



DATASHEET

Hybrid Cloud Architecture for IoT, Edge Computing, and AI Workloads

Vultr's hybrid cloud architecture enables secure IoT and AI deployments across industries. With low-latency infrastructure and global reach, enterprises can manage real-time data processing and integration, ensuring seamless and efficient IoT operations at scale.

[VULTR.COM](https://vultr.com)

IoT and AI workloads on composable cloud infrastructure

The IoT market is projected to reach \$407.6 billion across key industries through 2025 as businesses pursue real-time insights and operational efficiency. By 2027, 75% of new, industry-specific applications are expected to include embedded IoT middleware, a significant increase from 10% today.¹ Scaling and securing the future of IoT solutions requires hybrid cloud, edge computing, and AI, especially for enterprises managing complex, distributed workloads globally.

Vultr's composable cloud, with bare metal and powerful virtualized GPUs, offers the flexibility and performance essential for IoT and AI innovation. With 32 cloud data center regions worldwide, Vultr enables enterprises to deploy and manage latency-sensitive solutions at the edge, ensuring global reach and fast local performance. This infrastructure facilitates seamless integration between cloud and IoT devices, making efficient data processing and analysis crucial for real-time decisions and enhancing operational efficiency.

Challenges and solutions

Vultr and the [Vultr Cloud Alliance](#) work together to tackle IoT challenges in data, latency, and security head-on.

Data management and interoperability

IoT devices generate vast data volumes that need efficient storage, processing, and analysis for actionable insights. Vultr's composable cloud infrastructure with global storage and cloud GPUs enables scalable data handling and seamless IoT-AI integration, enhancing data management and enabling swift, data-driven decisions.

This can be extended with NetApp for hybrid cloud storage across edge and cloud, and with DDN for high-performance storage optimized for GPU workloads.

AI and data processing

Real-time AI data processing is crucial for time-sensitive IoT applications. Vultr's edge computing and

high-performance cloud GPUs deliver the speed and power needed for rapid IoT data handling.

This can be extended with SQream's GPU-accelerated analytics and Run:ai's workload management for an optimized AI setup at the edge.

Latency and bandwidth

Centralizing IoT data processing can often increase latency and bandwidth costs, impacting performance. With Vultr's global reach and powerful GPUs, processing happens closer to IoT devices, reducing latency and bandwidth needs. This setup enhances customer experience by delivering timely insights and real-time connectivity.

Vultr's reduced latency can be enhanced with Console Connect's private networking and EdgeSIM, which bypasses the public internet and improves data flow.

Security and networking

The vast number of IoT devices introduces significant security risks, particularly in data transit, and creates complex networking demands as OT and LoB devices converge with IT systems. This complexity strains current infrastructure, impacting latency, coverage, and security. Vultr's secure cloud infrastructure and private networking help safeguard IoT data in storage and transit.

For additional security, Console Connect's private networking and Qdrant's vector database for RAG offer secure, efficient storage and querying.

Scalability and performance

IoT and AI workloads must be rapidly scalable to accommodate data growth and fluctuating workloads. Vultr's flexible cloud solutions, including powerful GPUs and bare metal servers, allow organizations to scale as data needs expand. This agility enables enterprises to deploy IoT solutions across locations cost-effectively, with infrastructure that scales to meet growth and customer needs.

¹Gartner: [Cross-Industry Insight: IoT Market Opportunities and Top Spend Use Cases](#)

Hybrid cloud for IoT across industries

Financial services

IoT devices, including biometric sensors and connected payment systems, enable real-time fraud detection and secure transactions. Vultr's low-latency infrastructure provides the essential speed and security to support fraud prevention and risk assessment in financial networks.

Healthcare and life sciences

With IoT wearables and medical sensors enabling remote monitoring and diagnostics, Vultr's high-performance cloud and AI-driven insights allow for efficiently analyzing large datasets, supporting improved patient outcomes through timely, edge-based care.

Manufacturing and energy

IoT sensors track equipment to predict maintenance needs, reducing downtime and maximizing productivity. Vultr's scalable cloud solutions allow manufacturers to process IoT data near the source, optimizing workflows and maintaining operational efficiency.

Media, entertainment, and gaming

Immersive experiences via VR/AR require robust IoT support. Vultr capabilities enhance these applications, while its global network ensures that latency-sensitive, high-quality content reaches audiences seamlessly worldwide.

