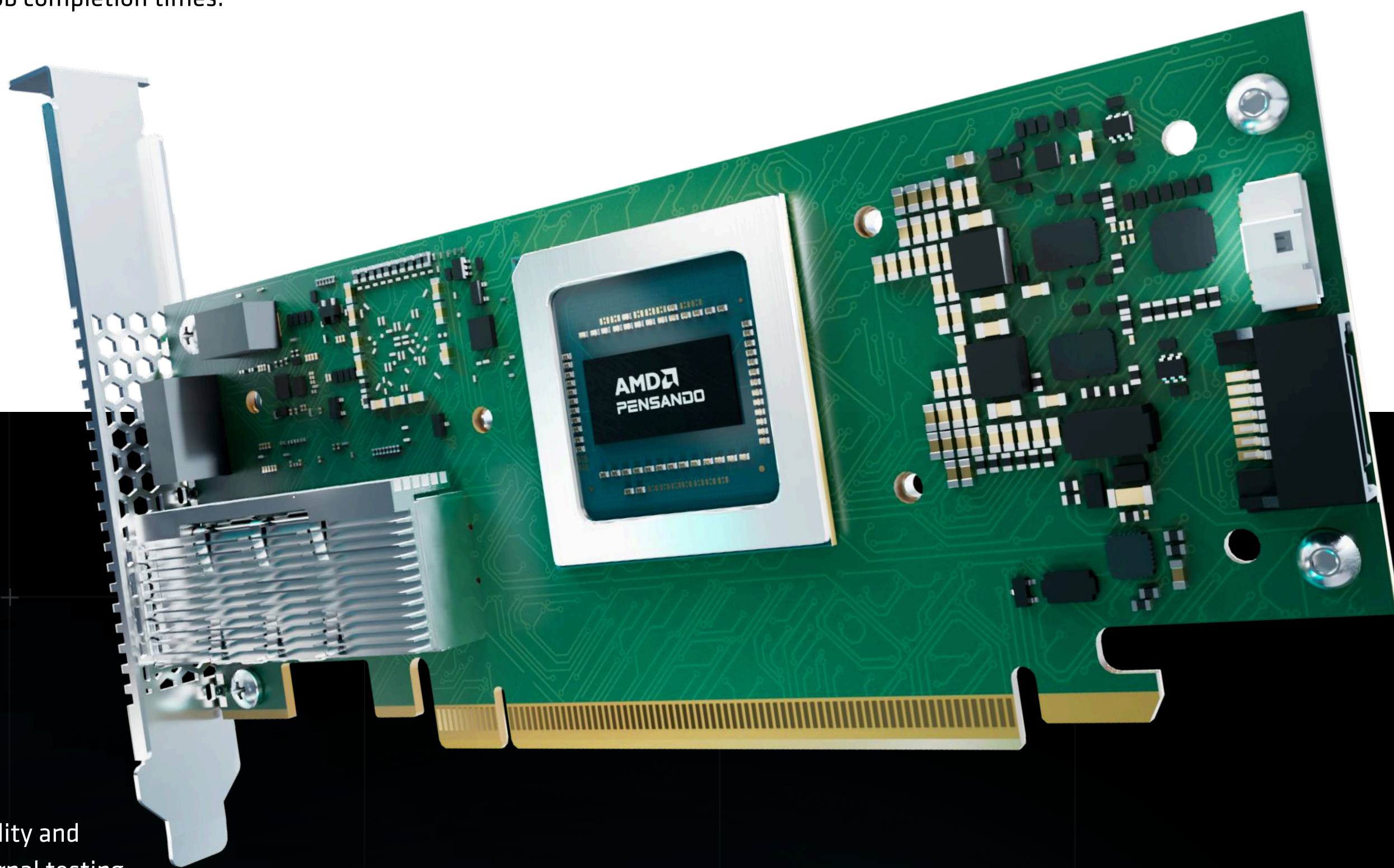
Continuous cluster uptime is vital for mission-critical AI applications, which rely on the highly efficient synchronization of data in AI GPU clusters. To achieve this, networking teams must prioritize reliability, availability, and serviceability (RAS) in designing high-performance networks.

The AMD Pensando™ Pollara 400 AI NIC employs aggressive fault detection mechanisms – a must to keep AI applications humming along at optimum performance. These mechanisms identify issues in milliseconds, enabling near-instantaneous failover and minimizing GPU idle time.

The ability to rapidly detect and isolate faulty paths allows the optimal allocation of network resources and helps ensure AI workloads move without interruption on healthy paths. This approach can enhance overall AI performance, potentially reducing training times, improving inference accuracy, and shrinking job completion times.

Add to this the NIC’s ability to intelligently manage network congestion using a sophisticated combination of real-time telemetry and network-aware algorithms – all without the need for a fully lossless network. Together, these features simplify network configuration, can reduce operational overhead, and avoid common issues, like congestion spreading, deadlock, and head-of-line blocking.

Also topping the list of networking teams’ objectives for AI infrastructure clusters is high network resilience. The AMD Pensando Pollara 400 AI NIC delivers operational resilience via rapid fault detection, fast failover, and loss recovery. Its support for multiplane architecture – which conveys transport traffic along multiple, independent, and parallel network planes or fabrics – enables high-availability failover with robust fault isolation. This can significantly improve networking RAS and supports the cost-effective scale-out of AI networks.



KEY BENEFITS:

50%

Maximize AI Availability

Boosts network reliability, availability and serviceability, by up to 50% in internal testing using RDMA vs. RoCEv2 on a Llama workload with a 0.1% traffic drop, which help improve overall cluster uptime².



System Resilience

Eliminates single points of failure, helping ensure that a failure in one plane does not compromise the integrity of others.



Flawless Delivery

Reduces packet drops, delays, and failures.



Swift Recovery

Rapid fault detection, fast failover, and fast loss recovery.

4. ACHIEVING NETWORK OPERATIONAL EXCELLENCE

Achieving network operational excellence is critical for keeping large AI clusters running reliably and at high utilization. In addition to its performance benefits, the AMD Pensando™ Pollara 400 AI NIC helps on the operational side by minimizing disruptions, accelerating troubleshooting, and streamlining deployments through on-NIC intelligence and automation.

