CIO

by **Paula Rooney**
Senior Writer

CIOs look beyond 'Big 3' cloud providers for AI innovation

Feature
Feb 13, 2025 • 7 mins
Artificial Intelligence GPUs IaaS

Athos Therapeutics' decision to go niche with a GPU-as-a-service from upstart Vultr could point the way forward for similar organizations with specialized AI needs.

CREDIT: GORODENKOFF / SHUTTERSTOCK

AWS, Microsoft, and Google may continue to dominate the enterprise cloud market, but a raft of second-tier cloud providers are proving to be valuable partners for organizations and innovators with specialized workloads and use cases — especially in the burgeoning AI era.

Athos Therapeutics is one such enterprise. Having previously run its AI workloads on premises, the clinical-stage biotechnology company chose Vultr Cloud GPU, a GPU-as-a-service that Athos runs on Dell's Nvidia-powered HGX H100 PowerEdge servers to train its homegrown AI models and build its AI-powered precision medicine platform aimed at solving auto-immune and cancerous diseases.

Before signing the deal last summer, the trio — Athos, Vultr, and Dell — collaborated on a pilot project that met Athos' needs, says June Guo, vice president of AI and ML at Athos.

Guo cited several reasons why Athos abandoned its small on-premises infrastructure in favor of Vultr's specialized cloud platform, including Athos' huge on-prem electricity and internet use as well as the need for security to train its large models. The move to Vultr has also proved less expensive, he says.

"We used to have our on-premises servers for training these AI algorithms but infrastructure problems such as power outages and internet bandwidth issues [prompted us to start] looking for a vendor to address these issues," Guo says, noting that the switch to a GPU-as-a-service model has helped accelerate the development of Athos' precision therapeutics platform as well.

Niche needs, niche solutions

Vultr, IBM, Alibaba, Akamai, OVHcloud, Tencent, and Huawei are among a handful of second-tier cloud IaaS players offering cost advances and the flexibility of AI model training, tuning, and inference services, as well as required security and confidentiality for intellectual property, genomics data, and proprietary AI algorithms, analysts say.

"Second-tier cloud providers like Vultr, Akamai, IBM, Alibaba, Tencent, and Huawei differentiate themselves by focusing on niche markets and specific customer needs, as opposed to the hyperscalers' broad, one-size-fits-all approach," says David Linthicum, a cloud and AI thought leader who previously served as managing director and chief cloud strategy officer at Deloitte Consulting. "These providers thrive in areas where specialization, flexibility, and cost efficiency matter most."

Intellectual property protection was a significant reason behind Athos' move to Vultr's GPU cloud, Guo says, as doing so would better protect its model IP, while conforming with industry regulations and compliance. The GPU-as-a-service model also minimizes the constantly evolving maintenance requirements of an AI infrastructure, including downloading massive amounts of genomics data, internet updates, and swapping Nvidia cards in and out, he says.

Athos could have opted for one of the big three hyperscalers — AWS, Google Cloud, or Microsoft Azure — but training its algorithms and scaling various types of scientific omics data would be prohibitively expensive on those platforms, Guo says.

Because Athos' data is derived from human blood, stool, and tissue samples, and because multi-omics data structures are not like text, audio, or video, Athos cannot use standard gen AI models. The sheer variety and volume of data used for precision therapeutics requires Athos to build its own AI algorithms and AI models, which it may commercialize to other biotechnology and pharmaceutical companies when fully baked.

Additionally, Vultr and Dell were able to provide more cost-effective engineering support for Athos' autonomous AI analytical platform and data lake than most cloud vendors can provide, Guo says.

For Athos, Vultr's Cloud GPU platform does it all — runs Athos' AI model training, finetuning, and inferences services — and that's not true of all clouds, he says. It also provide required security and confidentiality for Athos' intellectual property, large genomics data sets, and proprietary AI algorithms.

Athos Therapeutics, which is partnering with the Cleveland Clinic and Lahey Hospital & Medical Center, has developed multiple models with billions of parameters and multi-omics data and has performed analysis of more than 25,000 patients. The Vultr-Dell cloud partnership has helped pave the way.

