VULTR run:ai

SOLUTION BRIEF

Run:ai + Vultr for AI and ML Orchestration

Vultr's cloud infrastructure and Run:ai's advanced orchestration platform are optimized for better performance with NVIDIA GPU-based AI and ML workloads.

VULTR.COM

VULTR

Run:ai + Vultr: Advanced NVIDIA GPU Orchestration for AI/ML Workloads

Run:ai and Vultr provide enterprises with comprehensive AI/ML application tools and capabilities, including powerful NVIDIA GPUs and networking, advanced orchestration, scalability, global reach, optimized price-to-performance, resource management, and enhanced productivity.

As a part of the Vultr Cloud Alliance, Run:ai offers a cutting-edge AI workload orchestration platform crucial for developing advanced AI and ML applications. Running on Vultr's robust cloud infrastructure, including Vultr Cloud GPUs accelerated by NVIDIA and Vultr Kubernetes Engine, this premier orchestration technology supports businesses in implementing AI strategies by delivering an integrated platform that ensures high performance and cost efficiency worldwide.

Run:ai enables the dynamic allocation and management of GPU resources, ensuring efficient and optimized AI workloads. Through this collaboration, businesses can harness the full capabilities of GPUs for applications that require precise scheduling, resource sharing, and workload distribution at scale.

By leveraging Vultr's global infrastructure and Run:ai's orchestration capabilities, organizations can build, deploy, and manage their AI workloads effortlessly. This approach enhances resource efficiency and security, equipping businesses across the industries with the agility needed to navigate the evolving demands of modern AI environments.

AI orchestration on high-performance cloud infrastructure

For businesses and developers needing efficient and scalable AI infrastructure, Vultr and Run:ai offer a cutting-edge solution that enhances resource utilization, supports rapid AI deployment, and provides customizable scalability through their integrated infrastructure and advanced AI workload orchestration.

Kubernetes-native orchestration

Seamlessly orchestrate containerized AI workloads with Run:ai utilizing Vultr Kubernetes Engine, ensuring efficient and automated management of resources.

Global reach, low latency

Vultr offers a global cloud platform with NVIDIA-accelerated cloud GPUs, strategically positioned near end users' applications, connecting 90% of the world's population within 2-40ms. Combined with Run:ai's AI workload orchestration, it provides low-latency and high-performance cloud, enhancing AI-driven solutions.

Vultr Cloud GPU, accelerated by NVIDIA

Vultr offers access to the latest NVIDIA GPUs, including:

- NVIDIA GH200 Grace Hopper™ Superchip
- NVIDIA H100 Tensor Core GPU
- NVIDIA L40S GPU
- NVIDIA A40 GPU
- NVIDIA A16 GPU
- NVIDIA A100 Tensor Core GPU

Solve the challenges of AI deployment

Difficulty in scaling AI training jobs

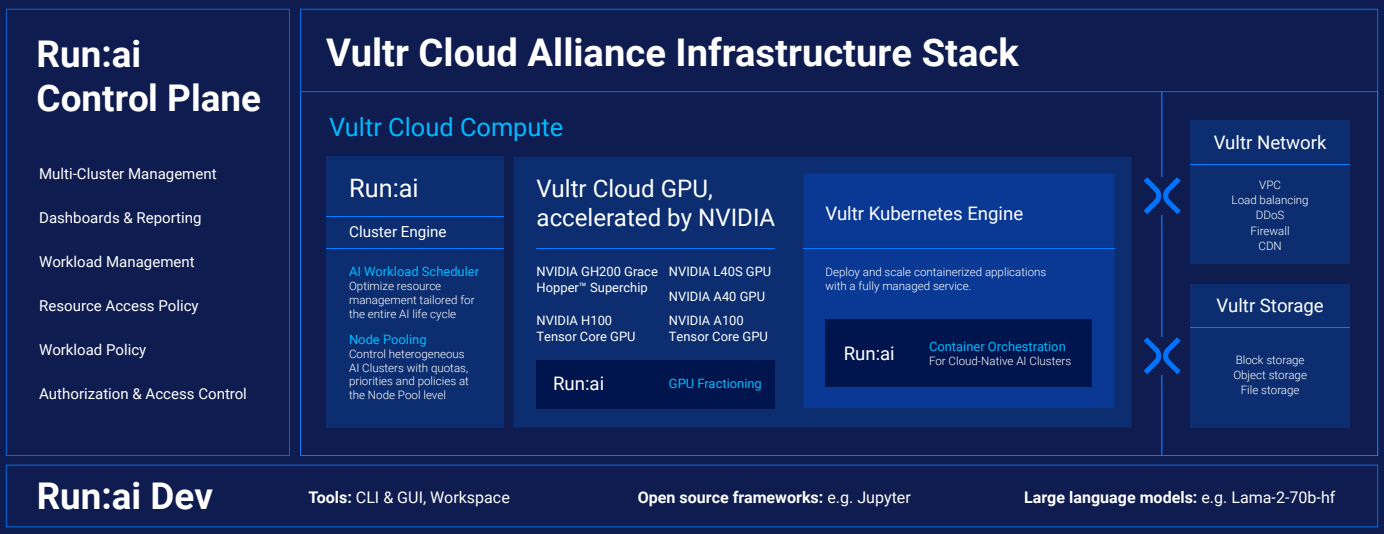
Simplify scaling AI training jobs across multiple GPUs and nodes with dynamic resource allocation and management, supported by Vultr's easily scalable instances connected with NVIDIA Quantum-2 InfiniBand networking for smooth and efficient scaling.

