

# IcebergProposal

## Iceberg Proposal

### Abstract

Iceberg is a table format for large, slow-moving tabular data.

It is designed to improve on the de-facto standard table layout built into Apache Hive, Presto, and Apache Spark.

### Proposal

The purpose of Iceberg is to provide SQL-like tables that are backed by large sets of data files. Iceberg is similar to the Hive table layout, the de-facto standard structure used to track files in a table, but provides additional guarantees and performance optimizations:

- Atomicity - Each change to the table is will be complete or will fail. "Do or do not. There is no try."
- Snapshot isolation - Reads use one and only one snapshot of a table at some time without holding a lock.
- Safe schema evolution - A table's schema can change in well-defined ways, without breaking older data files.
- Column projection - An engine may request a subset of the available columns, including nested fields.
- Predicate pushdown - An engine can push filters into read planning to improve performance using partition data and file-level statistics.

Iceberg does NOT define a new file format. All data is stored in Apache Avro, Apache ORC, or Apache Parquet files.

Additionally, Iceberg is designed to work well when data files are stored in cloud blob stores, even when those systems provide weaker guarantees than a file system, including:

- Eventual consistency in the namespace
- High latency for directory listings
- No renames of objects
- No folder hierarchy

### Rationale

Initial benchmarks show dramatic improvements in query planning. For example, in Netflix's Atlas use case, which stores time-series metrics from Netflix runtime systems and 1 month is stored across 2.7 million files in 2,688 partitions:

- Hive table using Parquet:
  - 400k+ splits, not combined
  - Explain query: 9.6 minutes wall time (planning only)
- Iceberg table with partition filtering:
  - 15,218 splits, combined
  - Planning: 10 seconds
  - Query wall time: 13 minutes
- Iceberg table with partition and min/max filtering:
  - 412 splits
  - Planning: 25 seconds
  - Query wall time: 42 seconds

These performance gains combined with the cross-engine compatibility are a very compelling story.

### Initial Goals

The initial goal will be to move the existing codebase to Apache and integrate with the Apache development process and infrastructure. A primary goal of incubation will be to grow and diversify the Iceberg community. We are well aware that the project community is largely comprised of individuals from a single company. We aim to change that during incubation.

### Current Status

As previously mentioned, Iceberg is under active development at Netflix, and is being used in processing large volumes of data in Amazon EC2.

Iceberg license documentation is already based on Apache guidelines for LICENSE and NOTICE content.

### Meritocracy

We value meritocracy and we understand that it is the basis for an open community that encourages multiple companies and individuals to contribute and be invested in the project's future. We will encourage and monitor participation and make sure to extend privileges and responsibilities to all contributors.

### Community

Iceberg is currently being used by developers at Netflix and a growing number of users are actively using it in production environments. Iceberg has received contributions from developers working at Hortonworks, [WeWork](#), and Palantir. By bringing Iceberg to Apache we aim to assure current and future contributors that the Iceberg community is meritocratic and open, in order to broaden and diversify the user and developer community.

## Core Developers

Iceberg was initially developed at Netflix and is under active development. We believe Netflix will be of interest to a broad range of users and developers and that incubating the project at the ASF will help us build a diverse, sustainable community.

## Alignment

Iceberg utilizes other Apache projects such as Avro, Hadoop, Hive, ORC, Parquet, Pig, and Spark. We anticipate integration with additional Apache projects as the Iceberg community and interest in the project grows.

## Known Risks

### Orphaned Products

Netflix is committed to the future development of Iceberg and understands that graduation to a TLP, while preferable, is not the only positive outcome of incubation.

Should the Iceberg project be accepted by the Incubator, the prospective PPMC would be willing to agree to a target incubation period of 2 years or less, knowing that every Incubator project incurs a certain cost in terms of ASF infrastructure and volunteer time.

### Inexperience with Open Source

Three of the initial committers are Apache members and Incubator PMC members. They will work with the other community members to teach them the Apache Way.

### Homogenous Developers

The majority of the committers work at Netflix, though we are committed to recruiting and developing additional committers from a wide spectrum of industries and backgrounds.

