Apache Iceberg
Apache Parquet

Kristo Raun

Data Engineering 2024 Fall

UNIVERSITY OF TARTU
1632

Apache Parquet

- Apache Parquet is an open source, column-oriented data file format designed for efficient data storage and retrieval.
- It provides high performance compression and encoding schemes to handle complex data in bulk and is supported in many programming language and analytics tools.

https://parquet.apache.org/

Apache Parquet

- Columnar storage
- Compression
- Open source
- Language agnostic
- Complex data types

https://parquet.apache.org/

Apache Parquet

- Row groups

	Product	Customer	Country	Date	Sales Amount
Row Group 1	Ball	John Doe	USA	2023-01-01	100
	T-Shirt	John Doe	USA	2023-01-02	200
Row Group 2	Socks	Maria Adams	UK	2023-01-01	300
	Socks	Antonio Grant	USA	2023-01-03	100
Row Group 3	T-Shirt	Maria Adams	UK	2023-01-02	500
	Socks	John Doe	USA	2023-01-05	200

https://data-mozart.com/parquet-file-format-everything-you-need-to-know/

Apache Parquet

- Encodings

Product	Index
Ball	0
T-Shirt	1
Socks	2
Socks	2
T-Shirt	1
Socks	2

Index	Product
0	Ball
1	T-Shirt
2	Socks

https://data-mozart.com/parquet-file-format-everything-you-need-to-know/

Apache Iceberg

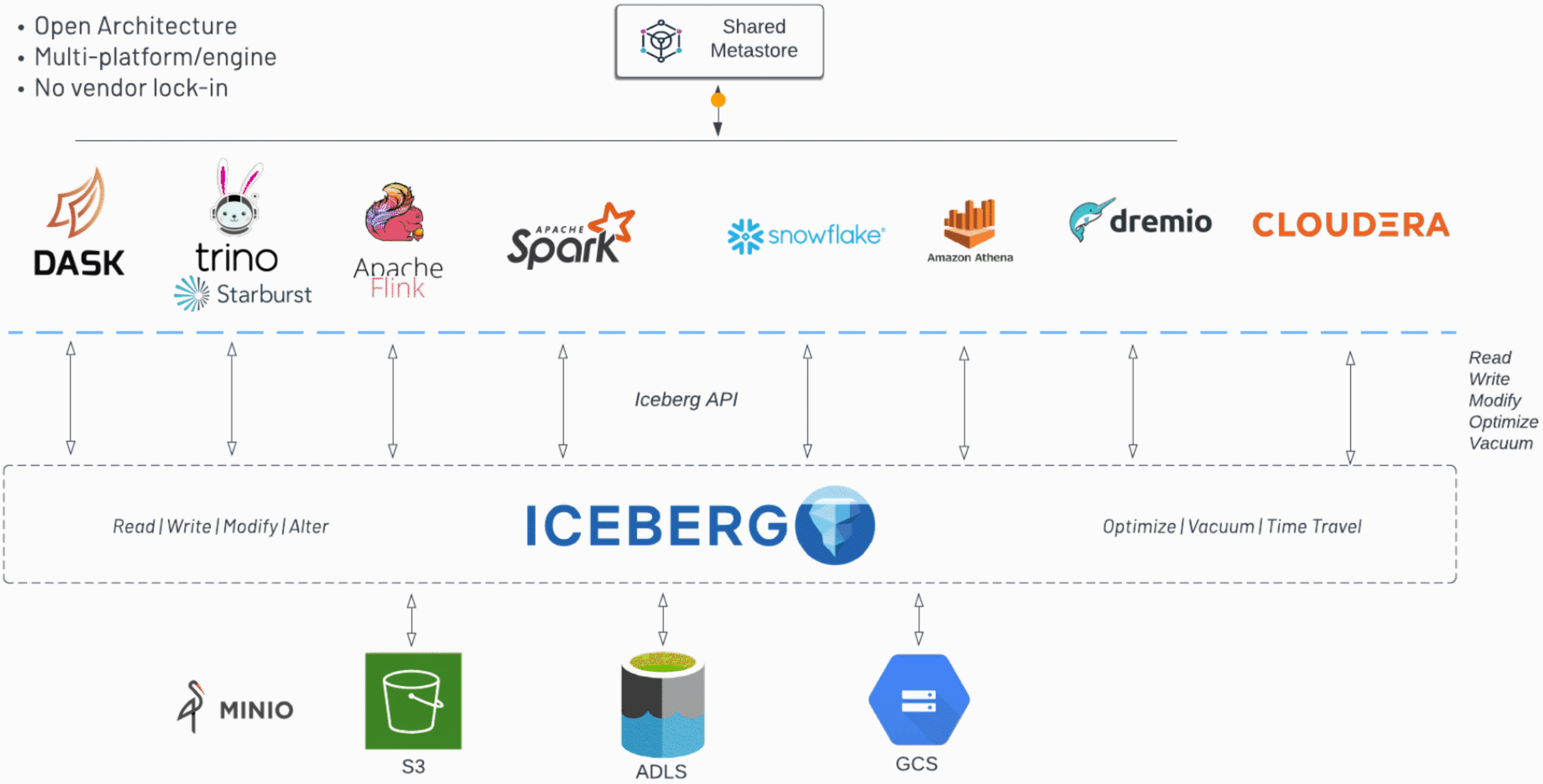
- Iceberg is a high-performance format for huge analytic tables.
- Iceberg brings the reliability and simplicity of SQL tables to big data, while making it possible for engines like Spark, Trino, Flink, Presto, Hive and Impala to safely work with the same tables, at the same time.

