Cloudera Public Cloud

Data services tools

Date published: 2019-08-22
Date modified: 2025-08-18

https://docs.cloudera.com/

Legal Notice

Contents

Apache Iceberg in Cloudera

Apache Iceberg is a high-performance open table format for organizing petabyte-scale analytic datasets on a file system or object store. Combined with Cloudera, users can build an open data lakehouse architecture for multi-function analytics and to deploy large scale end-to-end pipelines.

Open data lakehouse on Cloudera simplifies advanced analytics on all data with a unified platform for structured and unstructured data and integrated data services to enable any analytics use case from ML, BI to stream analytics and real-time analytics. Apache Iceberg is the secret sauce of the open lakehouse.

The following table shows the support for Iceberg in Cloudera and below the table Iceberg versions v1 and v2 are defined:

Table 1: Iceberg Support Matrix

Release	Iceberg support level	SQL Engine				
		Impala	Hive	Spark	NiFi	Flink
Cloudera Data Hub						
Cloudera Data Hub 7.3.1	GA	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert	v1, v2: create table, read, append, overwrite ***
Cloudera Data Hub 7.2.18	GA	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert	v1, v2: create table, read, append, overwrite ***
Cloudera Data Hub 7.2.17	GA	v1, v2: create table, read, insert, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert	v1, v2: create table, read, append, overwrite ***
Cloudera Data Hub 7.2.16.2	GA	v1, v2: create table, read,	v1: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, and delete	v1, v2: read and insert	v1: create table, read, and insert
Cloudera Base on premises						
Cloudera Base on premises 7.3.1	GA	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert	v1, v2: create table, read, and insert
Cloudera Private Cloud Base 7.1.9	GA	v1, v2: create table, read, insert, partition evolution (hidden partitioning)	No Iceberg support	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert	v1, v2: create table, read, and insert
Cloudera Private Cloud Base 7.1.7 SP2, 7.1.8	No Iceberg support					

Release	Iceberg support level	SQL Engine				
		Impala	Hive	Spark	NiFi	Flink
Cloudera Data Services on cloud						
Cloudera Data Services on cloud	GA	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert	N/A
Cloudera Data Services on premises						
Cloudera Data Services on premises 1.5.4	GA (Cloudera Private Cloud Base 7.1.9) Technical Preview (Cloudera Private Cloud Base 7.1.7, 7.1.8)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert (Cloudera Private Cloud Base 7.1.9)	v1, v2: create table, read, and insert (Cloudera Private Cloud Base 7.1.9)
Cloudera Data Services on premises 1.5.3	GA (Cloudera Private Cloud Base 7.1.9) Technical Preview (Cloudera Private Cloud Base 7.1.7, 7.1.8)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert (Cloudera Private Cloud Base 7.1.9)	v1, v2: create table, read, and insert (Cloudera Private Cloud Base 7.1.9)
Cloudera Data Services on premises 1.5.2	GA (Cloudera Private Cloud Base 7.1.9) Technical Preview (Cloudera Private Cloud Base 7.1.7, 7.1.8)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, delete, partition evolution (hidden partitioning)	v1, v2: read and insert (Cloudera Private Cloud Base 7.1.9)	v1, v2: create table, read, and insert (Cloudera Private Cloud Base 7.1.9)
Cloudera Data Services on premises 1.5.1 2023.0.13.0-20	Technical Preview (Cloudera Private Cloud Base 7.1.7, 7.1.8)	v1, v2: create table, read	v1, v2: create table, read, insert, delete, partition evolution (hidden partitioning)	v1, v2: create table, read, insert, update, and delete	No Cloudera on premises support	No Cloudera on premises support

** The support for delete operations, except from Flink, shown in this table is limited to position deletes. Equality deletes are not supported in these releases except from Flink.

*** Iceberg v2 updates and deletes from Flink are a technical preview in Cloudera on cloud 7.2.17.

The Apache Iceberg format specification describes the following versions of tables:

- v1

  Defines large analytic data tables using open format files.
- v2

  Specifies ACID compliant tables including row-level deletes and updates.

Table 2: Iceberg Docs and Availability Matrix

Release	Docs	Iceberg Support Level
Cloudera Base on premises 7.3.1	Iceberg support for Hive	GA
Open Data Lakehouse (Cloudera Private Cloud Base 7.1.9)	Iceberg in Open Data Lakehouse	GA
	Iceberg support for Atlas	GA

Release	Docs	Iceberg Support Level
	SQL Stream Builder with Iceberg (CSA 1.11) and Flink with Iceberg (CSA 1.11) Iceberg replication policies	GA
Cloudera Data Engineering on cloud	Using Iceberg	GA
Cloudera Data Warehouse on cloud	Iceberg features	GA
Cloudera Data Engineering on premises	Using Iceberg	Technical Preview
Cloudera Data Warehouse on premises	Iceberg introduction Moving data into Iceberg tables on Cloudera Data Warehouse	GA (Cloudera Base on premises 7.1.9), Technical Preview (Cloudera Base on premises 7.1.7-7.1.8)
Cloudera Data Hub 7.2.16 and later	Iceberg features Iceberg features	Technical Preview
Cloudera Data Hub 7.2.17 and later	Iceberg in Apache Atlas Iceberg in Apache Atlas	Technical Preview
	Streaming Analytics Iceberg support in Flink	GA
	Flink/Iceberg connector	GA
	Using NiFi to ingest data into Cloudera Data Warehouse on cloud in Iceberg table format	GA
Cloudera Data Hub 7.2.18	Iceberg in Apache Atlas Iceberg in Apache Atlas	GA
Cloudera Data Hub 7.3.1	Iceberg features No new features in this release	GA
Cloudera Flow Management for Cloudera on premises	Using the PutIcebergCDC processor	Technical Preview
	Using NiFi to ingest data into Cloudera Data Warehouse on premises in Iceberg table format	GA
Cloudera DataFlow	Using the Kafka to Apache Iceberg ReadyFlow	GA
Cloudera AI on cloud	Connection to Iceberg	GA