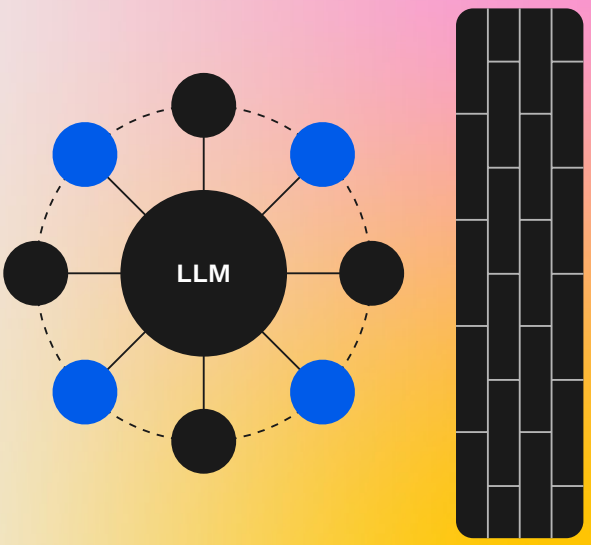
PANZURA



SOLUTION BRIEF

Panzura Symphony

Zero Trust Principled Data Broker

Enabling global trusted data access for AI in a Zero Trust environment.

A Zero Trust Data Broker is a secure intermediary platform or service designed to facilitate the transfer, sharing, or exchange of data between parties while adhering to zero trust security principles. The concept ensures that no entity, internal or external, is inherently trusted, and every action, user, or system must be continuously verified.

While Panzura Symphony was created as an unstructured data management platform intended for several applications outlined in the section below, this solution brief will focus on utilizing Symphony as a secure data intermediary, a Zero Trust Data Broker.

Data Assessment, Discovery, and Optimization
Panzura Symphony allows teams to conduct automated and comprehensive data discovery and assessments, providing a detailed overview of data including its structure, content, and relationships.

Compliance, Risk Management, and Security
Risk and Compliance Analysis automation allows organizations to more effectively manage regulatory requirements and mitigate risks, ensuring a more secure and compliant data environment.

Dynamic Data Movement Orchestration
This enables efficient data workflows, allowing for the seamless transfer and management of data across different environments. It helps in optimizing storage usage and reducing associated costs.

What is a Zero Trust Principled Data Broker?

Traditional data sharing relies on implicit trust in networks, users, or devices, which increases the risk of breaches or data misuse. A Zero Trust Principled Data Broker eliminates these vulnerabilities by enforcing strict verification and access controls, ensuring that shared data remains secure and protected at every step of the process.

Key Characteristics of a Zero Trust Data Broker

Zero Trust Security Framework

- **Least Privilege Access:** Users and systems are granted the minimum access necessary to perform their tasks.
- **Continuous Verification:** Every access request is authenticated and authorized in real-time, even if the user or system has been previously verified.

Data Protection and Privacy

- Sensitive information is masked or tokenized (ala webhooks) where necessary to prevent unauthorized exposure.
- Data-sharing policies and compliance regulations (e.g., GDPR, HIPAA) are enforced.

Data Governance

- Ensures that shared data adheres to pre-defined governance policies.
- Maintains control over data usage even after it leaves the organization, such as through rights management.

Real-Time Monitoring and Analytics

- Tracks all data access and sharing activities.
- Uses machine learning or rules-based systems to detect anomalies or suspicious behaviors.

Secure Intermediation

- Acts as a secure middleman that ensures data is only shared with authorized entities.

Applying Symphony as a Zero Trust Principled Data Broker in a Large Language Model Environment

Large Language Models (LLMs) have become integral to processing, analyzing, and generating insights from vast volumes of data. However, accessing unstructured data, which lacks inherent organization, poses significant challenges. This is where Panzura Symphony comes in as a Zero Trust Data Broker—an intermediary platform designed to facilitate secure and efficient access to data—Symphony addresses these challenges by enabling metadata-driven interactions between LLMs and unstructured data repositories. By utilizing Symphony, organizations can streamline data access, enhance security, and maintain compliance while optimizing LLM performance.

The Role of Metadata in Unstructured Data Access

Metadata plays a crucial role in bridging the gap between LLMs and unstructured data. Metadata provides descriptive information about unstructured content, including attributes such as keywords, creation dates, file types, data sources, and classification levels. By leveraging metadata, an LLM leverages the data broker for specific datasets without directly processing or accessing the raw data itself, which can be massive, and expensive to access in cloud environments.

This metadata-driven approach eliminates inefficiencies associated with processing irrelevant data and allows for focused interactions that align with specific business needs.

Secure and Controlled Data Access

Symphony supports robust security measures based on zero trust principles, ensuring that no requestor is inherently trusted and every interaction is evaluated, logged and verified. Through authentication and authorization processes, Symphony enforces the LLM's ability to access specific dataset locations. Data is filtered, prepared and provisioned based on predefined criteria, such as compliance regulations or business rules, ensuring that only relevant and permissible information is shared.

In scenarios involving sensitive information, Symphony can utilize JDBC, ReST and webhooks to finely tune provisioned datasets while enforcing limited access. This ensures that LLMs can perform their functions without exposing organizations to unnecessary risks.