Retail

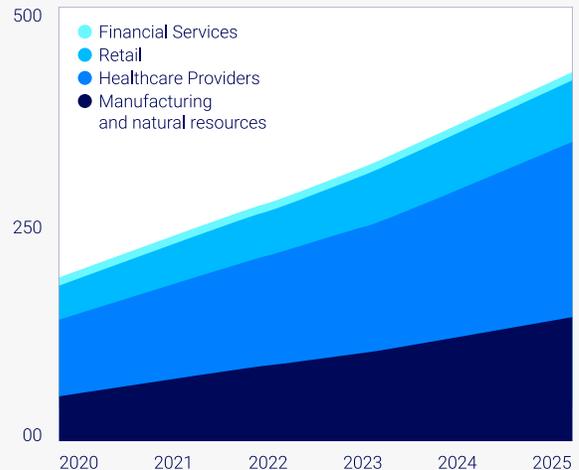
IoT devices monitor inventory and customer behavior, streamlining operations and personalizing the customer experience. Vultr's composable cloud infrastructure supports real-time insights and processing, driving more innovative inventory management and tailored engagement strategies.

Telecommunications

IoT devices track network performance, detect outages, and improve service. Vultr's hybrid cloud and IoT integration help telecoms maintain uptime and enhance user experiences through edge processing.

IoT Spend by Key Verticals, 2020-2025

IoT Spend (Billions of dollars)



Source: Gartner

IoT connectivity benchmark

Flexible IoT connectivity for mission-critical data

As IoT devices manage increasingly critical data, secure, low-latency network performance is crucial. Traditional solutions often fall short, exposing data to public internet risks or limiting flexibility with rigid VPNs. [Vultr Cloud Alliance](#) partner [Console Connect](#) offers a robust suite of connectivity solutions, including Edge SIM and Cloud Router Layer 2 and 3, providing flexible, secure, low-latency connections across IoT networks. This enables your IoT infrastructure to scale globally with seamless access to Vultr's high-performance data centers, ensuring reliable, real-time data transfer for mission-critical applications.

Latency for IoT-cloud integration

Benchmark objectives and components

This benchmark measures network latency between IoT devices and cloud providers, Vultr and AWS, in Paris, Singapore, and Toronto to evaluate real-time communication for latency-sensitive AI and IoT applications. We assess each provider's response speed using round-trip time (RTT) from an Edge SIM-enabled device. Network stability is ensured through multiple trials, providing reliable metrics across different times and conditions. Console Connect's direct paths simulate real-world scenarios, delivering latency results relevant to practical use cases.

Benchmark test

This benchmark test measures network latency between IoT devices, Vultr and AWS cloud instances. Identical instances are deployed, and an IoT device connects via Console Connect for direct, secure paths. Round-trip times (RTTs) are recorded immediately upon activation, with multiple readings to ensure consistency. The data is then analyzed to identify differences in response times across regions.

Results

Tested with Console Connect, this benchmark shows Vultr's superior low-latency network for IoT, which is ideal for efficient, scalable cloud integrations.

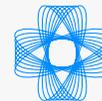
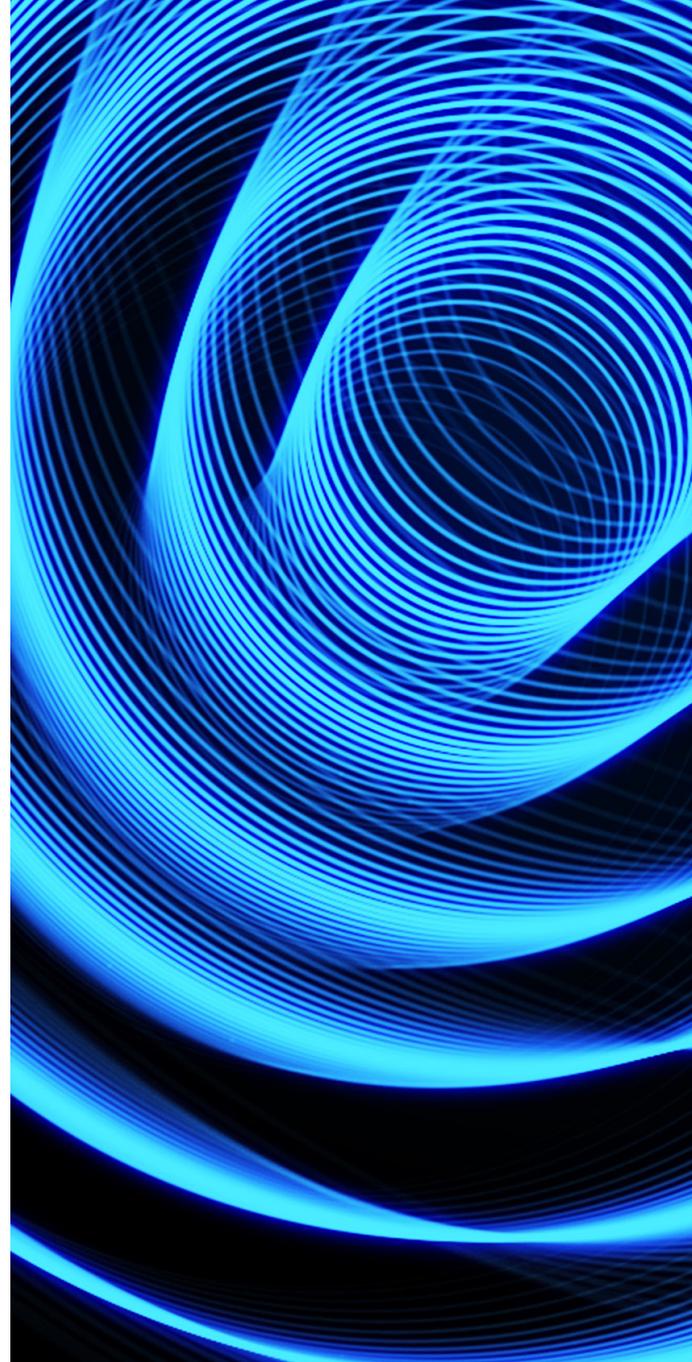
Latency Paris	AWS	Vultr
Average	83 ms	74.5 ms
Range	75-91 ms	71-78 ms

