

[cloudera@quickstart Desktop]\$ hbase shell

2025-05-07 06:26:32,488 INFO [main] Configuration.deprecation: hadoop.native.lib is deprecated.
Instead, use io.native.lib.available

HBase Shell; enter 'help<RETURN>' for list of supported commands.

Type "exit<RETURN>" to leave the HBase Shell

Version 1.2.0-cdh5.13.0, rUnknown, Wed Oct 4 11:16:18 PDT 2017

Hbase database creation

hbase(main):001:0> create 'flight_info','details','schedule','delay'

0 row(s) in 2.7600 seconds

=> Hbase::Table - flight_info

hbase(main):002:0> put 'flight_info','flight1','details:airlines','Delta'

0 row(s) in 0.1740 seconds

hbase(main):003:0> put 'flight_info','flight1','details:source','JFK'

0 row(s) in 0.0140 seconds

hbase(main):004:0> put 'flight_info','flight1','details:destination','LAX'

0 row(s) in 0.0260 seconds

hbase(main):005:0> put 'flight_info','flight1','schedule:departure_time','2025-05-01 08:00'

0 row(s) in 0.0140 seconds

hbase(main):006:0> put 'flight_info','flight1','schedule:arrival_time','2025-05-01 11:00'

0 row(s) in 0.0120 seconds

hbase(main):007:0> put 'flight_info','flight1','delay:dep_delay','15'

0 row(s) in 0.0130 seconds

hbase(main):008:0> put 'flight_info','flight1','delay:arr_delay','10'

0 row(s) in 0.0130 seconds

hbase(main):009:0> **alter 'flight_info','add','weather'**

Updating all regions with the new schema...

0/1 regions updated.

1/1 regions updated.

Done.

Updating all regions with the new schema...

1/1 regions updated.

Done.

0 row(s) in 5.1120 seconds

[cloudera@quickstart Desktop]\$ hive

Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties

WARNING: Hive CLI is deprecated and migration to Beeline is recommended.

Mapping Hive and executing queries

hive> **CREATE EXTERNAL TABLE flight_hive (**

> flight_id STRING,

> airline STRING,

> source STRING,

> destination STRING,

> departure_time STRING,

> arrival_time STRING,

> dep_delay INT,

> arr_delay INT

>)

> STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'

> WITH SERDEPROPERTIES (

> 'hbase.columns.mapping' = ":key,details:airline,details:source,details:destination,

> schedule:departure_time,schedule:arrival_time,

```
>          delay:dep_delay,delay:arr_delay"
> )
> TBLPROPERTIES ("hbase.table.name" = "flight_info");
```

OK

Time taken: 1.312 seconds

hive> **SELECT * FROM flight_hive**

```
> ;
```

OK

flight1	NULL	JFK	LAX	2025-05-01 08:00	2025-05-01 11:00	15	10
---------	------	-----	-----	------------------	------------------	----	----

Time taken: 1.683 seconds, Fetched: 1 row(s)

```
hive> SELECT SUM(dep_delay) AS total_departure_delay
> FROM flight_hive;
```

Query ID = cloudera_20250507064848_77409f07-bf3f-42be-a191-ecf9eff50a22

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks determined at compile time: 1

In order to change the average load for a reducer (in bytes):

```
set hive.exec.reducers.bytes.per.reducer=<number>
```

In order to limit the maximum number of reducers:

```
set hive.exec.reducers.max=<number>
```

In order to set a constant number of reducers:

```
set mapreduce.job.reduces=<number>
```

Starting Job = job_1744085536392_0008, Tracking URL =
http://quickstart.cloudera:8088/proxy/application_1744085536392_0008/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1744085536392_0008

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2025-05-07 06:49:25,154 Stage-1 map = 0%, reduce = 0%

2025-05-07 06:49:42,130 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.12 sec

2025-05-07 06:49:51,670 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.41 sec

MapReduce Total cumulative CPU time: 4 seconds 410 msec

Ended Job = job_1744085536392_0008

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.41 sec HDFS Read: 9022 HDFS Write: 3
SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 410 msec

OK

15

Time taken: 66.159 seconds, Fetched: 1 row(s)

hive> SELECT AVG(dep_delay) AS avg_departure_delay

> FROM flight_hive;

Query ID = cloudera_20250507065050_6eb79164-40af-4703-a0ba-daed64abebaf

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks determined at compile time: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>

Starting Job = job_1744085536392_0009, Tracking URL =
http://quickstart.cloudera:8088/proxy/application_1744085536392_0009/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1744085536392_0009

Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1

2025-05-07 06:51:19,669 Stage-1 map = 0%, reduce = 0%

2025-05-07 06:51:31,922 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.65 sec

2025-05-07 06:51:43,148 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.05 sec

MapReduce Total cumulative CPU time: 4 seconds 50 msec

Ended Job = job_1744085536392_0009

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.05 sec HDFS Read: 9623 HDFS Write: 5
SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 50 msec

OK

15.0

Time taken: 52.513 seconds, Fetched: 1 row(s)

hbase(main):010:0> CREATE INDEX idx_airline ON flight_info (details:airline);