Solution Brief

VAST AgentEngine

Production Infrastructure for Enterprise AI Agents



Enterprise AI agents have moved quickly from experimentation to production, but most organizations are discovering that building an agent is only the beginning. Long-running workflows require persistent state, secure access to enterprise data and tools, automatic recovery from failures, observability, governance, and the ability to scale reliably across distributed infrastructure. VAST AgentEngine extends the VAST AI Operating System to provide the production infrastructure for agentic AI, enabling organizations to deploy, operate, scale, recover, and govern AI agents with the same security, resilience, and operational simplicity as their enterprise data platform.

AgentEngine bridges the gap between AI prototypes and enterprise production by providing the infrastructure that allows agents to run reliably for minutes, hours, or days while securely interacting with enterprise data, tools, and other agents.

Designed to Operationalize Enterprise AI

Production Infrastructure for Long-Running AI Agent

Most AI frameworks focus on building agents but leave organizations responsible for assembling the infrastructure required to run them in production. AgentEngine provides the managed infrastructure required to deploy and operate AI agents reliably at enterprise scale. It automatically manages persistent sessions, execution state, service discovery, container lifecycle, scaling, and recovery from failures, allowing developers to focus on building agents rather than operating infrastructure. Unlike short-lived serverless functions, AgentEngine agents can run for hours or days, periodically checkpointing their execution state so they can automatically recover from interruptions and resume complex reasoning workflows without restarting from the beginning. Built on the VAST DataEngine execution environment, AgentEngine orchestrates these long-running workloads while providing the resilience and operational simplicity required for production AI.

Built on DataEngine, Informed by InsightEngine

AgentEngine extends the event-driven execution model of the VAST DataEngine with three new abstractions: Services, Agents, and Stacks. Services expose long-running REST APIs and MCP interfaces, Agents perform reasoning and orchestration across multiple steps, and Stacks define how agents, tools, and DataEngine pipelines interact

Why VAST

Production AI Infrastructure

Deploy and operate persistent AI agents with built-in checkpointing, recovery, scaling, service discovery, and lifecycle management.

Built on DataEngine

Extend serverless functions and event-driven pipelines with managed Services, Agents, and Stacks.

Real-Time Context

Combine InsightEngine retrieval with unified access to files, objects, tables, vectors, and event streams.

within a governed deployment. Together with the VAST InsightEngine, which continuously transforms enterprise data into AI-ready context, AgentEngine enables AI agents to retrieve current information, reason over enterprise knowledge, and execute actions within a single unified platform.

Unified Identity, Governance, and Observability

AgentEngine provides the operational capabilities required for enterprise AI without requiring developers to build them into every application. Agents automatically inherit the identity and permissions of the users they represent, ensuring secure access to enterprise data, MCP tools, and other services. Built-in tracing, logging, and execution lineage provide complete visibility into every workflow, allowing administrators to reconstruct every decision, tool invocation, and data access event for troubleshooting, compliance, and governance.

A Unified Platform for Data, Context, and Action

Enterprise AI often requires separate systems for storage, vector search, orchestration, container management, observability, and identity management. AgentEngine eliminates this fragmentation by operating natively within the VAST AI Operating System. Agents access files, objects, tables, vectors, event streams, and DataEngine pipelines through a common security and governance model while leveraging InsightEngine for retrieval and context. The result is a single platform where enterprise data becomes AI-ready context and agents take action without unnecessary data movement or infrastructure silos.

Scale Agentic AI Without Rebuilding Infrastructure

Because AgentEngine runs natively on the VAST AI Operating System, agentic applications scale alongside enterprise data using the same disaggregated, shared-everything architecture. Managed containers automatically recover from failures, while services scale within administrator-defined limits to support growing AI workloads. Developers remain free to build with popular frameworks such as Agno, CrewAI, and LangGraph, while AgentEngine provides the production infrastructure needed to deploy, operate, and govern those applications reliably at enterprise scale.



Enterprise Governance

Enforce consistent identity, authorization, observability, and auditability across agents, tools, and enterprise data.



Unified AI Platform

Eliminate separate infrastructure for orchestration, storage, vector databases, observability, and container management.



Framework Flexibility

Build with Agno,CrewAI,LangGraph, or custom frameworks while AgentEngine manages the production runtime.



Enterprise Scale

Scale agentic applications alongside enterprise data using the VAST AI Operating System.

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.

07/26 | 12923