The importance of this benchmark

This benchmark is crucial for real-time IoT and AI applications. It enables low latency for instant decisions, enhances user experience by reducing delays, and supports scalable deployments across regions.

Impact on the business

The benchmark results highlight substantial advantages for businesses with hybrid cloud and IoT setups. Vultr's lower latency (74.5 ms vs. AWS's 83 ms) enhances operational efficiency, enabling faster data processing for latency-sensitive tasks. This low latency supports smooth, responsive interactions between cloud and IoT, allowing seamless scalability across regions without impacting user experience. Reduced latency also improves cost efficiency by streamlining operations, supporting real-time automation, and maximizing ROI in real-time applications.



Use case

IoT, GPU cloud, and private network

A live demo of Vultr Cloud GPU, Console Connect, and Gravio showcases secure, AI-powered IoT connectivity, integrating sensor data and Visual Questions Answering (VQA) capabilities.

[Read the use case](#) →

Data processing across hybrid environments at the edge

Unified data processing with IoT and cloud

IoT devices generate immense volumes of data at the edge, offering valuable insights but posing challenges in harnessing their full potential. SQream's GPU-powered analytics, part of the [Vultr Cloud Alliance](#), work with Vultr's low-latency, high-performance infrastructure to combine data collected from multiple sources, including locally stored IoT data. This setup reduces delays, lowers bandwidth costs, and enables rapid, informed decision-making in latency-sensitive environments by consolidating this data for near-real-time processing closer to its source.

Edge data handling for IoT applications

Traditional centralized data processing often struggles with the size and complexity of modern IoT datasets, leading to delays in insights and higher infrastructure costs. Vultr's global network ensures low-latency connections, while SQream's GPU-powered analytics process terabytes to petabytes of data directly at the edge—on devices, sensors, or IoT systems. By reducing costly data movement and eliminating bottlenecks in preparation and querying, this partnership accelerates time-to-insight for IoT workloads, optimizes resource utilization, and delivers seamless scalability for distributed environments.

GPU-driven IoT data analysis

IoT data processing challenges

Legacy data processing workflows, burdened by complex data transfers and disconnected systems, struggle to manage the vast datasets generated by IoT applications in near-real-time. By utilizing GPU power through SQream and Vultr, businesses can efficiently process data near its source, minimizing delays and reducing infrastructure costs.

Data processing scenario

In industrial settings, IoT devices generate large volumes of data that demand immediate processing for critical applications such as anomaly detection and predictive maintenance. A typical example involves monitoring production lines to detect subtle changes in equipment performance. Delays in processing this data can lead to missed opportunities for optimization, unplanned downtime, and reduced product quality.

With SQream, businesses can process over 100TB of data daily with near-real-time insights on Vultr's infrastructure. This setup enables precise operational adjustments, enhances system reliability, and improves efficiency by ensuring timely responses to data-driven insights.

Performance insights

Data analytics with faster, smarter processing

SQream delivers scalable solutions, cutting data processing time by 92% on Vultr Cloud GPUs — turning hours into minutes and future-proofing your business.

[Read the benchmark datasheet](#) →

Agentic AI in hybrid cloud and IoT

The next technological shift is agentic AI, where autonomous agents make real-time, independent decisions, transforming how workloads are managed and executed. Agentic AI relies on hybrid cloud infrastructure with data center edge computing, private networking, and GPU acceleration to fully realize its potential. By processing data closer to its source, this setup reduces latency, ensures secure data transmission, and provides the computational power needed for autonomous agents to act swiftly and efficiently.

By integrating agentic AI into this architecture, businesses can unlock new levels of operational intelligence, allowing for seamless automation, improved decision-making, and the ability to scale IoT ecosystems with greater precision and reliability.

Learn more about Hybrid cloud architecture for IoT and AI

[Contact us at vultr.com](#) to get started. →