VAST Data Platform Built for Deep Learning

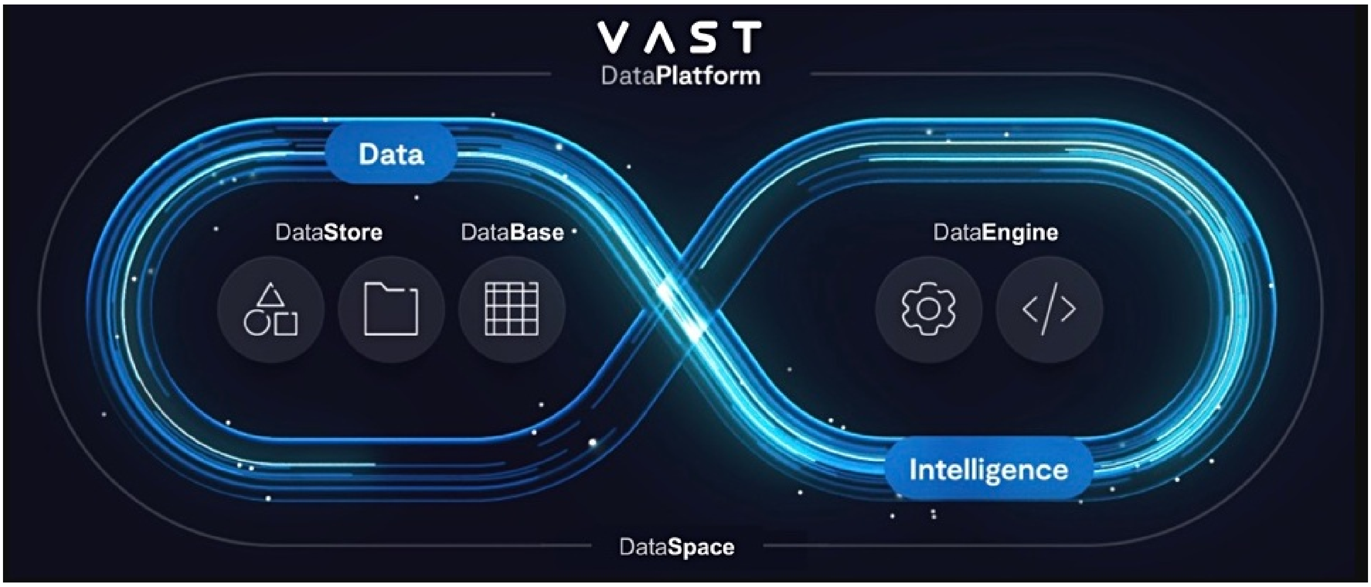
VAST Data Platform Built for Deep Learning User Guide

Home » VAST » VAST Data Platform Built for Deep Learning User Guide

Contents

VAST

VAST Data Platform Built for Deep Learning



Specifications

- **Data Encryption**: FIPS 140-3 validated ciphers
- **Key Management:** External key management
- **Access Control**: RBAC, ABAC, ACLs, SELinux labeling
- **Authentication:** Integration with Active Directory, LDAP, NIS
- **Data Protection:** Encryption at rest, certificate-based authentication
- **Audit:** Comprehensive logging of data access events