Fragmented AI development tools

Access pre-configured workspaces integrating tools like Jupyter Notebooks, PyCharm, and VSCode into a unified environment, enhanced by Vultr's robust cloud infrastructure for smooth operation and quick access.

Long queue times for AI experimentation

Run:ai's job queuing system efficiently queues and executes hundreds of prioritized batch jobs, while Vultr's scalable infrastructure processes these jobs quickly, reducing wait times and accelerating research.



Empowering industries globally

Manufacturing: Run:ai and Vultr support the manufacturing industry by optimizing GPU allocation and providing scalable cloud solutions. This enables efficient AI-driven smart factory operations, predictive maintenance, and industrial automation. Their integration enhances productivity, operational efficiency, and competitive advantage, aiding Industry 4.0 adoption.

Healthcare & life sciences: Unlock healthcare advancements with Run:ai AI orchestration on Vultr Kubernetes Engine to accelerate AI-driven personalized medicine. Enhance patient monitoring, genomics, drug discovery, and diagnostics securely, improving patient outcomes and driving innovation in life sciences.

Telecommunications: Run:ai optimizes GPU resource allocation, enabling efficient AI model training and deployment for network optimization. Combined with Vultr's high-performance cloud infrastructure, accelerated by NVIDIA GPUs and networking, this partnership harnesses the full potential of AI for network performance and service delivery.

Financial services: Run:ai enables efficient AI model training for fraud detection, predictive analytics, trading algorithms, and customer behavior modeling. Vultr offers sovereign clouds on-demand, ensuring compliance with a wide range of local requirements and regulations for secure and compliant operations, driving innovation in banking, fintech, and insurance.

Retail: Run:ai and Vultr enable real-time analytics, automated logistics, personalized customer experiences, and optimized stock levels with efficient GPU allocation and high-performance cloud infrastructure. Retailers can leverage AI for predictive analytics, dynamic pricing, demand forecasting, and inventory management, boosting efficiency, customer engagement, and innovation.

Media & entertainment: Run.ai and Vultr offer AI-driven solutions for content creation, video production, and data-driven advertising. This enables faster innovation, high-quality visuals, and personalized content at scale. It supports real-time rendering for gaming and optimizes AI for dynamic video editing, music production, and interactive media.

Key advantages for AI and machine learning workloads

Improved price-to-performance

Dramatically cut cloud costs — up to 90% compared to traditional hyperscalers — while delivering powerful AI capabilities. This price-to-performance advantage is essential for companies looking to maximize their technology investments without compromising quality.

Accelerated time-to-market

Speed up the deployment and management of AI initiatives by leveraging the integrated tools and optimized infrastructure from Run:ai and Vultr, simplifying AI development processes and allowing data scientists and engineers to focus on building and fine-tuning models without dealing with infrastructure complexities.

LLM management

Deploy and manage private large language models (LLMs) seamlessly, ensuring secure and efficient handling of AI models with Vultr's extensive security features like IAM and VPCs.

Learn more about Run:ai and Vultr

Contact us at vultr.com to get started. →