https://iceberg.apache.org/

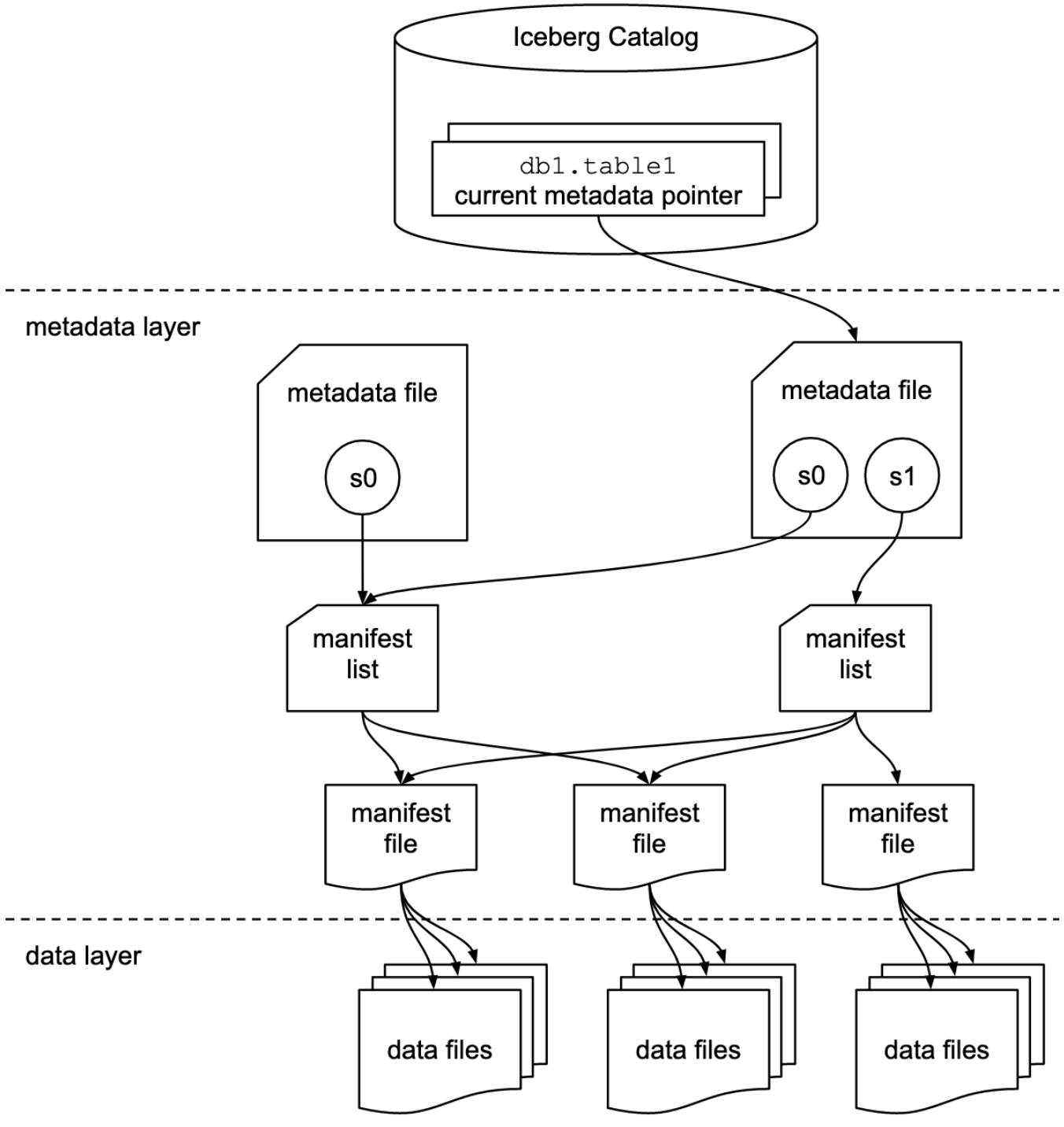
Apache Iceberg

- Schema evolution
  - supports add, drop, update, or rename
- Hidden partitioning
  - prevents user mistakes that cause silently incorrect results or extremely slow queries
- Partition layout evolution
  - can update the layout of a table as data volume or query patterns change
- Time travel
  - enables reproducible queries that use exactly the same table snapshot, or lets users easily examine changes
- Version rollback
  - allows users to quickly correct problems by resetting tables to a good state

- Open Architecture
- Multi-platform/engine
- No vendor lock-in



https://www.starburst.io/blog/introduction-to-apache-iceberg-in-trino/



Catalog is responsible for two things:

- maintain the current metadata location
- track table names and namespaces

Table state is maintained in metadata files.