In a zero-trust framework, webhooks serve as the backbone for real-time notifications and authorizations. When an LLM requests access to a dataset, Symphony provides advanced integration options; from JDBC to webhooks, as well as a comprehensive ReST API to enforce access restrictions, confirm compliance policy, or filter datasets using enhanced metadata. For example, if a dataset is reclassified as "high-priority," a webhook can immediately alert the LLM and re-provision the dataset ensuring the model works only with the latest artifacts. This dynamic interaction reduces delays and unnecessary polling while maintaining strict control over data delivery.

Symphony's range of integration options support streamlined interaction with external systems like identity and access management (IAM) tools. If an LLM requests data, a webhook may trigger an evaluation of permissions in real-time, ensuring that the request complies with established policies. Approved requests proceed securely, while unauthorized attempts are blocked and logged. Additionally, Symphony can automate actions to secure sensitive data, such as encrypting or masking it before sharing, ensuring that the LLM receives only permissible information.

Optimizing Data for LLM Ingestion

Symphony goes beyond simple data access by preparing unstructured content for LLM processing. Through Panzura Symphony's partnership with IBM Storage Fusion Data Catalog, as well as GRAU Data, it can normalize data formats, convert non-text formats (e.g., PDFs or images) to text, and extract 100's of proprietary formats used in Life Sciences, Genomic, Scientific, Geological, amongst others. Symphony's partners also extract metadata summaries that provide context for the data. Efficient use of metadata can help segment large datasets into manageable pieces, making it easier for the LLM to process the information efficiently.

This preprocessing minimizes the computational resources required by the LLM and allows it to focus on generating insights from relevant and structured content.

Metadata-Driven Query Optimization

Metadata significantly enhances the precision and efficiency of LLM queries. By using metadata tags, LLMs can refine their search parameters and retrieve only the data that meets specific criteria. This ensures that the LLM is working with the most relevant subsets of unstructured data, which reduces the time and resources required for processing.

Moreover, metadata enables the LLM to contextualize its results. For instance, when analyzing documents, metadata about the author, source, and relevance score can be included in the output, providing richer and more actionable insights.

Compliance and Auditability

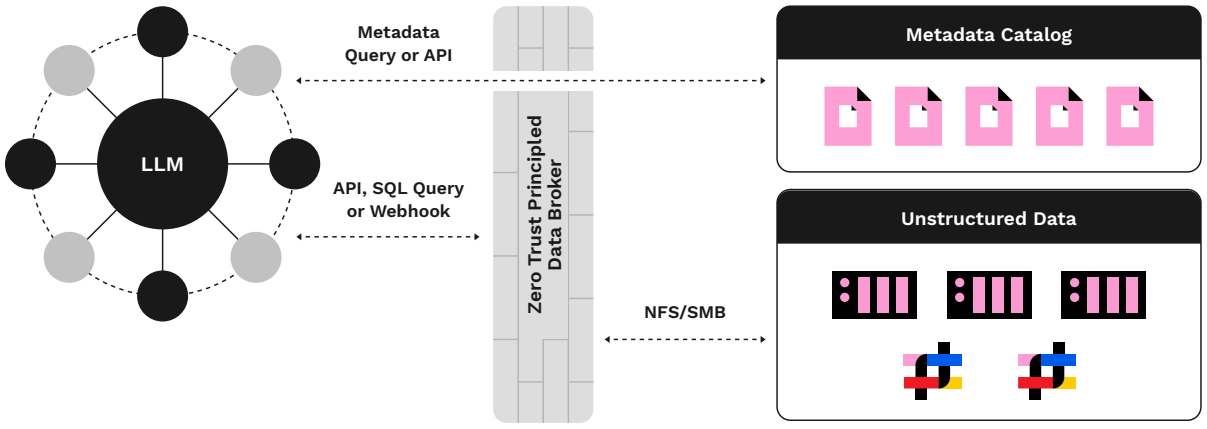
Maintaining compliance with data protection regulations, such as GDPR and HIPAA, is a critical concern for organizations handling unstructured data. Symphony simplifies compliance by providing access to relational database and enhanced metadata, supporting granular filtering. Every interaction is logged, creating a tamper-proof audit trail that shows what data was provisioned, to where, by what requestor and when. This transparency ensures accountability and reduces the risk of regulatory violations.

Advantages of Data Brokers such as Panzura Symphony for LLMs

By leveraging Symphony, organizations gain several advantages in using LLMs to process unstructured data. Symphony ensures efficient data access by delivering only relevant subsets of data, reducing computational overhead. It enforces strict security measures, mitigating risks associated with unauthorized access or breaches. Additionally, Symphony's ability to maintain compliance simplifies regulatory adherence and protects sensitive information.

The scalability of Symphony also ensures that it can handle large datasets and complex queries, making it an indispensable tool for organizations looking to maximize the value of their LLM investments.

Integrating Panzura's Symphony with LLM workflows provides a secure, efficient, and compliant method for accessing and processing unstructured data. By leveraging metadata, the data broker optimizes data access while ensuring that the LLM interacts only with relevant and permissible content. This approach not only enhances the performance of LLMs and saves cloud resources (and money), but also protects organizations from data misuse and compliance risks, making it a foundational component of modern data management strategies.



Zero Trust Principled Data Broker

Allows parsing of Metadata Catalog extracted from Unstructured Data. API, SQL, Webhook requests will be authenticated individually. Full artifact requests will be re-authenticated separately.