

Cutting Query Costs and Powering Threat Intelligence with VAST DataBase



A leading cybersecurity and threat intelligence company faced a critical inflection point with its data storage and analytics infrastructure. With over 2 billion malware and PCAP files requiring fast, concurrent access for analysis and detection, their legacy systems were beginning to face challenges. The IT team sought a solution that could efficiently handle massive data volumes, support high-performance analytics, and simplify operations.

Their infrastructure, which relied on legacy storage and Elasticsearch clusters, was becoming increasingly difficult to manage. As the data footprint grew to 2 petabytes, managing disparate data lakes and the complex Elasticsearch environment created mounting challenges. Several key pain points emerged:

- **High operational complexity:** Managing large, fragmented datasets across different storage platforms and dealing with frequent scaling issues in Elasticsearch created significant overhead.
- **Costly cloud storage:** Use of cloud providers for data queries led to high query costs, deterring analysts from fully utilizing the available data.
- **Inefficient data management:** The lack of a centralized platform for managing both structured and unstructured data made it difficult to perform holistic data searches and analyses.

Seeking a more modern, efficient solution, the organization turned to VAST Data to consolidate and enhance its threat research infrastructure. By implementing the VAST Data Platform, they streamlined data management and enabled more powerful analytics. One engineer on the team explained: “What enticed me about VAST is that it’s not just some behemoth that is picked up and dropped on top. It’s something that is new and well thought out, void of technical debt.”



What enticed me about VAST is that it's not just some behemoth that is picked up and dropped on top. It's something that is new and well thought out, void of technical debt.

– Engineer, Anonymous
Cybersecurity and Threat
Intelligence Company

Key Advantages



Consolidated Data Infrastructure: VAST enabled the organization to unify its data management, integrating both file and object with a multi-view system that supports NFS and S3 simultaneously. This consolidation removed the need for multiple storage solutions and reduced operational complexity.



Improved Data Reduction and Cost Efficiency: VAST's advanced data reduction capabilities delivered a 5:1 data reduction ratio, allowing the team to store and access large volumes of data more cost-effectively. They successfully replaced their existing legacy storage with VAST at a lower upfront cost than legacy HDD-based solutions.



Enhanced Performance for Threat Detection: VAST's ability to handle high concurrency and rapid read/write operations proved essential for data-intensive applications, including malware analysis and Splunk-based querying. VAST's architecture, which decouples compute from storage, ensures scalable performance as their data and analytics needs grow.



Operational Simplicity and Scalability: VAST's simplified data management significantly reduced the burden on the IT team. Adding capacity or compute resources is straightforward, allowing the organization to scale infrastructure easily as data grows. The streamlined maintenance process replaced the laborious updates and scaling hurdles previously experienced with Elasticsearch and the legacy storage vendor.



Cost-Effective On-Premises Data Management: For their specific use cases, keeping data on-prem proved more cost-effective than cloud storage. VAST's centralized Data Platform not only reduced operational costs but also eliminated the high costs associated with Amazon Athena queries. Running queries on large datasets, particularly mail metadata and threat intelligence, could previously reach mid-five figures for a single query.



Native Support for Structured Data with the VAST DataBase: The VAST DataBase offered a new way to store and query data at scale, replacing the need for Elasticsearch in some cases. With its SQL-based interface and seamless integration with Trino, the VAST DataBase provided faster, more efficient querying capabilities, reducing reliance on cloud-based data lakes.

Running queries is not cheap, and it scares people away from trying to query data. It's a bad combination—here's all the cool data, but you really can't touch it because it's really expensive to look at.



– Engineer, Anonymous Cybersecurity and Threat Intelligence Company

The Implementation Process

The team deployed VAST across its on-premises data centers, migrating 2 petabytes of critical threat research data. They then gradually transitioned more applications, including Splunk and QA tools, onto VAST. One engineer described the process: "We moved our existing data set one-to-one from our legacy provider onto VAST. VAST was less expensive, reduced our footprint, and simplified management—eliminating the operational headaches of our previous system."

The results were transformative. The organization saw significant improvements in data reduction, cost savings, and operational efficiency. The VAST Data Platform enables them to continue their cybersecurity mission while providing a foundation for future growth. As another team member noted: “I’m building out tools to make research better. Querying massive amounts of data at scale is the primary thing we were after, and VAST allows us to build solutions that make researchers’ work easier.”

VAST DataBase

The team then began exploring the VAST DataBase to catalog and query their structured and unstructured data more efficiently. “We were looking to displace some databases we had in Elastic, and the debut of the VAST DataBase hit just right. It seemed like a good fit to try out,” said one engineer.

With Trino integration, the VAST DataBase enabled them to run complex queries across their datasets, including malware samples and PCAP files. “Running queries is not cheap, and it scares people away from trying to query data. It’s a bad combination—here’s all the cool data, but you really can’t touch it because it’s really expensive to look at.”

The VAST DataBase’s free queries and lack of additional licensing costs simplified their cost model compared to previous Elastic and AWS-based solutions. “The VAST DataBase’s integration allowed us to consolidate multiple databases into a single, cost-effective platform, reducing cost and complexity.”

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.

The VAST DataBase’s integration allowed us to consolidate multiple databases into a single, cost-effective platform, reducing cost and complexity.

– Engineer, Anonymous Cybersecurity and Threat Intelligence Company



Another team member shared: “Migrating to the VAST DataBase was the clear choice for handling our structured queries that Elasticsearch previously managed. It freed up operational time and compute resources, without any added costs beyond storage. Now, we can focus on consolidating more data sources onto VAST. It is transforming how we manage data at scale. The data reduction capabilities make it an incredibly cost-efficient platform, allowing us to centralize and query vast amounts of data seamlessly across the company.”

By modernizing its threat research infrastructure with VAST Data, the organization has established a more agile, scalable, and cost-efficient data environment. This empowers analysts and researchers to work more efficiently and deliver better insights, helping to protect customers from emerging threats in real time. “VAST’s modern architecture replaced outdated legacy systems, empowering us to innovate our threat research infrastructure and unlock new levels of performance and scalability.”