Introduction

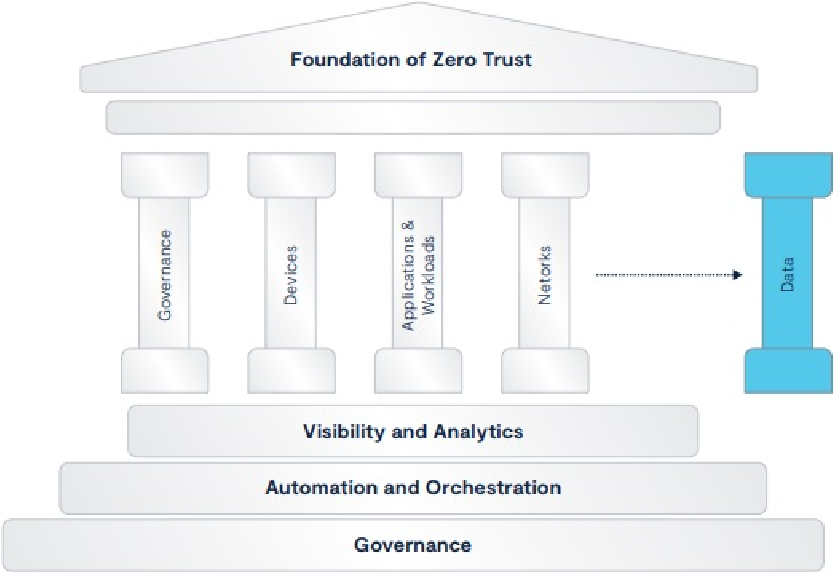
The VAST Data Platform provides a comprehensive suite of security capabilities to protect data confidentiality and enable secure multi-tenancy for unstructured data workloads. It integrates advanced access controls, encryption, auditing, and secure software development practices to meet stringent security and compliance requirements.

At its core, the platform leverages Multi-Category Security (MCS) from Security-Enhanced Linux (SELinux) to assign categories to files containing sensitive unstructured data like documents, images, and videos. Only authorized users and processes associated with those categories can access the data, preventing unauthorized access. This is complemented by secure tenancy features that create isolated logical or physical environments for different groups, with granular controls over resource allocation, networking, and access permissions.

The platform implements robust authentication and authorization mechanisms, including integration with Active Directory, LDAP, NIS, local user management, role-based access control (RBAC), and attribute-based access control (ABAC). It supports single sign-on (SSO), protocol access control lists (ACLs), and SELinux labeling for files and directories accessed via NFS, SMB, and S3 protocols.

Data protection is strengthened through encryption of data at rest using FIPS 140-3 validated ciphers, external key management, certificate-based authentication, and crypto erase capabilities. Comprehensive auditing logs all data access events, which can be stored in the platform’s database for analysis.

The platform’s secure software supply chain incorporates the NIST Secure Software Development Framework, software composition analysis, automated security testing, vulnerability scanning, and strict access controls throughout the development lifecycle. By combining advanced MCS, secure tenancy, encryption, access controls, auditing, and secure development practices, the VAST Data Platform delivers a robust security solution tailored for AI/ML and enterprise workloads on unstructured data.



Data Encryption and Key Management
The VAST Data Platform employs AES-XTS-256 encryption for data at rest and TLS 1.3 for data in transit. It supports external key management solutions like Thales CipherTrust and Fornetix VaultCore.

- NIST Control: SC-12 (Cryptographic Key Establishment and Management), SC-13 (Cryptographic Protection)
- Admin Guide Reference: Section: “Data Encryption” [p. 128]

This feature ensures that data is encrypted both at rest and in transit, protecting it from unauthorized access and ensuring compliance with cryptographic standards. The use of external key management further enhances security by centralizing and securing key management processes.

Access Control and Authorization
Feature: The platform integrates Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) to provide dynamic and granular access control.

- NIST Control: AC-2 (Account Management), AC-3 (Access Enforcement), AC-5 (Separation of Duties), AC-6 (Least Privilege)
- Admin Guide Reference: Section: “Attribute-Based Access Control (ABAC)” [p. 269]

RBAC and ABAC ensure that access to resources is granted based on user roles and attributes, enforcing the principle of least privilege and ensuring that users only have access to the resources necessary for their roles. This minimizes the risk of unauthorized access and potential data breaches.