A key driver of this excellence is the NIC’s ability to handle advanced operational tasks – like graceful and hitless software upgrades, consistency validation, and rich telemetry – that have traditionally required complex switch features or manual intervention. Graceful and hitless upgrades enable seamless maintenance without interrupting GPU jobs, while cluster-wide consistency validation via the AI NIC checker catches configuration drift early, helping reduce operational risk.

Together, these features – enhanced observability with high-resolution metrics on queue pairs, interfaces, and paths, plus programmable telemetry – accelerate mean time to detect and repair issues before they cascade across the cluster.

Having these capabilities on the NIC simplifies day-to-day operations, from proactive monitoring via standard tools and Grafana dashboards to automated Kubernetes deployments with the AMD Network Operator. In large-scale builds like 128K GPU clusters, this cuts configuration toil and boosts uptime without deep expertise.

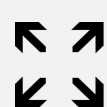
The operational flexibility of the AMD Pensando Pollara 400 AI NIC derives in part from its tight integration with modern orchestration frameworks and open standards. As a result, it supports diverse environments – bare metal, virtualized, or multi-CNI – while enabling rapid scaling and evolution. Significantly, this approach reduces complexity in upgrades and monitoring, making the AMD Pensando Pollara 400 AI NIC a powerful solution for resilient AI fabrics that support continuous innovation.

KEY BENEFITS:



Seamless Maintenance

Graceful and hitless upgrades for seamless maintenance without downtime.



Consistency Validation

Proactive consistency validation to reduce configuration errors across clusters.



Faster Troubleshooting

Rich, high-resolution observability for faster issue detection and troubleshooting.

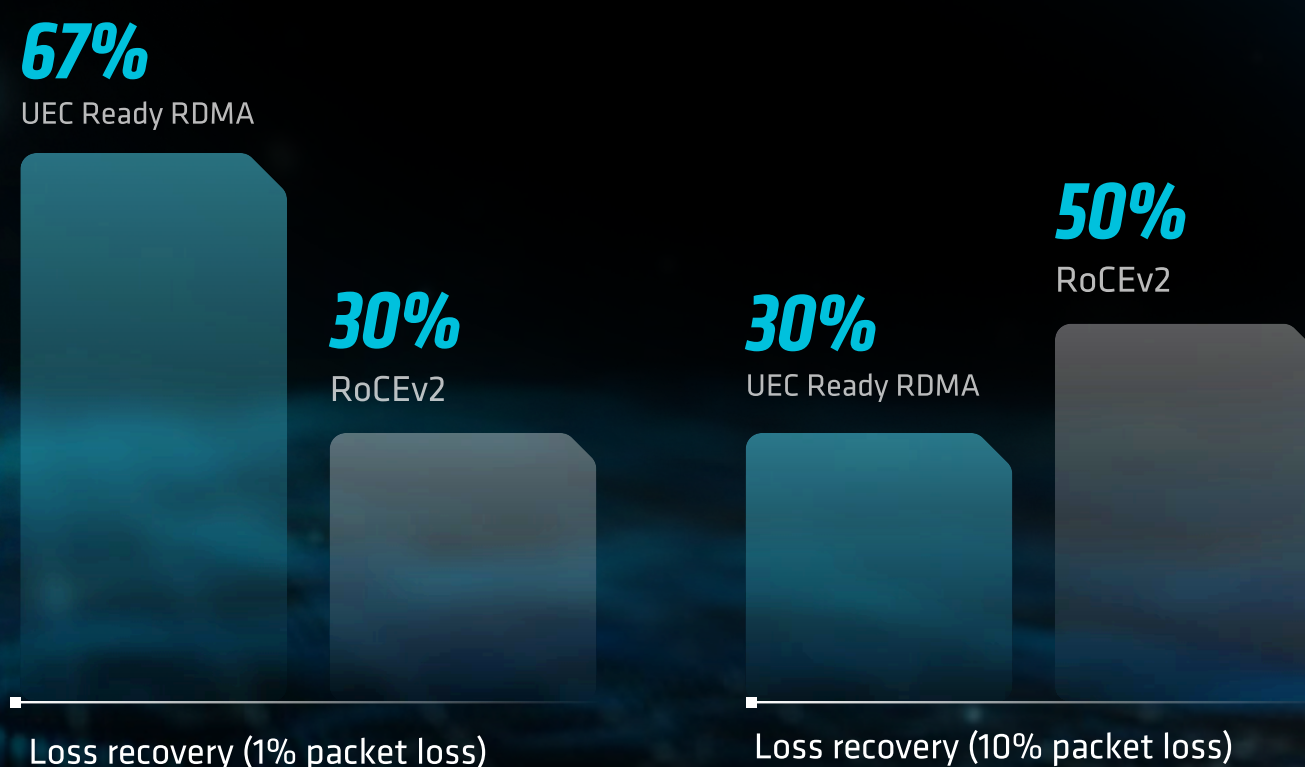


Automated Deployment

Automated Kubernetes deployment via Helm charts and flexible CNI support to accelerate time to production.

RESILIENT NETWORKING:

Network Reliability & Resiliency: UEC Ready RDMA vs RoCEv2 50% Improvement in AI Cluster Uptime with Pollara 400

