Solution Brief

Accelerating GPU Cloud Operations

Multi-Tenant AI Orchestration with Mirantis k0rdent AI and VAST AI OS

The Challenge: Data Platform Orchestration and Management

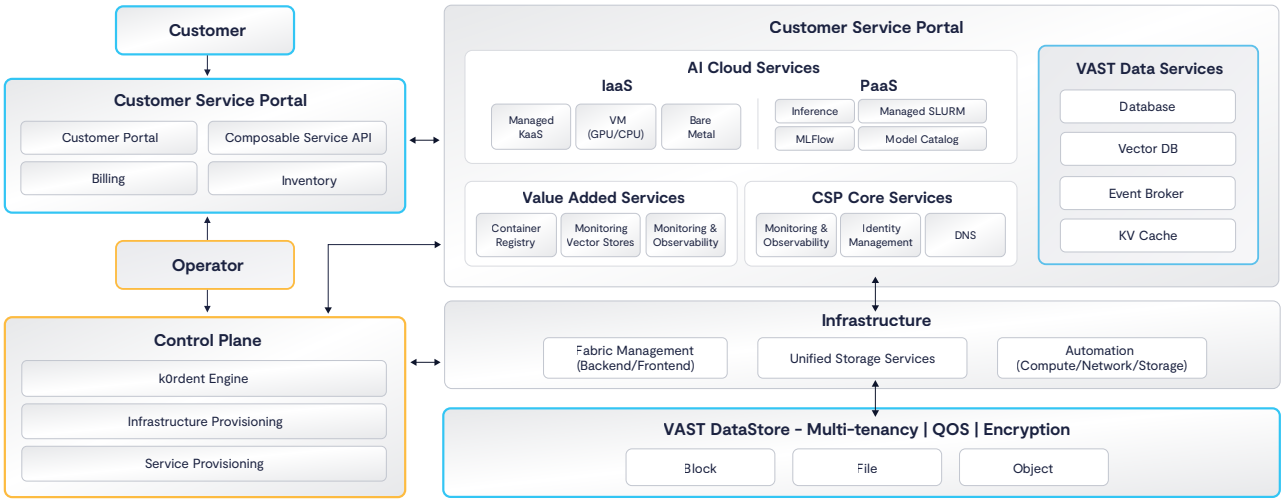
Neoclouds, Cloud Service Providers (CSPs), and modern enterprises face increasing complexity in platform engineering. Platform architects need ways of designing and operationalizing developer platforms swiftly using preferred technologies while empowering safe, secure, and policy-compliant use across global organizations. Organizations also require agility: instantiating and operating platforms across diverse clouds, bare metal, and edge environments from a single point of control and visibility.

Meanwhile, modern MLOps, AI, and data-driven workloads introduce unprecedented demands for compute, networking, specialized hardware (GPUs), and high-performance storage. These pipelines must reliably feed GPUs at sustained throughput, enforce strict tenant isolation, deliver predictable performance, and support dynamic data access across massive workflows, all while remaining accessible to infrastructure and platform teams with varied levels of technical specialization.

Single-layer provisioning tools exist, but GPU cloud operators still face significant control-plane fragmentation. Managing Kubernetes clusters, storage classes, tenant quotas, and network fabrics through disconnected automation silos creates configuration drift, increases integration overhead, and slows time to market.

Solution: Orchestrated, Multi-Tenant AI Data Infrastructure

Mirantis k0rdent AI with the VAST AI Operating System (VAST AI OS), bridges the gap between infrastructure orchestration and enterprise data management. Mirantis acts as the intelligent orchestration plane, while VAST provides the high-performance, multi-tenant data foundation, delivering an end-to-end platform for AI clouds.



