Pong Yuen uses AI and the metaverse to improve decision making

■ NetApp

PY PONG YUEN

Using AI and the metaverse to help businesses make better decisions and mitigate risks.

Decision makers want to make predictions about the future to help them assess risk. Simulations in the metaverse allow organizations in different industries to make predictions in the simulated world, accelerate decision making, and manage risk. Pong Yuen Holdings Limited, established in 1991, is a leading service provider and expert in artificial intelligence of things and machine learning solutions.

"The metaverse is not simply about social networking and gaming," says Anri Kitami, software engineering specialist at Pong Yuen. "In fact, the metaverse is extremely useful across different industries, especially when it comes to decision making and risk management."

6 times the training performance

"It simplifies and accelerates AIOps, as well as greatly reduces the time to market."

Anri Kitami
Software Engineering Specialist at Pong Yuen

Searching for a more effective AI infrastructure
Pong Yuen is dedicated to guiding organizations through digital transformation and helping them gain a competitive advantage from the metaverse.

"One of our customers simulates road signs, lane markings, obstacles, and other vehicles in the virtual highway, so they can capture all synthetic data to develop and train autopilot machine learning models for self-driving cars," says Kitami.

"We can also simulate a fire accident in a large building in the metaverse, so we can train machine learning models for intelligent fire detection and evacuation, as well as construction risk management."

To speed up analytics, training, and inference, Pong Yuen requires an effective AI infrastructure.

Removing the complexity and guesswork for an optimized AI infrastructure
Traditionally, assembling and integrating AI infrastructure involves computing, storage, and networking, as well as software components and interdependencies, which makes designing an effective AI infrastructure challenging.

NetApp® ONTAP® AI is a proven architecture powered by NVIDIA DGX A100 servers. This architecture enables organizations like Pong Yuen to customize and deploy AI with preconfigured solutions that are built on a field-proven, validated reference architecture and sized for the most common use cases.

"It is crystal clear to us that only NetApp ONTAP AI can allow us to build a certified and approved AI cluster as it is a proven architecture researched by NVIDIA and NetApp," says Kitami. "It consolidates a data center's worth of analytics, training, and inference computing into a single, integrated system. We can easily get our AI and machine learning workload up and running without worrying about setting up our own AI infrastructure." With ONTAP AI, Pong Yuen eliminated the guesswork as well as the complexity of design and management.

Seamless and resilient ML operations
Experimentation leads to improved results in AI projects. Data scientists at Pong Yuen use different versioning of models and datasets, in order to come up with different combinations of configurations and model architectures. Generating a vast number of copies of models and datasets can be challenging and time-consuming. By leveraging NetApp Snapshot™ and FlexClone® technologies, Pong Yuen can create copies for machine learning regardless of the dataset size in just a couple of seconds, rather than hours.

Even in a worst-case scenario like the corruption or accidental deletion of the current model, the data science team doesn't have to start the project over from the beginning. They can revert the application and data and continue to develop their machine learning projects.

Speedy results with a high-performance solution
AI model training routines demand massive amounts of computing power and storage. With NetApp all-flash storage, you can expect to get more than 2GBps of sustained throughput (5GBps peak). The state-of-the-art storage with unparalleled data-reduction ratio can handle thousands of training images per second. ONTAP AI is a best-in-class data solution that delivers ultra-high training performance and keeps up with the most demanding training workloads.

Fast time to market and return on investment

ONTAP AI delivers maximum performance for efficient AI model training to achieve faster time to market, as well as higher return on investment in AI infrastructure.

According to Kitami, "It simplifies and accelerates AIOps, as well as greatly reduces the time-to-market." Compared with CPU systems, ONTAP AI requires a fraction of the physical space and power, while reducing overall AI model training costs.

Optimized and secured Kubernetes cluster

Pong Yuen leveraged the NetApp AI Control Plane to simplify AI data management by integrating Kubernetes and Kubeflow with a data fabric enabled by NetApp. This integration allows Pong Yuen's data science team to easily perform numerous data management tasks while using the same tools to securely control and protect their data in flight, in use, or at rest, and to meet compliance requirements with confidence. If any issues arise, NetApp's proven support model can provide guidance and help Pong Yuen troubleshoot.

Future-proof AI solution

NetApp ONTAP AI is one of the first converged infrastructure stacks to incorporate the NVIDIA DGX A100 system, which sets a new gold standard for computing density for AI infrastructure. The infrastructure offers scalability and data freedom from edge to core to cloud to support Pong Yuen's growth and changing business needs. ONTAP AI offers the flexibility to scale and embrace cloud AI innovations now and in the future.

"Partnering with NetApp allows us to democratize the capability of building a metaverse and to unleash their power of data for our customers." Kitami said.

Learn more here.

NetApp products

AI Control Pane

ONTAP AI

FlexClone

Snapshot

Partner products

NVIDIA Ampere

NVIDIA DGX A100

+1 877 263 8277

About NetApp

In a world full of generalists, NetApp is a specialist. We're focused on one thing, helping your business get the most out of your data. NetApp brings the enterprise-grade data services you rely on into the cloud, and the simple flexibility of cloud into the data center. Our industry-leading solutions work across diverse customer environments and the world's biggest public clouds.

As a cloud-led, data-centric software company, only NetApp can help build your unique data fabric, simplify and connect your cloud, and securely deliver the right data, services and applications to the right people—anytime, anywhere. To learn more, visit www.netapp.com

NetApp