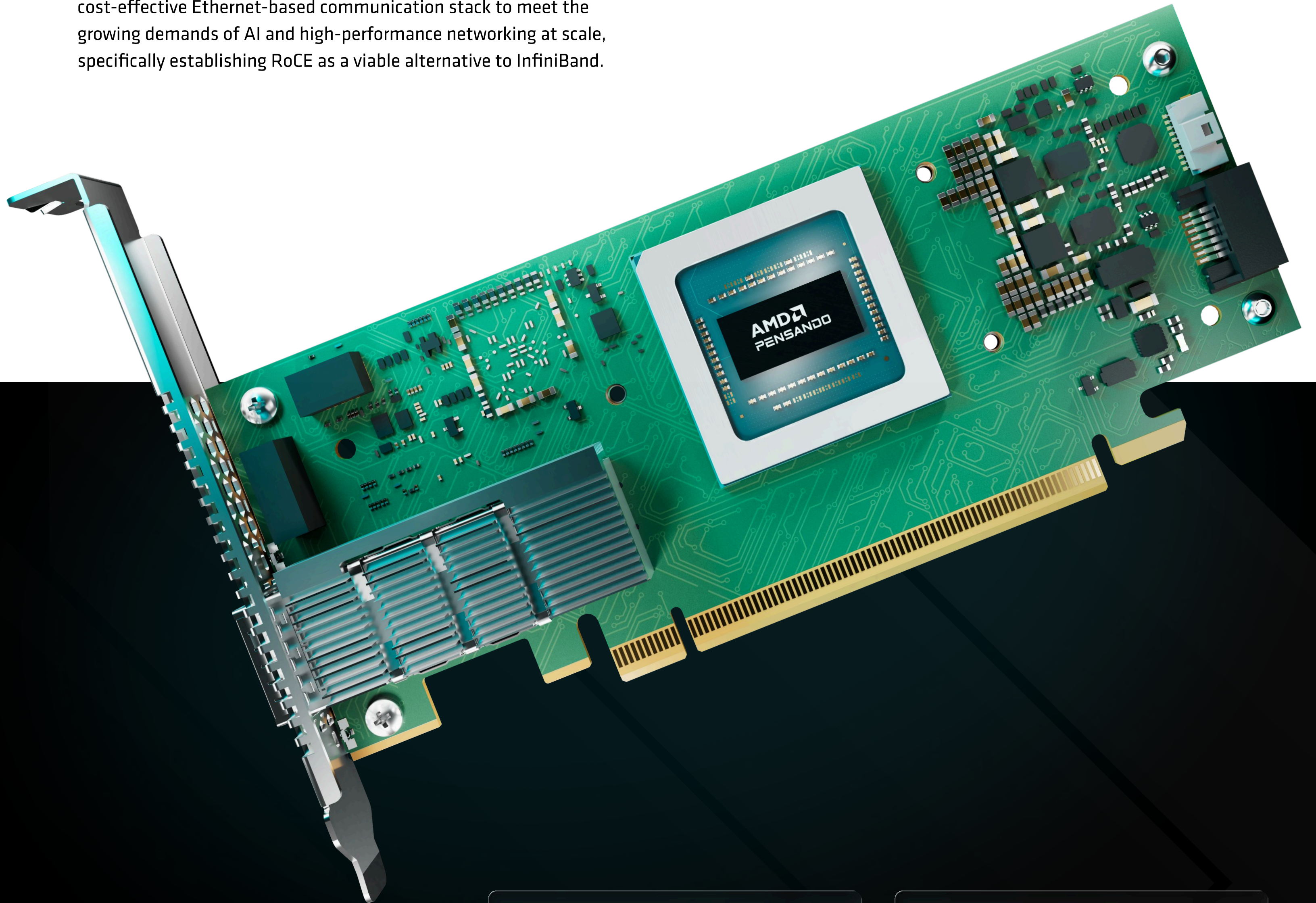


5. AN OPEN ECOSYSTEM

In today's dynamic and evolving AI landscape, remaining free of the constraints of proprietary technologies and retaining freedom of vendor choice can give enterprises a competitive edge. This freedom is a cornerstone of the AMD philosophy and the root of our commitment to an open ecosystem, as evidenced in our lead role in the founding of the UEC.

The consortium's objectives include developing an open, scalable, and cost-effective Ethernet-based communication stack to meet the growing demands of AI and high-performance networking at scale, specifically establishing RoCE as a viable alternative to InfiniBand.

As the first AI NIC to support **UEC standards**, the AMD Pensando™ Pollara 400 AI NIC works with any UEC-compliant network switches. Want to mix and match NICs from different vendors? Or switch out a tried and true for something new? Have at it. The AMD Pensando Pollara 400 AI NIC can grow with your infrastructure as you build it out.



KEY BENEFITS:



Achieve up to 16% lower network infrastructure CapEx⁸.



Preserve customer choice and avoid vendor lock-in.

6. AN AI NIC PURPOSE-BUILT FOR GENAI-SPECIFIC INFRASTRUCTURE

Of note, there's a trend among enterprises toward building and deploying custom GenAI-specific infrastructure on premises due to considerations of cost, data privacy, and governance constraints. A 2025 study by **Enterprise Strategy Group** found that 89% of organizations have built and deployed their own custom GenAI infrastructure on premises or will do so within the next 24 months⁹.

In addition to enterprises and cloud service providers, this trend presents opportunities for providers of neoclouds – emerging cloud platforms with a strong focus on GPU infrastructure and performance that are purpose-built for AI-era workloads. On-premises infrastructure is also favored by sectors like finance, government, and healthcare, where data sovereignty and increasingly sovereign AI – controlling where data and infrastructure live and how they're managed – are critical to compliance and trust.

Building on-premises infrastructure for GenAI and agentic AI carries very specific technical requirements that include connecting servers and storage both within back-end networks and to the rest of the IT infrastructure outside of the clusters through the front-end network (the network's user-facing portion).

The AMD Pensando™ Pollara 400 AI NIC has both back-end and front-end use cases, serving both as an RDMA NIC for accelerating scale-out and optimizing data transfer within back-end AI networks and as a host NIC for connectivity in front-end networks. This makes the solution a powerful component in the design of on-premises GenAI-specific infrastructure, unlocking exceptional performance and efficiency.

KEY BENEFIT:



Suited to on-premises GenAI-specific infrastructure with back-end and front-end use cases.

7. UNPRECEDENTED, FULLY PROGRAMMABLE ADAPTABILITY TO MEET EMERGING NEEDS

Whether scaling AI systems, tuning them to adopt new innovations or standards, or customizing transport protocols to handle new AI workloads, organizations reap multiple benefits from being able to program their network infrastructure. The AMD Pensando Pollara 400 AI NIC is the first **fully programmable AI NIC**, making infrastructure upgrades as simple as a software update.