"Compared to the on-premises servers that we used before; the acceleration is amazing. We have 10 times the acceleration that we had before," Guo says.

Options for innovation abound

The AI gold rush has brought with it the rise of niche players offering pureplay GPU-as-a-service, as well as second-tier providers who host a fuller menu of cloud options, much like the big three hyperscalers, including GPUs on the cloud. IDC analyst Dave McCarthy says GPU-as-a-service has emerged as a new term with the debut of startups like CoreWeave that offer the GPU service model exclusively, but larger players, such as AWS, Microsoft, and Google, have been offering it for many years.

Vultr, McCarthy says, is "more akin" to the big three hyperscalars because it offers a range of cloud services besides GPU-as-a-service. Vultr, which launched in 2014, was founded by developers who designed its early cloud platform to offer a unique control pane and automation to help customers scale rapidly. The platform was originally very popular with gaming companies and SaaS startups.

As Vultr grew, scaled up, hired a C-suite, and formed partnerships with Nvidia and Dell, the company now targets enterprises globally with flexible, scalable, global cloud compute, cloud GPU, bare metal and cloud storage solutions that support AI workloads.

For instance, Vultr's CPU-a-a-service offers access to AMD's Instinct MI300X Accelerator and Nvidia GPUs, including GH200, HGX H100, A100, L40S, A40, and A16 for AI/ML and high-performance computing,

Cloud expert Linthicum points out that Vultr appeals to developers, startups, and SMBs with its straightforward pricing and simplicity, while Akamai dominates in edge computing and content delivery for media, gaming, and latency-sensitive applications. IBM, Linthicum adds, leads in hybrid cloud solutions, targeting enterprises managing legacy systems with a focus on compliance-heavy industries such as healthcare and finance in the US and EU.

Geographically, Alibaba, Tencent, and Huawei are dominant players in Asia, he says.

For US and European companies, the appeal of second-tier cloud providers often lies in vendor diversity, cost savings, and the ability to support AI and specialized workloads that do not require the scalability and complexity of AWS, Google Cloud Platform, or Azure, Linthicum adds.

Accenture's Cloud First analyst Sulabh Srivastava says second-tier cloud players offer a variety of specialized services for SMBs, regional companies, and regulated industries with key benefits such as lower costs for AI/ML training and inferencing use cases. Srivastava says the cost of a hyperscaler's GPU computing infrastructure to train AI models is commonly much higher than the same resources provided by specialty providers such as Vultr.

For its part, Vultr is working to expand its footprint in the enterprise and only time will tell if the larger hyperscalers lower their prices to snag more of the AI model training business.

"For enterprise segments, Vultr has limited penetration," Srivastava says. "As an Nvidia Preferred Partner, Vultr appears to be positioning itself at the forefront as cost-effective alternative for AI workloads, potentially making AI/ML training more accessible."

For Athos Therapeutics, the Vultr-Dell solution is the optimal choice for developing its platform for building and training its homegrown models.

"We don't have even have any legacy data, code, or platform. We do use the public cloud for processing some data, but for training our models with genomics data, Vultr and Dell gave us a proposal that is very reasonable in this space," Guo says.

See also:

- CIOs recalibrate multicloud strategies as challenges remain
- Cost concerns put CIOs' AI strategies on edge
- Private cloud makes its comeback, thanks to AI

by **Paula Rooney**
Senior Writer
Paula Rooney is a senior writer at CIO.com, where she focuses on how CIOs deploy AI, cloud, and digital technologies to transform their organizations. A veteran IT journalist, Paula has reported for Linux.com, The Register, TechTarget, ZDnet, and UBM, among other outlets. She holds a master's degree in journalism from Columbia University and was most recently recognized with ASBPE Regional Silver and Regional Bronze awards for her enterprise news story "AI to go nuclear? Data center deals say it's inevitable" and her case study "LA Public Defender CIO digitizes to divert people to programs, not prison."