

Fluree Data Fabric for Multi-Cloud Secure Data Sharing

Fluree is a new data collaboration platform that enables data to be securely shared across multiple on-prem and cloud environments without copying.



The Challenge

Organizations Need to Support Multiple Cloud and On Prem Environments for Apps and Data Support

Over the last few years, large and heavily-regulated enterprises have accelerated their adoption of the cloud to host their information technology services.

In parallel, regulators have been active in highlighting and providing targeted guidance around the risks of dependency with a single-cloud vendor, resulting in organizations moving to a multi-cloud strategy.

One emerging and common pattern in the marketplace is that organizations are selecting one cloud vendor for Application Development and to host business app databases, while another independent cloud vendor is selected to provide data warehousing and advanced analytics.



The Impact

Synchronizing Data Across Multiple Clouds Costs Time and Money

An unintended consequence of this hybrid, multi-cloud design is the extremely high cost and latency resulting from trying to synchronize operational data sources resident in one cloud providers with the analytic warehouses on another cloud.

At terabyte scale, this could result in: (1) millions of dollars of overhead (ingress/egress costs from on-prem to cloud); (2) timeliness issues (data can easily be out of sync for months); and (3) data quality and fidelity issues arising from the volume of extraction, transformation and loading jobs that need to run to keep data in sync.

The expected net benefits and ROI for enterprise cloud initiatives have very often been coming in far below expectations.



The Solution

Multi-Cloud Data Fabric for Zero-Copy Integration

At Fluree, we believe strongly in a vision of data being hosted anywhere, once, but being available to anyone within policy – without the need for copying and without sacrificing security and control. Fluree is an ideal solution for dealing with data on-premises and across multiple cloud providers.

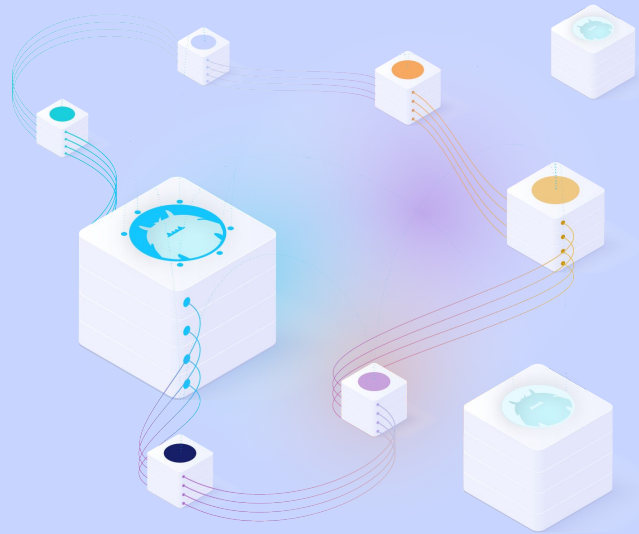
Fluree Data Fabric for Multi-Cloud Secure Data Sharing



Fluree's Approach

In the Fluree data fabric solution, each process that creates data would write their data and persist it in its local environment (e.g., on-prem file system or Cloud Storage Object on AWS S3, Azure Data Lake, or Google Cloud).

Fluree uses distributed ledger technology (DLT) to cryptographically broadcast the availability of data to members of a data collaboration network. Members in the network can be other business applications that need to consume data operationally, or analytic processes (e.g., data scientists, models or algorithms) that need to perform more complex analytic queries.



Dynamic Real-Time Query Across Clouds

Member consumers can dynamically spin-up query services on the fly to query and consume data using IT resources from its local environment. This can be from a different cloud or on-prem environment from the source environment where the data is hosted. Data can be queried and combined across multiple systems without having to be cleansed and copied. Data can remain in-place while servicing consumers.



Zero-Trust, Zero-Copy Integration

Fluree is a "Zero Trust" solution, meaning that at the point of execution, each request to consume data from the ledger must be authenticated from an authoritative identity source, and must be authorized via Smart Data Contracts that are built directly into the ledger. These smart contracts can be used to enforce ownership, creation, viewing, editing, retention and destruction permissions for any object of data on the ledger. Using the Fluree approach, data never needs to move from its original locations; instead, consumers interact with cryptographically-hashed pointers on the ledger that refer back to the original source data at query time based on policy.



Global Interoperability Across Systems

Fluree leverages Knowledge Graph technology to activate data using the power of semantic ontologies. This provides instant translational interoperability of data across systems, allowing data to be easily understood by data scientists, algorithms and AI models that may not have understood the data in its source schema or form.

Contact us to get started: sales@flur.ee