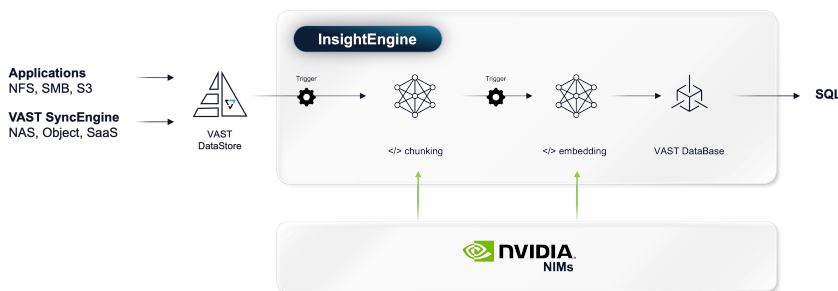


# Cisco Unified AI Data Platform



## Scale Agentic AI with Confidence

Cisco AI Data Platform with VAST AI OS makes enterprise data continuously usable for AI by unifying files, objects, tables, streams, metadata, vectors and events in one software platform. Instead of copying data across storage tiers, ETL jobs, standalone vector stores, and external indexing services, VAST keeps information live, governed, and ready for retrieval, inference, analytics, and agents. VAST InsightEngine extends that foundation by continuously classifying content, extracting context, generating embeddings, and turning enterprise knowledge into searchable, AI-ready signals.



Within Cisco AI PODs, VAST combines a transactional vector database, DataEngine's event-driven architecture, and managed Kubernetes services to ingest, enrich, embed, index, and serve data in place. This allows NVIDIA NIM microservices to operate as a service close to the data, reducing latency, data movement, and pipeline sprawl while keeping outputs current. New data can be processed as it arrives, so models and agents work from fresher context without waiting for batch refresh cycles or disconnected orchestration layers.

Aligned to Cisco Secure AI Factory with NVIDIA, the solution brings together Cisco CNode-X systems for GPU-accelerated data services and NVIDIA NIM orchestration with Cisco EBox systems for resilient all-flash data management, all on Cisco UCS and Cisco networking. Together, CNode-X and EBox form the Cisco-sold hardware foundation for Cisco AI PODs: EBox delivers the persistent, high-throughput data layer while CNode-X accelerates vector search, real-time pipelines, analytics, and AI services close to the data. The result is a modular, secure architecture that simplifies operations and speeds production AI.

### The VAST + Cisco Alliance

By partnering with VAST, Cisco was first to market with a unified AI Data Platform, launching VAST InsightEngine on Cisco AI PODs in September 2025. As a core component of Cisco Secure AI Factory with NVIDIA, the solution combines Cisco's AI infrastructure and networking leadership with VAST's advanced data platform to help enterprises unify, govern, and activate data for AI. The result is a faster, simpler path to secure, production-scale AI with performance, scalability, and enterprise readiness customers expect from Cisco.