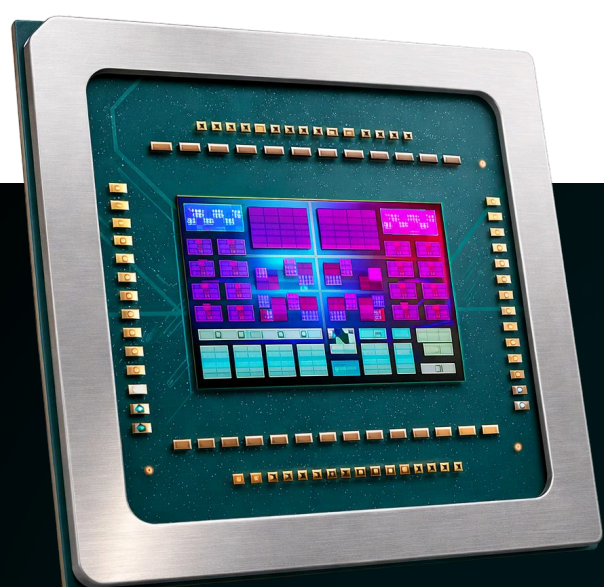
This versatility derives from the proven third-generation P4 engine that the AMD Pensando Pollara 400 AI NIC is built on. A compact, fully programmable unit designed to optimize network workloads, the P4 allows customers to customize network behavior, implement bespoke RDMA transports, and optimize performance for specific AI workloads, while maintaining compatibility with UEC standards.

This helps ensure that AI infrastructures can evolve as quickly as the AI technologies they support. Programmable network infrastructure also helps ensure future and backwards compatibility across generations so enterprises can innovate without the cost or disruption of rip-and-replace cycles.

Importantly, this architectural approach creates a seamless path to the next generation programmable AI NIC, “Vulcano.” Built on an advanced 3nm process and designed for 800G line-rate performance, “Vulcano” offer higher scale out bandwidth per GPU, while supporting emerging I/O technologies like UAL and PCIe® Gen6. It also maintains forward and backward software compatibility with the AMD Pensando Pollara 400 AI NIC, helping ensure that customers deploying the AI NIC today are investing in a networking architecture designed to scale with next generation AI systems and rack-scale platforms such as AMD “Helios” AI Rack Solutions.

P4 programmable architecture allows organizations to:

- Rapidly adopt software innovations without costly infrastructure rebuilds.
- Quickly adopt new UEC standards to become future ready.
- Simplify implementation of UEC RDMA or RoCEv2 transports as well as congestion control protocols optimized for individual AI training cluster deployments.
- Enable compatibility across future and backward generations so enterprises can innovate without the cost or disruption of rip-and-replace cycles.



FULLY PROGRAMMABLE OPEN ECOSYSTEM

Optimizing GPU-NIC Performance for Large Clusters

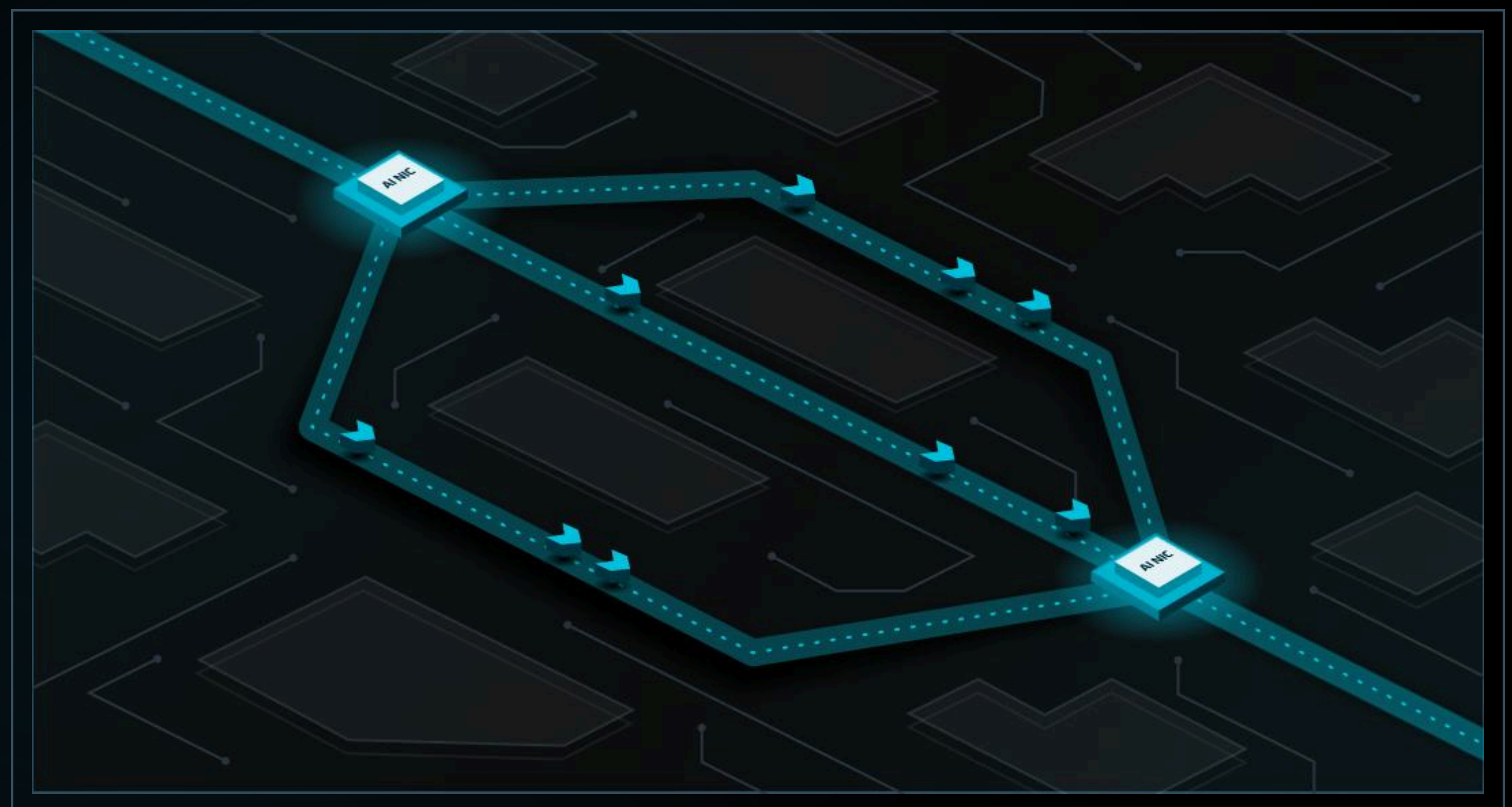
KEY FEATURES OF THE AMD PENSANDO™ POLLARA 400 AI NIC

1. PERFORMANCE

- **400 Gbps Connectivity:** Lossless and lossy fabric support for flexible, high-throughput deployment.
- **In-order Message Delivery:** Maintains sequencing across multipathing, protecting AI workload integrity.

