

Success Story

Building the Future of Safer Driving with AI and VAST Data

VAST

A leading autonomous driving software company has taken a bold step forward in the race for safer mobility by reimagining its data infrastructure. Focused on delivering the most advanced driver assistance systems (ADAS) and autonomous vehicle capabilities, the company is training cutting-edge AI models that power perception, planning, and decision-making in real-time — helping vehicles better understand their environments and respond intelligently to complex driving scenarios.

The Challenge: From Still Images to Intelligent Motion

With a mission to lead in safety-first autonomy, the company has been steadily evolving its approach to deep learning. Traditionally, the team trained models on still images extracted from recorded video. But as part of their next-generation initiative, they're shifting toward training on motion video to enhance accuracy and contextual awareness in perception models. This advancement brings not only new demands for throughput, performance, and flexibility — but also new complexity in managing massive AI workloads and global-scale datasets.

Across a distributed infrastructure with hundreds of GPUs, the software team needed an intelligent data platform that could keep pace with large-scale training, manage petabyte-scale video repositories, and adapt to unpredictable research workflows.



Why VAST

This company turned to the VAST Data Platform, driven by three key differentiators:

Predictable Performance at Scale: Workloads currently maintain 50GB/s — a number easily sustained by VAST. More importantly, VAST enables them to independently scale performance without scaling capacity, allowing them to feed GPUs efficiently during training surges without over-provisioning infrastructure.

Extreme Efficiency with Data Reduction: Training datasets are built from high volumes of JPEG and PNG files — notoriously compressed formats. Despite this, the VAST Data Platform achieves a 2.5:1 data reduction ratio by leveraging advanced similarity-based data reduction algorithms, significantly outperforming legacy systems that were constrained to 1.1:1 ratios due to architectural limitations supporting only traditional compression methods. The results: more room for innovation, less waste, and radically better economics.

Disaggregated Licensing and Architecture: VAST's Gemini licensing model has proven transformative. With the ability to add compute and storage nodes independently of capacity licenses, the company can scale metadata performance and meet inode demands cost-effectively. This flexibility was crucial as the team scaled small-file workloads and overcame outages tied to metadata constraints.

About VAST Data

VAST Data is the data platform company built for the AI era. As the new standard for enterprise AI infrastructure, organizations trust the VAST Data Platform to serve their most data-intensive computing needs. VAST Data empowers enterprises to unlock the full potential of their data by providing AI infrastructure that is simple, scalable, and architected from the ground up to power deep learning and GPU-accelerated data centers and clouds. Launched in 2019, VAST Data is the fastest growing data infrastructure company in history. For more information, please visit and follow VAST Data on X (formerly Twitter) and LinkedIn.

Scan here to learn more.

Innovation in Action

The organization operates several VAST clusters, totaling multiple petabytes, across separate data centers. Together, these clusters support a hybrid research environment for hundreds of data scientists — many of whom compete for GPU time in a shared scheduler. VAST enables seamless collaboration, agile development, and uninterrupted experimentation.

And with new innovations like global snapshots, the team is taking a leap forward in how data is shared across locations. This functionality allows users to access complete datasets without replicating large volumes, enabling smart caching, scrambling for model variation, and cross-site collaboration without data bloat. For IT teams, it's a major operational win. For researchers, it's frictionless access to what they need, when they need it.

A Platform to Build the Future On

This customer isn't just storing data — they're building the future of mobility with it. With VAST, their AI infrastructure is no longer a bottleneck, but a strategic advantage. They can explore new training paradigms, manage massive data footprints, and continue evolving toward safer, smarter vehicles — all on a platform that grows with them.

©2025 VAST Data, Inc. All rights reserved. All trademarks belong to their respective owners.
04/2025