FORTIFIED DATA LAKES IN THE CLOUD

Fortified Data Lakes in the Cloud

zero-trust by design beyond cloud-native capabilities

In an era of data-driven decision-making, public clouds have emerged as the go-to platforms for their high availability and scalability. Particularly, Cloud Data Lakes, collecting and storing data in its raw format (e.g., text, PDFs, videos, images, parquet), have found widespread applications in various sectors. Be it Life Science companies conducting clinical research, Financial institutions detecting fraud, Military Forces analyzing drone video streams, or Insurance agencies expediting claims resolution through customer-provided collision photographs, the scope is vast.

Leveraging immense amounts of data, that may contain sensitive information, to gain a competitive advantage with real-time AI/ML computing, requires strict adherence to data privacy and compliance standards. Neglect of proper access control, aligned with the principles of least privilege, significantly heightens the risk of a breach.

In this scenario, a "zero trust" approach to cloud security becomes essential, especially for a dynamic environment. But what does "zero trust" mean, and how can organizations implement it successfully?

Challenges

of Cloud Data Lakes
zero-trust security

Cloud Data Storage

Cloud platforms often use object storage (e.g., AWS S3) as the bedrock for data lakes. However, this approach has a flaw: when users gain access to a container, they are automatically granted access to all files. As a result, individuals and software exercise extraneous data access privileges while sysadmins retain visibility into all organizations' data. Cloud server-side encryption guards against external threats, but it is insufficient against malicious insiders or accidental errors.

Traditional Security Approaches

Existing third-party encryption solutions impede cloud speed and are not viable for real-time AI/ML computing.

Data tokenization is effective for certain data types but unsuitable for unstructured data like video or audio. Tokenizing free-form textual content is costly and time-consuming since sensitive areas must be pre-identified. Additionally, tokenization can impact the integrity of the AI model.

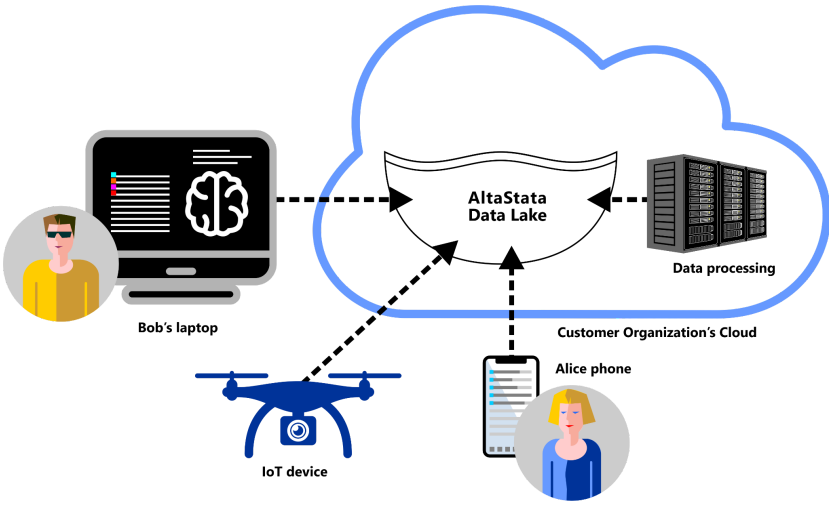
SIEM solutions only come into play after a breach has occurred, making them ineffective as proactive security measures.

Fortified Data Lakes in the Cloud

with end-to-end encryption

Founded by MIT Lincoln Labs researchers AltaStata, Inc. developed and patented a Fortified Data Lakes formation system with security at the core. This system provides end-to-end encryption throughout the collection, aggregation, computing, and exchange files and streams. With AltaStata, organizations can focus on their data-related tasks, assured of automatic data protection that doesn't affect cloud speed or scalability.

AltaStata Toolkit includes an Admin tool for Fortified Data Lakes formation in the cloud in minutes.