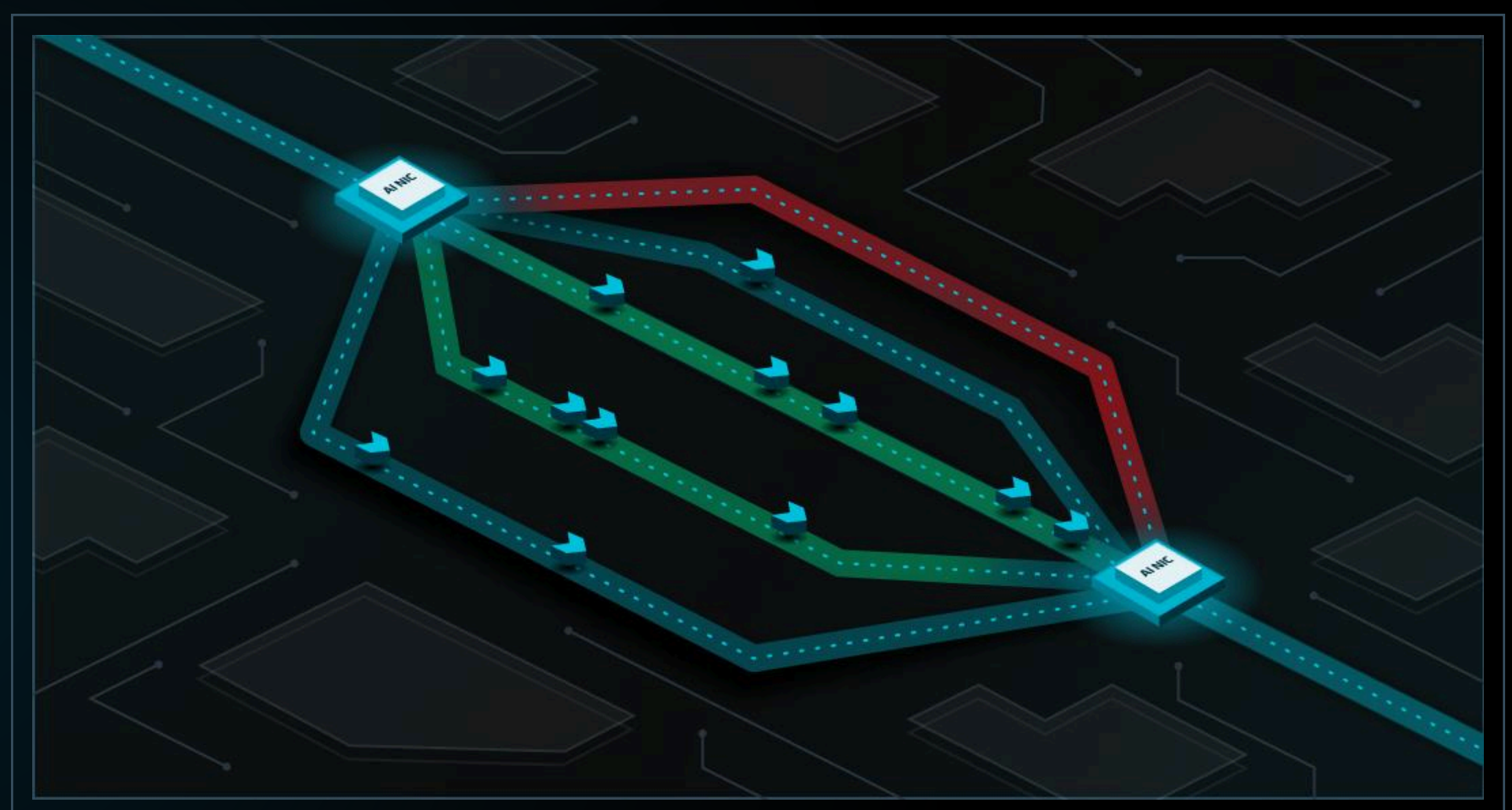
INTELLIGENT PACKET SPRAY

Distributes packets efficiently across multiple paths, maximizing utilization and scalability.



PATH-AWARE CONGESTION CONTROL AND DYNAMIC LOAD BALANCING

Proactively reroutes traffic around congestion, helping prevent performance cliffs and simplifying network performance management.



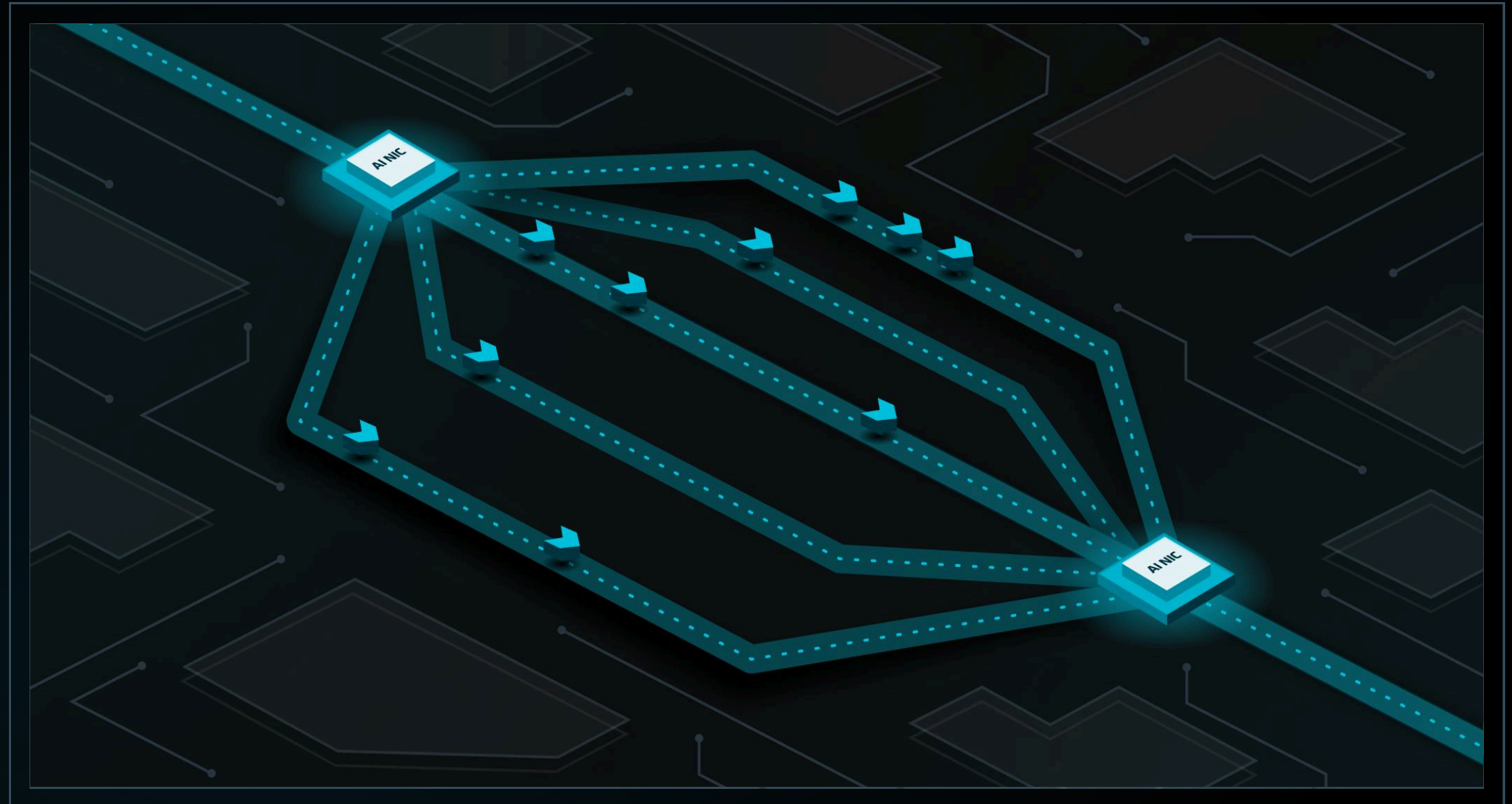
KEY FEATURES OF THE AMD PENSANDO™ POLLARA 400 AI NIC

