VMware Tanzu enhances support for generative AI and agents with Tanzu AI Solutions

June 11, 2025

by Greg Macatee

The vendor brings an updated set of AI-specific features in Tanzu Platform with Tanzu AI Solutions. They aim to reduce operational friction between developers, platforms and data science teams as businesses look to enhance their AI capabilities in private cloud, within the same application platform.

Introduction

VMware Tanzu brings an updated set of AI-specific products and features to the table in Tanzu Platform with Tanzu AI Solutions. These are intended to reduce operational friction between otherwise disparate developer, platform and data science teams as businesses look to build and enhance their AI capabilities in the private cloud, all within the same application platform.

THE TAKE

As customers move from AI experimentation to implementation, they stand to benefit from closer integrations between technology components as they redefine "Ops" frameworks within their businesses. VMware Tanzu AI Solutions are designed to do just that, with specific enhancements like AI middleware that boosts performance, fosters security and reduces time to value when it comes to operationalizing AI models and applications. Tanzu AI Solutions are polyglot for those who are unfamiliar with preferred data science languages or use other languages, and should appeal to Java developers with the launch of Spring AI. Spring AI also includes dedicated feature sets to control, observe, and evaluate models and, ultimately, agents, as organizations' AI capabilities progress. Model context protocol (MCP) is also supported in Tanzu Platform.

Context

Broadcom Inc. does not disclose VMware revenue figures since completing its acquisition, but does report an infrastructure software line item in its financial reporting that includes VMware's contribution. According to S&P Capital IQ Pro, Broadcom's infrastructure software revenue totaled $21.5 billion for its fiscal year 2024 that ended in early November 2024, and quarterly revenue for the segment in the first quarter of 2025 was $6.7 billion on an annual increase of 47%, which is notable since it is the first time a like-for-like comparison can be made with a year-ago quarter that fully includes VMware revenue. Broadcom also noted during its most recent investor call that its VMware Private AI Foundation collaboration with NVIDIA Corp. has 39 enterprise customers.

Customer interest in generative and, more recently, agentic AI remains high. According to 451 Research's Voice of the Enterprise: AI & Machine Learning, Use Cases 2025 survey, over 98% of respondent organizations report interest in exploring the capabilities of agents for a variety of use cases that include data entry and analysis, task automation, process optimization, customer service, content creation and personal assistants. However, they face a number of challenges in terms of agent development, deployment and measurement of ROIs as they begin to build out their capabilities.

Tanzu AI solutions

VMware Tanzu Platform takes a more developer-centric, platform as a service (PaaS)-oriented approach, building on the Cloud Foundry-based PaaS, Tanzu Application Service. The company hopes to better enable developers to build and deliver applications within their private cloud environments at scale, which better aligns with Broadcom's overarching strategy for VMware since completing the acquisition of the latter. However, it should be noted that Tanzu Platform continues to be offered across major cloud service providers such as Amazon Web Services, Microsoft Corp.'s Azure and Google Cloud Platform. Tanzu Platform also supports brokering of AI models running on NVIDIA CPUs and GPUs.

The Tanzu Platform portfolio includes Tanzu AI Solutions and other offerings that build upon these capabilities, bringing developer, platform and data team workflows closer together via a deeper set of tools and capabilities that are intended to support generative and, increasingly, agentic, AI. These include the following:

VMware Tanzu Greenplum. Tanzu Greenplum is a massively parallel data warehouse and analytics platform built on PostgreSQL that can help with data preparation for model training and with speeding AI model and application deployment. It supports structured, semi-structured, unstructured, vector, geospatial and graph data and contains tools that can assist with data science workflows including data ingestion, training, visualization and embedding via pgvector.

VMware Tanzu GemFire. Tanzu GemFire is an in-memory data grid equipped to deal with high throughput and low-latency operations. More specifically, Tanzu GemFire looks well suited to address vector databases and related applications such as retrieval augmented generation (RAG), where fast performance is needed to deliver data to models. It also supports caching that can improve performance and save on costs in addition to offering real-time analytics capabilities into data.

VMware Tanzu Data Services. Tanzu Data Services provides developers and administrators with sets of capabilities they need to access and manage their enterprise data. For developers, this includes self-provisioning capabilities they need that are designed to facilitate a more cloud-like experience. For administrators, this includes feature sets to ease data life cycle management and security with configurations for high availability, backup and replication. Tanzu Data Services also includes validation and commercial support for popular open-source software including PostgreSQL, MySQL, RabbitMQ and Valkey.

Spring AI. Spring AI is an extension of the Tanzu Spring Framework. Spring AI provides abstractions for Java to make AI-specific workflows and application building more approachable for Java developers. Spring AI supports most major model providers and model types (e.g., chat, text-to-image, audio and text-to-speech). RAG capabilities are also enhanced by features such as document ingestion and vector database support. AI model evaluation and observability of tools and embedding models are also supported to scan for potential hallucinations and other potentially unwanted behavior by models. Spring AI also includes chat memory support that can benefit both generative workflows and agents.

Tanzu AI Solutions also includes support for MCP, an open-source interface for giving models the context they need to access resources and execute tasks. With MCP, Tanzu Platform users can more easily connect their models to other data sources. This can help improve context and, ultimately, accuracy of models and decision making by agents. MCP also offers software development kits (SDKs) for a variety of different languages, including Java, which can further help lessen friction for developers when creating MCP clients and servers. While the value of MCP is clear, questions around security and observability are inevitable given the linkages that can be made to valuable enterprise data assets by host applications and models. This is a problem that will likely be amplified when introducing autonomous, non-deterministic agents into the mix. That said, a potential silver lining is the growing level of standardization on (or at least support of) MCP by technology vendors that can help contribute to a shared knowledge base around how to best navigate these potential pitfalls.

Competition

Tanzu Platform may encounter competition from suppliers such as Red Hat and SUSE in private cloud and/or on-premises environments and Amazon Web Services, Microsoft Azure and Google Cloud Platform PaaS offerings in the public cloud. It should also be noted that several of these vendors offer products that are specifically tailored around model training, development and deployment of AI workloads (e.g., Red Hat AI and SUSE AI).

SWOT Analysis

STRENGTHS

VMware continues to optimize its story around AI with Tanzu while lessening the friction between otherwise disparate enterprise personas in developers, platform engineers and data scientists. Tanzu also offers a deep set of features intended for Java developers with Spring AI, who are unlikely to be as familiar with preferred data science languages.

WEAKNESSES

Many VMware customers and partners are still coming to terms with the overarching changes made to licensing, pricing and partnership programs since it was acquired by Broadcom. Both current and potential Tanzu customers should benefit from a clearly defined, longer-term product road map, especially when evaluating it as a potential platform for AI development.

OPPORTUNITIES

Organizations continue to show interest in adding AI capabilities to new and existing applications, but many struggle to do so in practice. Facilitating a lighter initial setup for AI initiatives should help boost adoption.

THREATS

Just about every VMware competitor has a dedicated play for picking off customers. Certain customers have stronger affinities for using public cloud services when training models and building and deploying their AI applications.

CONTACTS

Americas: +1 800 447 2273
Japan: +81 3 6262 1887
Asia-Pacific: +60 4 291 3600
Europe, Middle East, Africa: +44 (0) 134 432 8300

www.spglobal.com/marketintelligence
www.spglobal.com/en/enterprise/about/contact-us.html