

The VAST Data Platform

The Data Platform For The AI Era

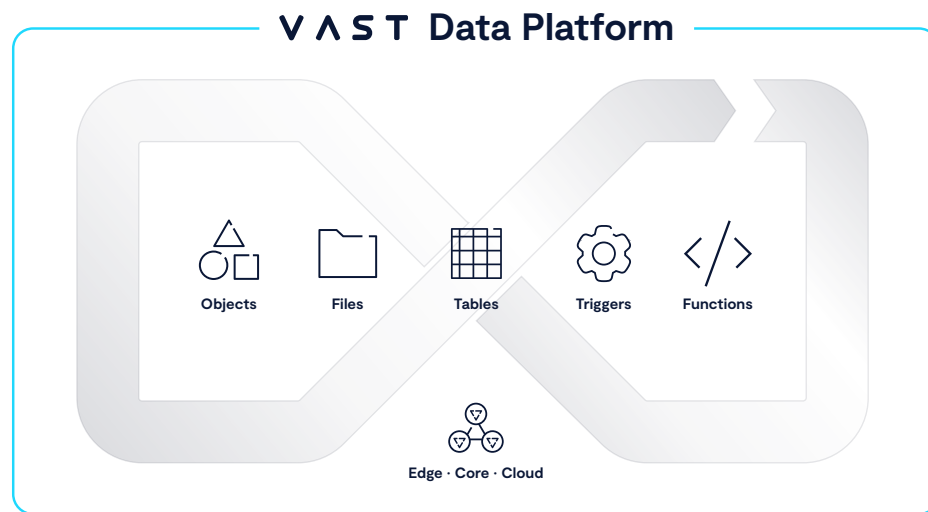


Unlocking AI-powered discovery for all your data.

Data continues to grow at unprecedented rates, largely fueled by unstructured content. To harness its full potential, we must reimagine computing. We wanted to give computers direct access to the natural world, efficiently store vast amounts of unstructured data, transform that raw data into an understanding of its underlying characteristics, and build a comprehensive knowledge base that could be queried. This isn't just about data – it's a leap towards new discoveries, pioneering the next wave of intelligent applications and robotics.

Introducing the VAST Data Platform

The VAST Data Platform is a breakthrough approach to data-intensive computing that serves as the comprehensive software infrastructure required to capture, catalog, refine, enrich, and preserve data through real-time deep data analysis and deep learning. Introducing the VAST Data Platform.



Automatically bring structured context to all unstructured data



Accelerate all transactional and analytical workloads



Eliminate data silos with a unified global namespace from edge-to-cloud



Enable AI-based workloads on all your data



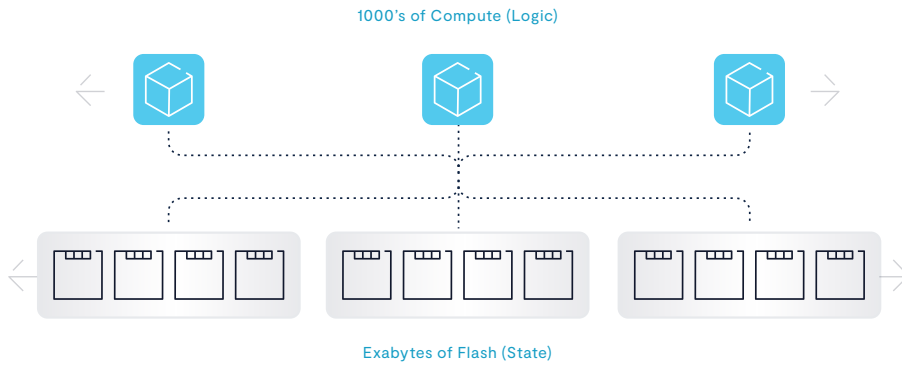
Simplify and automate AI and deep learning with data-driven framework

Intelligent storage for all of your unstructured and structured data. Automatic structuring of unstructured data upon ingestion using standard protocols, facilitating instant analysis of all data types.

High performance ingestion and analysis of structured data. Process millions of data rows per second, accelerating query speeds by over 20X, and enabling fast, data-driven decisions by combining streaming with analytics.

Access all your data from edge to cloud. Eradicates silos with one global namespace, achieving instant data availability and strict consistency across locations, while maintaining local performance levels.

Accelerate AI workloads on all of your data. Optimized for AI, the VAST Data Platform supports all data and major protocols, ensuring AI workloads access data natively and meet the demands of GPU-intensive tasks without additional systems.

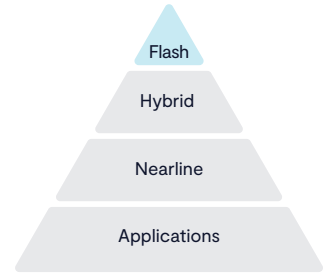


Disaggregated, Shared-Everything—To meet the performance, scale, and economics needed to enable today’s workloads on all data, we had to reimagine the definition of a scalable system. In 2019, we introduced the Disaggregated Shared Everything (DASE) architecture, which decouples system state and logic. By inventing a shared transactional data structure that ensures data consistency and integrity without the need for any east-west traffic, we enabled high-performance parallel data access from all compute nodes. By adding enterprise features, resilience, write and read optimizations, and data efficiencies, customers are now able to consolidate all their data into one performant cost-effective namespace.

Experience the world’s first deep learning data platform

The VAST Data Platform is not just a product; it’s a promise. A promise of a future where data is not just stored but understood, where insights are not just derived but acted upon, and where businesses are not just data-driven but vision-led. With over 10 exabytes of data under management, VAST enables companies like Disney, Zoom, NASA, G42, and others to embrace a future where the boundaries of what’s possible are constantly redefined.

Shared-Nothing
Multiple tiers required due to design limitations



Disaggregated, Shared-Everything
Consolidate all workloads with flash-based performance at archive economics



VAST Data Platform Key Features

Data Access

- **Unstructured**—Multi-Protocol: NFS, NFSoRDMA, SMB, NFS, S3, GPUDirect™, K8S CSI
- **Structured**—Support for Spark, Trino, VAST SQL, and more
- 1 million transactions per second with terabytes of query throughput at exabyte scale

Enterprise Features

- Snapshot and Object immutability
- Async replication with automated failover
- Quality of Service (QoS)
- Access Audit
- Encryption at rest and flight

Data Management

- Global Namespace across private and public clouds
- Local write performance with decentralized lock management
- Synchronization policies for optimized data placement
- Consistency modes for global read/write or point-in-time copies

Computing Services

- Fully-managed containerized computing environment that deploys on CPUs, GPUs, and DPUs - from edge to cloud
- Built-in triggers and functions to unlock instant value
- Bring your own code to discover the full power of your data
- Stream processing

Infrastructure

Scalable to Exabytes · Performance for AI workloads · Archive Economics · Full ACID transactions

Trusted By Industry Leaders