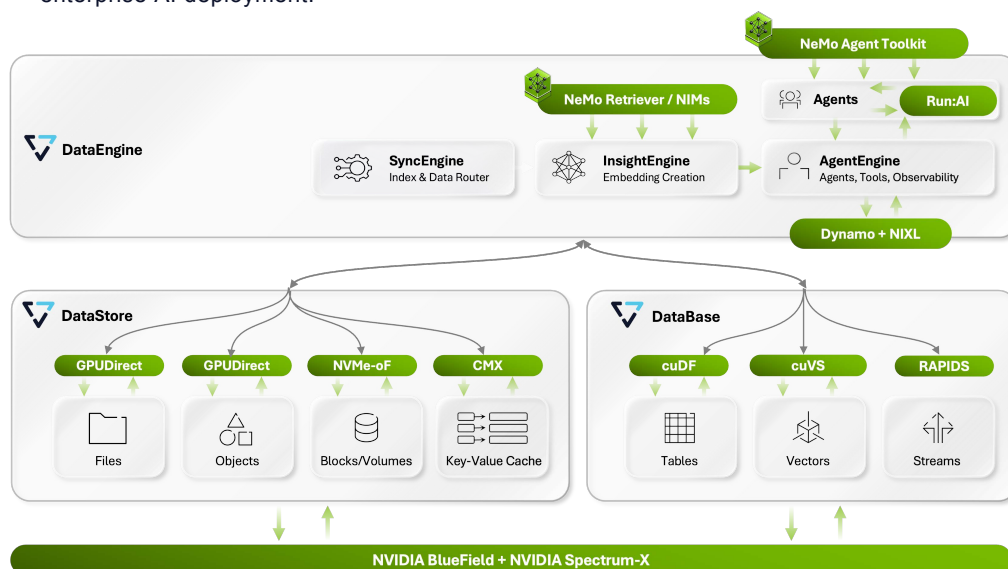
**Cisco CNode-X**  
UCS C845A M8 | NVIDIA RTX PRO 6000

**Cisco EBox**  
UCS C225 M8



## The VAST AI OS

VAST AI OS gives Cisco AI PODs a unified AI data platform that turns enterprise data into live, governed context for models, agents, and modern AI applications. Instead of stitching together storage, pipelines, vector databases, and separate processing tiers, VAST consolidates data services in one software platform that ingests, enriches, indexes, embeds, and serves data in place. Built on VAST's Disaggregated Shared-Everything architecture, the platform scales performance and capacity independently across flash and stateless compute to reduce bottlenecks, data copies, stranded capacity, and operational complexity. Within Cisco AI PODs and Cisco Secure AI Factory with NVIDIA, VAST provides the always-current data layer for retrieval, inference, analytics, and agentic workflows, helping enterprises keep data secure, relevant, and continuously usable at production scale. It simplifies operations while speeding time to insight and enterprise AI deployment.



## Example Use-Cases

### Enterprise RAG

Instantly query all your enterprise data using natural language across every file and document, with absolutely no model retraining required.

### Real-Time Analytics

Detect anomalies and respond to live streaming data events the moment they occur, enabling faster and more informed decisions across systems.

### Multimodal AI

Search and analyze text, images, and video content together seamlessly within one powerful unified AI platform built for enterprise scale now.

### Agentic AI

Autonomous AI agents that continuously retrieve, reason, and act on live, accurate, and fully contextual data to drive intelligent automation.

## Conclusion

Cisco and VAST deliver more than a reference architecture. Together they provide a production-ready AI data foundation that aligns modular Cisco AI POD infrastructure, Cisco Secure AI Factory with NVIDIA, and the VAST AI OS in one operational model. Cisco contributes trusted compute, networking, validated designs, and lifecycle support. VAST provides a unified software platform that keeps enterprise data continuously usable for retrieval, inference, analytics, and agents. The result is simpler operations, fewer silos, faster deployment, and a stronger path from experimentation to secure enterprise AI at scale today.

**Turn enterprise data into faster, trusted AI outcomes with Cisco and VAST.**

## Key Benefits



### AI Meets Data In Place

NVIDIA GPUs process data exactly where it lives within the system, fully eliminating costly and time-consuming data movement penalties



### Security Follows the Data

Identity-based access controls extend seamlessly from source documents all the way through to vectors and final AI responses



### Validated by Cisco

A pre-validated Cisco reference design with NVIDIA alignment reduces deployment setup time by up to 50% compared to manual integration



### Unified Stack Support

Cisco CX provides full-stack accountability for the entire solution, resolving complex multi-vendor issues an average of 43% more quickly

## About VAST Data

VAST Data is the leading data platform software company bringing businesses into the AI era. By accelerating time-to-insight for workload-intensive applications, the VAST Data Platform delivers scalable performance, simplifies data management, and enhances productivity. Launched in 2016, VAST has become the fastest-growing data infrastructure startup in history.

