WEKA® × Physical Intelligence (π)

CASE STUDY

Physical Intelligence Ignites a Robotics Revolution

Physical Intelligence (Pi) is bringing general-purpose AI into the physical world by developing foundation models and learning algorithms to power the robots of today and the physically actuated devices of the future. Physical Intelligence focuses on building a single generalist robot intelligence that can control a wide range of robots and perform various tasks of varied types and complexities.

Challenges

- Massive, diverse training data set
- Unused local GPU-attached NVMes
- Long model checkpoints
- Rigid Multi-Cloud Data Management

Solution

NeuralMesh™ Axon™ by WEKA® on Oracle Cloud Infrastructure

Benefits

- 10-15% faster model checkpoint times
- 80% reduction in data infrastructure costs
- Portability across diverse deployment types

Key Features

- NVMe Flash Storage
- Intelligent Tiering
- Adaptive Caching
- Flexible Deployment Options

In October 2024, the Physical Intelligence team released their first general-purpose robot foundation model, π0. This is a first step toward the long-term goal of developing artificial physical intelligence so that users can simply ask robots to perform any task they want, just like they can ask large language models (LLMs) and chatbot assistants. The foundation model can control a variety of different robots and can either be prompted to carry out desired tasks or fine-tuned to specialize in more specific complex scenarios.

Challenge:

Unused Local GPU-Attached Memory

Training a model to perform such a wide range of tasks requires massive amounts of data. The full training mixture for the π0 prototype includes both a large, diverse dataset of dexterous tasks that the PI team collected across a fleet of robots, and external data. The Physical Intelligence team is focused on an infrastructure strategy that prioritizes performance, ease of use, and flexibility. Pi built and trained early models in Oracle Cloud (OCI). However, the data storage common solutions were not sufficiently matched with the performance criteria. Long model checkpoint times delayed model training times by 10-15%. While exploring options to accelerate the underlying storage, the team quickly became convinced that leveraging the local flash (NVMe) storage attached to the GPU compute nodes in the training cluster offered the most promising path forward. Other solutions required significant investment in configuration and set-up, versus the out- of- the- box functionality with NeuralMesh by WEKA. Finally, Pi needed a data management solution that could work across their current multi-cloud deployment, and in the future, possibilities for diverse deployment types, including on-premises.

The Solution:

NeuralMesh Axon for Oracle Cloud

Physical Intelligence uses NeuralMesh Axon as its storage solution for distributed training of their robotics foundation models. The petabyte-sized single NeuralMesh namespace spans both the NVMe-based performance tier and Oracle Cloud Object storage capacity tier. Intelligent Tiering automatically manages data movement between the high-performance NVMe tier and the cost-effective Oracle Cloud object storage tier, eliminating the need for manual data management. Model training occurs on Oracle Cloud Infrastructure (OCI) bare- metal instances using NeuralMesh Axon for OCI. In NeuralMesh Axon, the data storage environment resides on the same physical infrastructure as the GPU-accelerated training environment. This innovative approach enables Physical Intelligence to leverage the existing local NVMe resources attached to the GPU/CPU resources in the cluster.

“WEKA turns unused local NVMes into a high-performance shared storage system that is resilient and available.”

The Impact:

Increased Data Performance

With NeuralMesh Axon, the Physical Intelligence team improved data performance across multiple dimensions. Model checkpoint times have collapsed, and performance is predictable because the team is able to completely parallelize the checkpoint write process. Overall, model training times are now 10-15% faster. The critical enabling capability is the use of local NVMe resources and adaptive caching. This effectively creates an extended memory layer shared across the entire training cluster, creating a persistent, resilient storage layer with memory (sub-millisecond) speeds.

Reduced Data Costs

The highly resource-efficient approach WEKA provides has enabled Physical Intelligence to get the most out of their infrastructure investments, and helped reduce data storage costs by 80%.

Simplified Data Management

While several approaches exist to use local NVMes for AI model training, they typically require complicated network configurations or manual data operations to perform data caching or pre-fetching operations. With NeuralMesh Axon, the Physical Intelligence team is able to eliminate all these manual tasks and complicated architectures. NeuralMesh Axon streamlines checkpoint management and makes checkpoints accessible across different deployment types. The Physical Intelligence team is excited about NeuralMesh Axon data resilience and availability features to survive even multiple cluster nodes dropping offline (a common occurrence in large-scale GPU training clusters).

Overall, management of NeuralMesh Axon is minimal and can be done by IT generalists or AI practitioners, not just storage specialists.

Unlimited Flexibility, Increased Resource Utilization

WEKA enables the Physical Intelligence infrastructure strategy to remain highly flexible and agile. The Pli team was happy to find that NeuralMesh Axon works within a broad ecosystem and supports diverse deployment types. This ensures the PI team can shift training from one deployment model to another as needed to support the business's needs.

“Since we are a lean team dedicated to research, we needed a data storage solution that works out of the box for large-scale AI training and data management—handling the heavy lifting and the small details for us so we can stay focused and move fast.”

About NeuralMesh™ by WEKA®

NeuralMesh™ by WEKA® is the world's only storage system purpose-built for accelerating AI at scale. But it's more than just storage. It's a high-performance, flexible foundation for enterprise AI and agentic AI innovation. WEKA's software-defined, containerized microservices architecture doesn't just handle scale—it feeds on it, getting faster, more efficient, and more resilient at exabytes and beyond. Designed for ultimate flexibility, it adapts effortlessly to any deployment strategy from bare metal to multi-cloud and everything in between. NeuralMesh helps enterprises, AI cloud providers, and AI builders achieve unmatched real-world performance, maximize GPU utilization, and lower the cost of innovation.