Mirantis k0rdent AI & VAST AI OS

Joint Architecture: Multi-Tenant Orchestration & Federated Data Services

- **Mirantis k0rdent AI Control Plane:** Provisions and operates bare metal, virtual machines, and Kubernetes clusters across entire GPU fleets from a single API, so every tenant environment is built in a reliable and secure fashion, and arrives AI-ready.
- **VAST DASE & Multi-Protocol Access:** The Disaggregated Shared Everything (DASE) architecture, separates compute from capacity so each can scale independently, and it gives every node access to the full global dataset. VAST provides a unified namespace with concurrent NFS file, S3 object, and block access.
- **Automated Hard Multi-Tenancy:** Every tenant cluster k0rdent AI creates gets its own isolated tenant on VAST, with credentials that provide access to that tenant's data and nothing else. k0rdent AI sets this up as part of cluster creation, and VAST enforces at-rest, in-flight encryption, and RBAC within each tenant, so environments are secure and isolated by default.
- **Predictable Performance & Composable QoS:** Mirantis manages compute resource allocation while VAST enforces granular Quality of Service (QoS), guaranteeing bandwidth, IOPS, and tenant isolation to eliminate "noisy neighbors" in shared environments.
- **Integrated High-Value AI Data Services for CSPs:** Beyond basic storage, VAST consolidates native Vector Search, Kafka-API compatible event streaming (VAST Event Broker), structured analytics (VAST DataBase), and KV-cache offloading directly into the data plane. This allows CSPs to offer tenants complete, turnkey data pipelines for RAG, real-time analytics, and agentic workflows, eliminating the need to stitch together fragmented third-party databases or message queues.
- **Unified Governance & Simplified Ops:** Platform teams obtain single-pane visibility across compute, Kubernetes, and storage. Operators see committed capacity against physical capacity on every VAST Cluster, and tenants see their own consumption.

Solution Highlights

Mirantis k0rdent AI delivers VAST storage as part of every cluster it provisions. Instead of relying on separate storage management workflows, platform operators and tenants can provision, expand, and remove storage through the same k0rdent AI API and portal they use for compute. k0rdent AI unifies GPU compute, network fabrics, and VAST storage arrays under a single control plane.

- **Storage Delivered with the Cluster:** When a tenant orders a Kubernetes cluster and declares how much storage it needs, k0rdent AI provisions that storage on VAST, installs and configures the VAST CSI driver in the cluster, and offers storage matched to the tenant's IOPS, latency, and replication profile. k0rdent AI keeps the driver current for the life of the cluster. Bare metal tenants receive the same storage as a share ready to mount.
- **Automated Tenant View & Namespace Creation:** When a cluster is created, k0rdent AI creates a dedicated VAST tenant and VAST View (isolated storage path) for it, sets its quota on the array, and generates credentials scoped to that tenant alone. Every call k0rdent AI makes to VAST is authenticated with mTLS, and VAST enforces end-to-end encryption (at-rest and in-flight) and strict RBAC policies per tenant.
- **High-Throughput Data Path (NFS-over-RDMA):** k0rdent AI delivers GPU worker nodes ready for NFS-over-RDMA and dynamic multipathing for AI training and inference. GPU worker nodes bypass server CPU bottlenecks, and read directly from VAST EBoxes or DBoxes across high-speed fabrics such as Spectrum-X, InfiniBand, or RoCE Ethernet.
- **Drift Protection:** If a node's configuration, a storage mount, or a quota is changed outside k0rdent AI, the platform detects the change and restores what the operator set, with no manual audit.

Summary

The Mirantis k0rdent AI and VAST Data solution provides platform engineers with a unified point of control for operating sovereign, multi-tenant AI infrastructure. By combining Mirantis single control plane and automated lifecycle management with VAST's multi-protocol data platform, providers streamline onboarding, enforce strict security, and maximize GPU ROI.

About VAST Data

VAST Data is the AI Operating System company – powering the next generation of intelligent systems with a unified software infrastructure stack that was purpose-built to unlock the full potential of AI. The VAST AI OS consolidates foundational data and compute services and agentic execution into one scalable platform, enabling organizations to deploy and facilitate communication between AI agents, reason over real-time data, and automate complex workflows at global scale. Built on VAST's breakthrough DASE architecture – the world's first true parallel distributed system architecture that eliminates tradeoffs between performance, scale, simplicity, and resilience – VAST has transformed its modern infrastructure into a global fabric for reasoning AI. Learn more at vastdata.com and follow VAST Data on LinkedIn, YouTube and X.

©2026 VAST Data, Inc. All rights reserved. All trademarks belong to their respective owners.

10/26 I 13462