OUT-OF-ORDER PACKET HANDLING

Direct placement into GPU memory without buffering enables fast data throughput.

IN-ORDER MESSAGE DELIVERY

Maintains sequencing across multipathing, protecting AI workload integrity.



SELECTIVE RETRANSMISSION

Only retransmits dropped or corrupted packets, helping reduce recovery latency and avoid redundant traffic.



2. COST EFFECTIVE AND SCALABLE

- **Multiplane Fabric Architecture:** Localizes faults and isolates service issues, significantly improving RAS and supporting the cost-effective scale-out of AI networks.
- **Open Ecosystem Integration:** Leverages full programmability to support UEC standards, enabling broad interoperability and maintaining freedom of vendor choice.
- **P4 Programmability:** Customize transports (UEC, RDMA, RoCEv2, custom protocols) without costly rebuilds.
- **Roadmap Continuity:** Establishes a seamless architectural path to our next generation programmable AI NICs, “Vulcano,” enabling smooth progression to higher bandwidth, scalable Ethernet based AI fabrics.

3. RELIABILITY, AVAILABILITY, SERVICEABILITY (RAS)

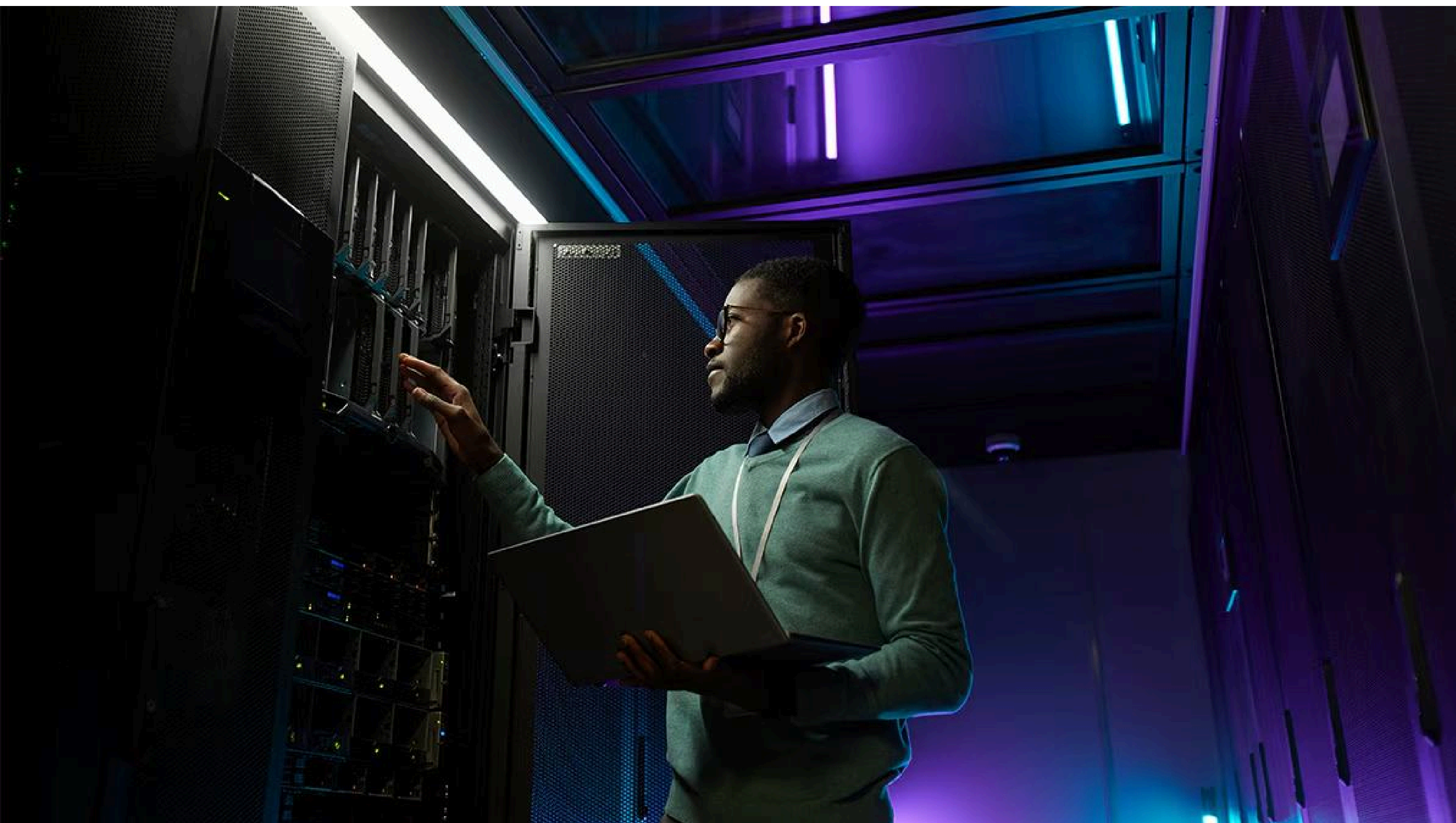
- **Rapid Fault Detection:** Real-time visibility into latency, congestion, and drops; near-instant failover helps reduce GPU downtime.
- **Advanced Telemetry and Observability:** Proactive monitoring, up to 50% cluster uptime improvement².
- **High-Availability Failover and Loss Recovery:** Maintains uptime under faults with fast detection and recovery.
- **Unified Support Model:** Simplifies Day-2 operations and troubleshooting across complex deployments.