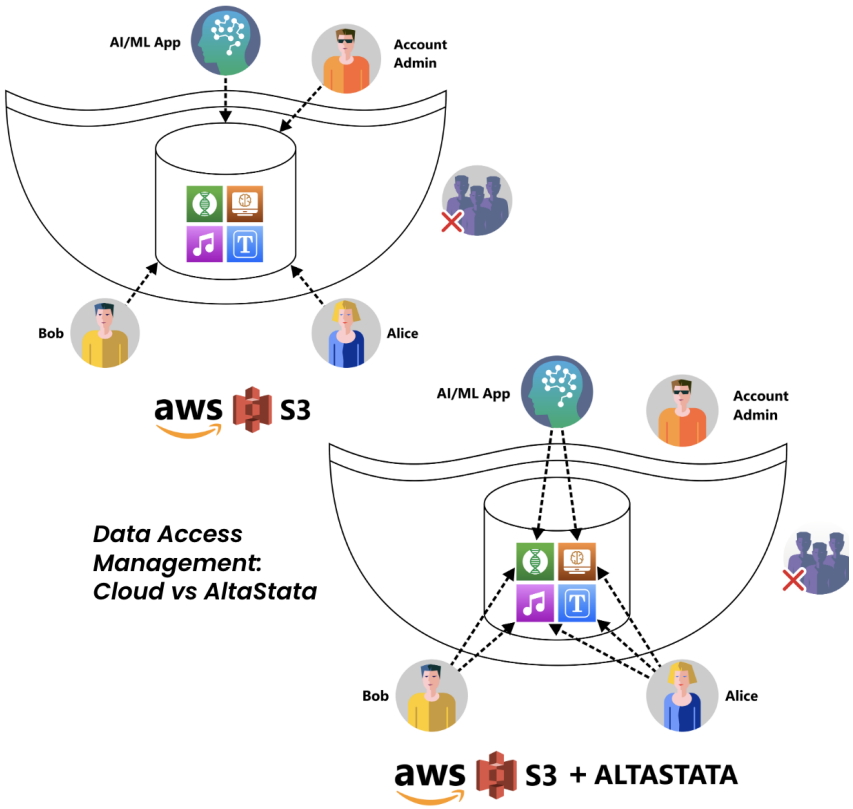


Via AltaStata's UI application and SDK, individuals, software, and autonomous devices can effortlessly and securely perform a multitude of actions. These include listing, uploading, downloading, and previewing encrypted files, streaming encrypted audio and video content, updating file versions, and granting or revoking file access as required.

AltaStata's Data Governance system automatically enforces various organizations' data access security policies and detects anomalies and potential threats.

AltaStata Method

Diverging from traditional cloud methods that use a single key to encrypt the entire storage, AltaStata generates a unique AES-256 data key for each file or stream. In addition to end-to-end encryption, this facilitates fine-grained control, enforcing user access at the file level as opposed to accessing the entire storage container.

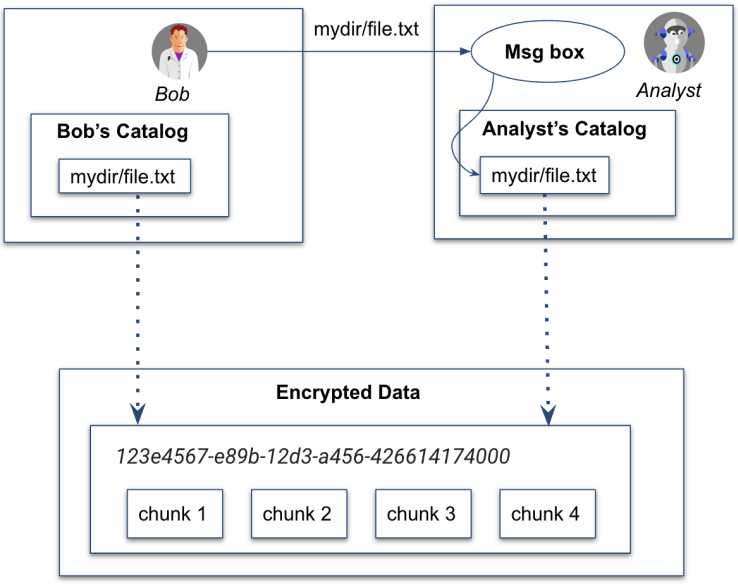


Data Access Management: Cloud vs AltaStata

Our unique method of structuring and storing encrypted data in the cloud enables sharing, parallel processing, streaming, search, and AI/ML computing at speeds comparable to unencrypted data processing. Furthermore, it demonstrates superior speed when handling certain file types, such as text and CSV.

AltaStata's security approach ensures high performance with three key concepts:

1. All files are split into fragments to enable streaming and parallel data access.
2. These fragments are compressed, reducing data transfer time.
3. These fragments are stored only once, without duplication. AltaStata facilitates data access sharing among users by transmitting copies of metadata objects stored in users' catalogs.



Business Impact

of Fortified Data Lakes

"AltaStata's solution ensures compliance with GDPR requirements for personal data protection, which is critical for the European market."

Robert Nica,
Chief Architect, BioBam

"AltaStata provides the first of its kind secured cloud strategy that enables the ability to bring integrity and high-availability of data and ensure that that data cannot be compromised."

Pierre Bourgeix
CEO, ESI Convergent

Seamless integration of AltaStata Fortified Data Lakes into existing multi-cloud environments and software ecosystems lets companies unlock the power of their data without compromising privacy or security.

Ensuring Zero-Trust Security Model and Compliance

Maintain granular access control over unstructured and semi-structured data, ensuring confidential data sharing among individuals, software, and organizations.

Multi-Cloud Flexibility

Enable multi-cloud flexibility with a single AltaStata Toolkit for all providers, thereby reducing cloud storage costs by up to 63% and optimizing performance.

AI/ML Computing

Perform real-time AI/ML computing while keeping the data end-to-end encrypted, preserving the integrity of your AI/ML models.

Get in touch today for your free, personalized demo.