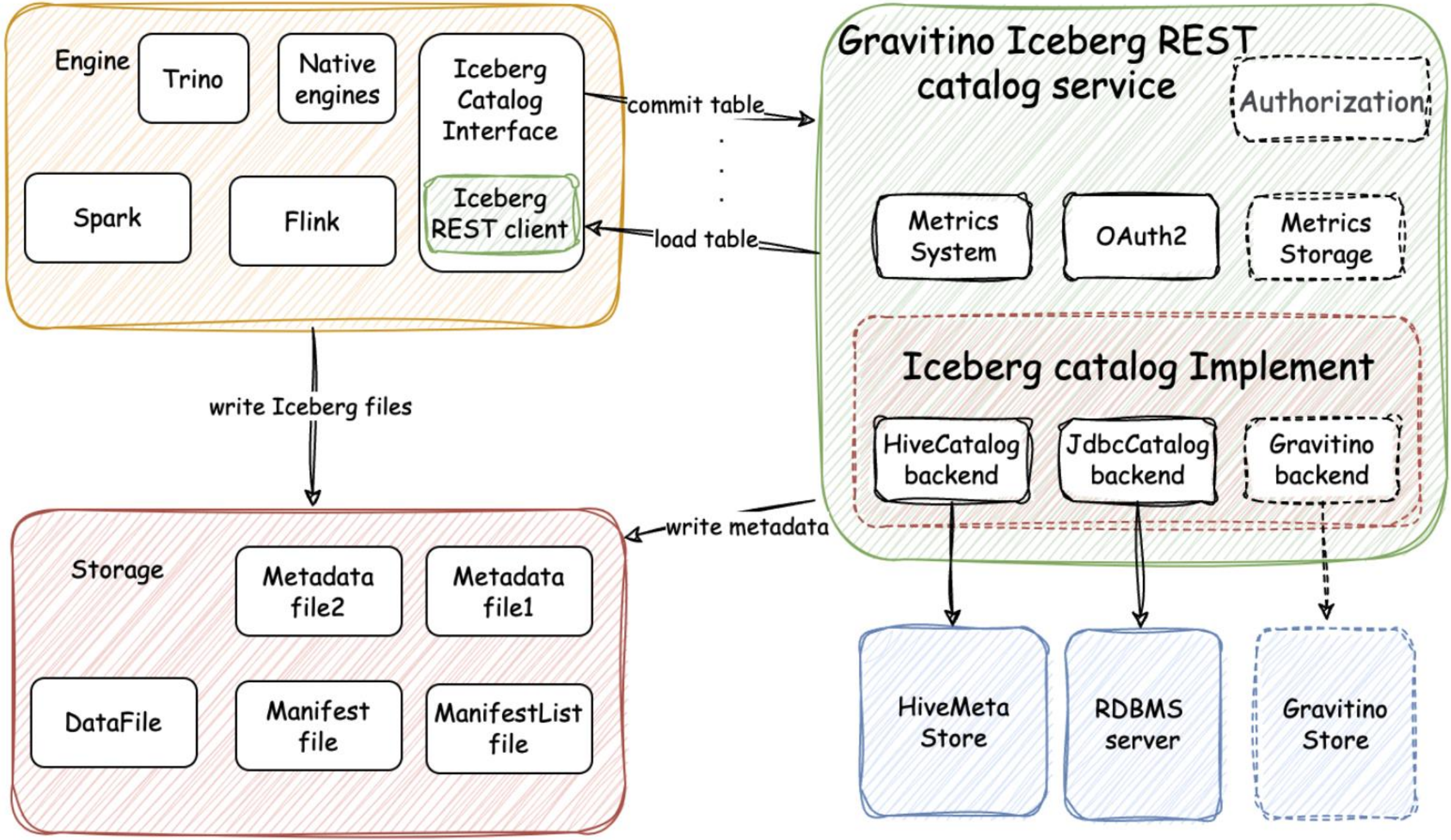
Metadata files:

- table schema
- partitioning config
- custom properties
- snapshots.

A snapshot represents the state of a table at some time and is used to access the complete set of data files in the table.

Data files in snapshots are tracked by one or more manifest files that contain a row for each data file in the table, the file's partition data, and its metrics.

The data in a snapshot is the union of all files in its manifests.



https://datastrato.ai/blog/gravitino-iceberg-rest-catalog-service/