KEY BENEFITS:

50%

Reliability, Availability, Serviceability (RAS)²

Better than competition.



ADVANCING THE FUTURE OF AI NETWORKING

With organizations pushing ever more complex AI into production, the **AMD Pensando™ Pollara 400 AI NIC** delivers the performance, scalability, and reliability required for modern AI infrastructure. More than a network interface, it serves as a foundational component in largescale AI systems—accelerating job completion times, improving cluster uptime, and reducing network cost while aligning with open Ethernet based fabrics and standards such as those defined by the UEC. As the world’s first UEC-ready AI NIC within an expanding RoCE ecosystem, the AMD Pensando Pollara 400 AI NIC enables measurable improvements in throughput, operational efficiency, and resiliency using familiar Ethernet frameworks. Its fully programmable design also sets the stage for the next evolution in AMD networking.

The “Vulcano” programmable AI NIC designed for 800G line rate performance, built on an advanced 3nm process, supporting UAL and PCIe® Gen6, and delivering up to 8 times more scale out bandwidth per GPU*, maintains full forward and backward software compatibility with the AMD Pensando Pollara 400 AI NIC.

This continuity helps ensure today’s AMD Pensando Pollara 400 AI NIC deployments can seamlessly evolve toward next generation AI systems and rack-scale architectures such as AMD “Helios” AI Rack Solutions.

Whether deployed as an RDMA NIC for backend AI fabrics or as a high-performance host NIC in frontend networks, the AMD Pensando Pollara 400 AI NIC reduces operational complexity and minimizes the learning curve for networking teams.

To discover how the AMD Pensando™ Pollara 400 AI NIC can help maximize the value of AI infrastructure investments and prepare your organization for what’s next, request a consultation.

**TO DISCOVER HOW THE
AMD PENSANDO™ POLLARA 400 AI NIC
CAN HELP MAXIMIZE THE VALUE
OF YOUR AI INVESTMENTS AND
PREPARE YOUR ORGANIZATION FOR
WHAT’S NEXT**

[Request a Consultation](#)



Sources:

1. PEN-013: Testing was conducted by AMD Performance Labs as of 3/30/2025 on the Pollara 400 vs Broadcom Thor2 on a test system using identical GPU cluster configurations (16 SuperMicro server nodes, 128 AMD Instinct™ MI300 GPU, 128 Pollara or Thor2 NICs, rail based network topology using 64-port x 400G Broadcom Tomahawk5 based Ethernet switching).

CPU on all are: 2P Intel® Xeon® Platinum 8468

Memory: 2048G DDR5 4800mhz 64GB dual rank dims, OS: Ubuntu® 22.0.4, Kernel 6.5.0-45-generic LTS, 2 rear m.2 NVME, Bios version: 2.3.5., ROCm™ 6.3.0-39.

Results may vary due to factors including but not limited to software versions, network speeds and system configurations.

2. PEN-019: Testing conducted by AMD Performance Labs as of [15 September 2025] on the AMD Pensando Pollara AI NIC, on a test system comprising of SMC-300X server for GPU-GPU communication: 2x AMD Pensando Pollara AI NIC, 2P AMD EPYC 9454 48-Core -2P Processor, 8x AMD Instinct MI300X GPU, Ubuntu 22.04.5 LTS, kernel 5.15.0-139-generic, ROCm 6.4.1.0-83-69b59e5. Testing running Llama-3.1-8B, Model Configuration: SEQ_LEN=2048, TP=1, PP=1, CP=1, FP8=1, MBS=10, GBS = 5120. Iteration = 2, No. of paths/QP : 128. Results may vary based on factors including but not limited to system configuration and software settings.

3. Dubey, Abhimanyu & Jauhri, Abhinav & Pandey, Abhinav & Kadian, Abhishek & Al-Dahle, Ahmad & Letman, Aiesha & Mathur, Akhil & Schelten, Alan & Yang, Amy & Fan, Angela & Goyal, Anirudh & Hartshorn, Anthony & Yang, Aobo & Mitra, Archi & Sravankumar, Archie & Korenev, Artem & Hinsvark, Arthur & Rao, Arun & Zhang, Aston & Zhao, Zhiwei. (2024). **The Llama 3 Herd of Models**. 10.48550/arXiv.2407.21783. Table 5, Page 13. Claim reflects technology used in AMD Pensando Pollara 400 NICs, however testing and data not specific to the AMD Pensando Pollara 400 AI NIC. Results may vary.

4. PEN-015 - Testing conducted by AMD Performance Labs as of [13th May 2025] on the [Pollara AI NIC and Nvidia CX7 NIC], on a test system comprising of 8 Nodes of 8xMI300X AMD GPUs (64 GPUs); Broadcom Tomahawk-5 based leaf switch (64x800G) Model Dell z9864f-r0; RAIL Topology; AMD AI NIC Pollara – 64 NICs, ROCm™ version 6.3.2.0-66-cbc70b5 OR Nvidia CX7 SmartNIC - 64 NICs, RCCL version 2.24.3-develop:7961624; CPU Model in each of the 8 Node - Dual socket AMD EPYC 9454 48-Core Processor; Operating System Ubuntu® 22.04.5 LTS; Kernel 5.15.0-139-generic. All the application software libraries (RCCL and ROCm) and test environment are exactly same except the low-level drivers which are hardware specific. For Nvidia CX7 card the drivers are installed from openly available Linux drivers' installation methodology link on NVIDIA website.