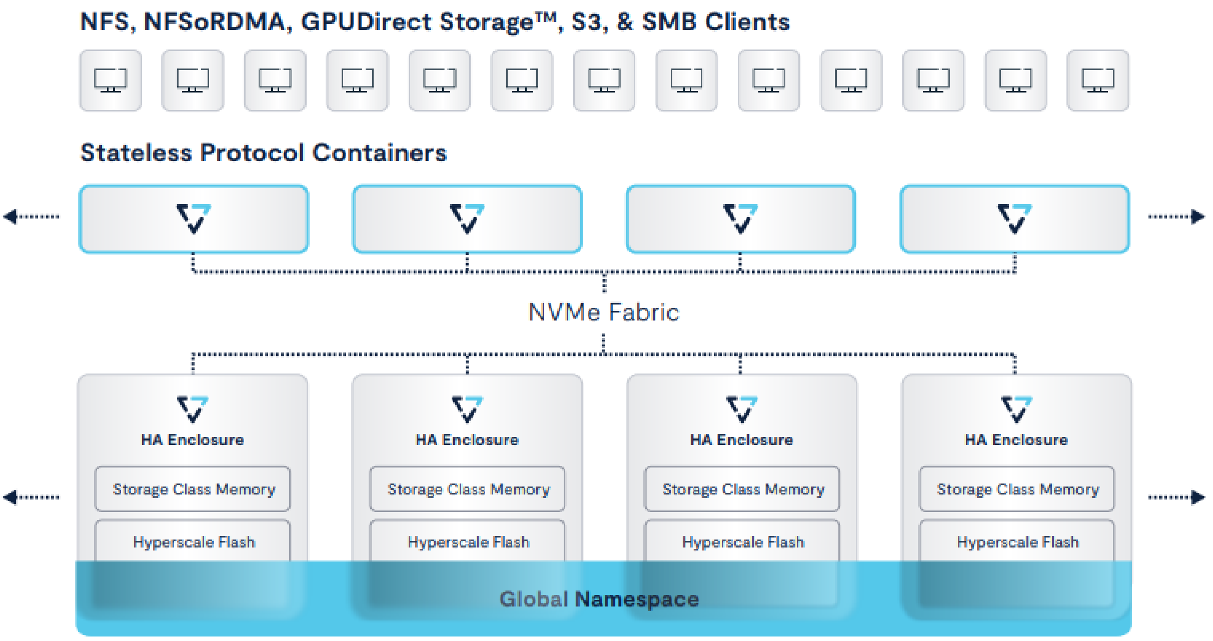
Audit and Accountability

- **Feature:** Comprehensive auditing capabilities including protocol and admin audit logs.
- **NIST Control:** AU-2 (Audit Events), AU-3 (Content of Audit Records), AU-6 (Audit Review, Analysis, and Reporting)
- **Admin Guide Reference:** Section: Protocol Auditing [p. 243]

The auditing features provide detailed logs of all access and administrative actions, ensuring that all activities can be tracked and reviewed. This is crucial for detecting and responding to unauthorized access attempts and ensuring compliance with regulatory requirements.

The VAST Cluster Architecture

Scale Capacity Independently From Performance


NFS, NFSoRDMA, GPUDirect Storage™, S3, & SMB Clients
Stateless Protocol Containers
NVMe Fabric
HA Enclosure
Storage Class Memory
Hyperscale Flash
Global Namespace

Data Flow and Segmentation

- **Feature:** VLAN tagging and binding, network segmentation, and control over protocol access.
- **NIST Control:** SC-7 (Boundary Protection), SC-8 (Transmission Confidentiality and Integrity)
- **Admin Guide Reference:** Section: "Tagging Virtual IP Pools with VLANs" [p. 147]

By segmenting the network and controlling data flow through VLAN tagging and binding, the platform ensures that data is isolated and protected from unauthorized access. This segmentation helps in maintaining the confidentiality and integrity of data as it moves across the network.

Data Sharing and Replication

Feature: Global Access allows you to make a subset of a cluster's namespace read and write accessible to clients of other clusters. This enables secure data sharing while maintaining control over access.

- **NIST Control:** AC-4 (Information Flow Enforcement), SC-7 (Boundary Protection)
- **Admin Guide Reference:** Section: "Global Access" [p. 413]

This feature provides granular access control down to the directory level, configurable lease expiration times for access, and auditing of access events, ensuring secure and controlled data sharing between clusters.

Asynchronous Replication

- **Feature:** Asynchronous replication allows replicating a subset of a cluster's data to a remote peer cluster for disaster recovery or data distribution purposes.

- **NIST Control:** CP-9 (Information System Backup), SC-8 (Transmission Confidentiality and Integrity)
- **Admin Guide Reference:** Section: “VAST Asynchronous Replication” [p. 381]

This feature ensures secure encrypted replication over WAN, granular replication at the directory level, read-only access at the replication target, and monitoring of replication status, providing robust data protection and disaster recovery capabilities.

