

Apache Iceberg is an open table format for huge analytic datasets. Iceberg adds tables to compute engines including Spark, Trino, PrestoDB, Flink, Hive and Impala using a high-performance table format that works just like a SQL table.

User experience

Iceberg avoids unpleasant surprises. Schema evolution works and won't inadvertently un-delete data. Users don't need to know about partitioning to get fast queries.

Schema evolution supports add, drop, update, or rename, and has no side-effects

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