<https://docs.nvidia.com/networking/display/connectx7vpi/linux+driver+installation>

For AMD Pensando Pollara NIC.

The drivers are used from internal build, but they are planned to be publicly available in coming months.

Following collective communication operations were measured.

Allreduce, Allroall, Alltoallv, Broadcast, Reduce, Scatter, Allgather.

5. PEN-016 - Testing conducted by AMD Performance Labs as of [28th April 2025] on the [AMD Pensando™ Pollara 400 AI NIC], on a production system comprising of: 2 Nodes of 8xMI300X AMD GPUs (16 GPUs): Broadcom Tomahawk-4 based leaf switch (64x400G) from MICAS network; CLOS Topology; AMD Pensando Pollara AI NIC – 16 NICs; CPU Model in each of the 2 nodes - Dual socket 5th gen Intel® Xeon® 8568 - 48 core CPU with PCIe® Gen-5 BIOS version 1.3.6 ; Mitigation - Off (default).

System profile setting - Performance (default) SMT- enabled (default); Operating System Ubuntu 22.04.5 LTS, Kernel 5.15.0-139-generic.

Following operation were measured: Allreduce.

Average 25% for All-Reduce operations with 4QP and using UEC ready RDMA vs the RoCEv2 for multiple different message size samples (512MB, 1GB, 2GB, 4GB, 8GB, 16GB). The results are based on the average at least 8 test runs.

6. PEN-018: AMD comparison and pricing as of July 6, 2025, for network fabric costs to support 128,000 GPUs. Comparison of a Pollara NIC with multiplane fabric and packet spray on an 800G Tomahawk 5-based multiplane design versus a generic fat-tree fabric built on fully scheduled, big-buffer (Jericho3/Ramon3) 800G switching platforms. The generic system is assumed to use a competitive NIC, with NIC costs considered comparable.

6. PEN-018: The Pollara-based design is estimated to deliver up to 58% network switching cost savings by enabling the use of more cost-effective Tomahawk 5-based switching in a multiplane architecture. AMD comparison and pricing as of 4/23/2025 of a Tomahawk 5 system with Pensando Pollara NIC featuring exclusive multiplane fabric and packet spray versus a generic big-buffer 800G switching platform; the generic system would employ a competitive NIC, costs of NICs are assumed to be comparable. Deploying Pollara with multi-fabric support and packet spray, allows customers to build cost-effective multiplane network fabrics, instead of a fat-tree design using less network switches to deliver the same amount of network bandwidth across the fabric, and dramatically reducing both switch platform cost, and cost associated with cables, optics.

- Fat-Tree Big Buffer Fully Scheduled Network (Leaf/Spine/Core) estimated Cost: \$1.22B;
- 3556 leaf (Jericho3-AI) units at \$104,998 each = \$373M;
- 1557 spine/core (Ramon3) units at \$147,998 each = \$247M;
- 128K AOC-10m cables at \$1059 each = \$136M;
- 568,889 QDD-SR4-400G transceivers at \$819 each = \$466M;
- Total (Switching & Optics) = \$1.22B;
- Naddod Tomahawk5 800G Multiplane Fabric Network estimated Cost: \$511M;
- 3,000 Leaf and Spine Units (Naddod N9600-640C) at \$26,999 each = \$81M;
- 384K (QDD-SR4-400G) transceivers at \$819 each=\$313M;
- 64K Switch (OSFP-2x400G-DR4) transceivers for NIC connections at \$759 each = \$48M;
- 256K MPO Cables at \$26 each = \$6.6M;
- 2K Optical Shuffle box, modules and internal cables at \$30K per rack =\$60M;
- Total (Switching & Optics) = \$511M.

Prices subject to change. Comparison for specific network configurations only and may not be representative of all possible network configurations and comparisons.

7. InfiniBand Scale Size 48,000 nodes:

https://network.nvidia.com/pdf/whitepapers/InfiniBandFAQ_FQ_100.pdf

100K RoCEv2:

<https://www.tomshardware.com/desktops/servers/first-in-depth-look-at-elon-musks-100-000-gpu-ai-cluster-xai-colossus-reveals-its-secrets>

UEC Scales to 1,000,000 nodes:

https://network.nvidia.com/pdf/whitepapers/InfiniBandFAQ_FQ_100.pdf

8. PEN-014: AMD comparison and pricing as of 4/23/2025 of a generic big-buffer Jericho3 based DEF Leaf-Spine Network versus a Naddod Tomahawk5 leaf-spine Network. The Tomahawk system would employ an AMD Pensando™ Pollara NIC in the AI network; the generic system would employ a competitive NIC, costs of the NICs are assumed to be comparable.

Naddod Tomahawk5 800G Leaf-Spine Network Cost: \$26.65M

188 Leaf & Spine units (Naddod N9600-640D) at \$26,999 each = \$5,075,812

[https://www.naddod.com/products/102322.html?](https://www.naddod.com/products/102322.html?srsltid=AfmBOor58tjyTO2mhb8ISaa3A13INrL_RSIS4BOHYTsCQN6X99fj-tPZ8K)

<https://www.fs.com/products/146389.html>

16K SR4 (QDD-SR4-400G) optics at \$819 each = \$13,104,000

[https://www.fs.com/products/226577.html?](https://www.fs.com/products/226577.html?attribute=94273&id=3632994)

[attribute=94273&id=3632994](https://www.fs.com/products/226577.html?attribute=94273&id=3632994)

Comparison Large Buffer, Fully Scheduled 800G Leaf-Spine (Jericho3-AI/Ramon3) Network
Cost: \$31.84M

Network comprised of 222 units of Leaf-Spine Network Cost: \$31.84M;

8K AOC 10m cables at \$1059 each = \$8,472,000

17,778 SR4 Opticss @ \$819 each = 14,560,182

Cost Savings (%): 16%

8. PEN-014: Prices subject to change. Comparison for specific network configurations only and may not be representative of all possible network configurations and comparisons.
9. Frey, Jim; Marsh, Emily. "Research Report: Architecting the Network for an AI-powered World," Enterprise Strategy Group, February 2025: <https://research.esg-global.com/reportaction/515201976/Marketing>.

*Preliminary specification subject to change