Backup to S3

Feature: You can back up data from a VAST cluster to an S3-compliant object store, enabling sharing access to that data.

- **NIST Control:** CP-9 (Information System Backup), MP-5 (Media Transport Protection)
- **Admin Guide Reference:** Section: “Backup to S3” [p. 376]

This feature ensures secure transfer to external S3 targets, granular backup at the directory level, data immutability at the S3 target, and monitoring of backup status, providing secure and reliable data backup and sharing capabilities.

Global Snapshot Clones

- **Feature**: Create read/write clones of snapshots from a remote peer cluster, enabling shared access to point-in-time data copies.
- **NIST Control:** CP-9 (Information System Backup), SC-8 (Transmission Confidentiality and Integrity)
- **Admin Guide Reference:** Section: “Global and Local Snapshot Clones” [p. 425]

This feature provides secure encrypted transfer, granular cloning at the snapshot level, background syncing of changes, and auditing of access events, ensuring secure and controlled data sharing and recovery.

Zero Trust Architecture (ZTA) Implementation

- **Feature:** Automated data labeling, anomaly detection, and indestructible snapshots.
- **NIST Control:** CA-7 (Continuous Monitoring), SI-4 (Information System Monitoring)
- **Admin Guide Reference: Section**: “Zero Trust Data Pillar” [p. 269]

These features support continuous monitoring and anomaly detection, which are key components of a Zero Trust Architecture. Automated data labeling ensures that data is appropriately classified and protected, while indestructible snapshots provide a reliable means of data recovery and integrity verification.

Conclusion

The VAST Data Platform stands at the forefront of the industry by integrating advanced security features and compliance measures aligned with NIST Zero Trust Architecture (ZTA) principles. By implementing robust data encryption, access control, auditing, and data flow segmentation, the platform ensures comprehensive protection of unstructured data workloads. These features not only meet but exceed the stringent requirements set forth by NIST, positioning VAST Data as a leader in secure data management solutions.

The platform’s adherence to Zero Trust principles is evident through its meticulous implementation of continuous monitoring, automated data labeling, and anomaly detection. These capabilities ensure that data is consistently protected and that any potential threats are swiftly identified and mitigated. The use of Multi-Category Security (MCS) from Security-Enhanced Linux (SELinux) to assign categories to files containing sensitive data further exemplifies VAST Data's commitment to Zero Trust principles, ensuring that only authorized users and processes can access critical information.

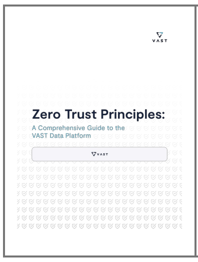
As the first in the industry to offer such a comprehensive suite of security features tailored for AI/ML and enterprise workloads on unstructured data, VAST Data is setting a new standard for data protection. By leveraging these advanced capabilities, organizations can confidently manage and secure their data, meeting stringent regulatory requirements and safeguarding against evolving cyber threats. The VAST Data Platform not only leads the industry in innovation but also provides a robust foundation for implementing a Zero Trust Architecture, ensuring that data remains secure in an increasingly complex digital landscape.

For more information on the VAST Data Platform and how it can help you solve your application problems, reach out to us at hello@vastdata.com.

FAQ

- **What is Zero Trust Architecture (ZTA)?**
  - Zero Trust Architecture is a security model based on the principle of maintaining strict access controls and not trusting any entity by default, whether inside or outside the network perimeter.
- **How does the VAST Data Platform enhance data security?**
  - The VAST Data Platform provides advanced security capabilities such as encryption at rest, robust access controls, comprehensive auditing, and integration with secure authentication mechanisms to protect data confidentiality and integrity.

Documents / Resources

 Zero Trust Principles: A Comprehensive Guide to the VAST Data Platform	VAST Data Platform Built for Deep Learning [pdf] User Guide Data Platform Built for Deep Learning, Data, Platform Built for Deep Learning, Built for Deep Learning, Deep Learning, Learning

References

- User Manual