### Reliance on Salaried Developers

It is expected that Iceberg development will occur on both salaried time and on volunteer time, after hours. Most of the initial committers are paid by Netflix to contribute to this project. However, they are all passionate about the project, and we are both confident and hopeful that the project will continue even if no salaried developers contribute to the project.

### Relationships with Other Apache Products

As mentioned in the Rationale section, Iceberg utilizes a number of existing Apache projects (Avro, Hadoop, Hive, ORC, Parquet, Pig, & Spark), and we expect that list to expand as the community grows and diversifies. Any Apache project in the big data space that needs to store or process tabular data would be potentially relevant.

### An Excessive Fascination with the Apache Brand

We are applying to the Incubator process because we think it is the next logical step for the Iceberg project after open-sourcing the code. This proposal is not for the purpose of generating publicity. Rather, we want to make sure to create a very inclusive and meritocratic community, outside the umbrella of a single company. Netflix has a long history of contributing to Apache projects and the Iceberg developers and contributors understand the implication of making it an Apache project.

## Required Resources

### Mailing lists

- [dev@iceberg.incubator.apache.org](mailto:dev@iceberg.incubator.apache.org)
- [commits@iceberg.incubator.apache.org](mailto:commits@iceberg.incubator.apache.org)
- [private@iceberg.incubator.apache.org](mailto:private@iceberg.incubator.apache.org)

The podling may also create a user mailing list, if needed.

### Source Control and Issue Tracking

The Iceberg podling would use Apache's gitbox integration to sync between github and Apache infrastructure. The podling would use github issues and pull requests for community engagement.

## Current Resources

- Initial source: <https://github.com/Netflix/iceberg>
- Java documentation: <https://netflix.github.io/iceberg/current/javadoc/index.html?com/netflix/iceberg/package-summary.html>
- Table specification: [https://docs.google.com/document/d/1Q-zL5IScle6NEEfyYsXYzX\\_Q8Qf0ctMyGBKslOswA/edit](https://docs.google.com/document/d/1Q-zL5IScle6NEEfyYsXYzX_Q8Qf0ctMyGBKslOswA/edit)

## Source and Intellectual Property Submission Plan

The Iceberg source code in Github is currently licensed under Apache License v2.0 and the copyright is assigned to Netflix. If Iceberg becomes an Incubator project at the ASF, Netflix will transfer the source code and trademark ownership to the Apache Software Foundation via a Software Grant Agreement.

## External Dependencies

External dependencies licensed under Apache License 2.0

- Guava <https://github.com/google/guava>
- Jackson <https://github.com/FasterXML/jackson-core>
- Joda-Time <http://www.joda.org/joda-time/>

External dependencies licensed under the MIT License

- SLF4J <https://www.slf4j.org/>
- Mockito <https://github.com/mockito/mockito>

ASF Projects

- Apache Avro
- Apache Hadoop
- Apache Hive
- Apache ORC
- Apache Parquet
- Apache Pig
- Apache Spark

## Cryptography

We do not expect Iceberg to be a controlled export item due to the use of encryption.

## Initial Committers and Affiliations

- Ryan Blue [blue@apache.org](mailto:blue@apache.org) (Netflix)
- Parth Brahmhatt [parth@apache.org](mailto:parth@apache.org) (Netflix)
- Julien Le Dem [julien@apache.org](mailto:julien@apache.org) (WeWork)
- Owen O'Malley [omalley@apache.org](mailto:omalley@apache.org) (Hortonworks)
- Daniel Weeks [dweeks@apache.org](mailto:dweeks@apache.org) (Netflix)
- James Taylor [jamestaylor@apache.org](mailto:jamestaylor@apache.org) (Lyft)

## Sponsors and Nominated Mentors

- Champion and mentor: Owen O'Malley [omalley@apache.org](mailto:omalley@apache.org)
- Mentor: Ryan Blue [blue@apache.org](mailto:blue@apache.org)
- Mentor: Julien Le Dem [julien@apache.org](mailto:julien@apache.org)
- Mentor: James Taylor [jamestaylor@apache.org](mailto:jamestaylor@apache.org)

## Sponsoring Entity

The Apache Incubator