Name – Atharv Papat

Roll No – I3226

Div - 2

hdoop@hadoop:~/apache-hive-3.1.2-bin$ jps

38256 NodeManager

38149 ResourceManager

37766 DataNode

39337 Jps

37916 SecondaryNameNode

37661 NameNode

hdoop@hadoop:~/apache-hive-3.1.2-bin$

hdoop@hadoop:~/apache-hive-3.1.2-bin$

hdoop@hadoop:~/apache-hive-3.1.2-bin$ hive

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/home/hdoop/hadoop-3.2.3/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.

SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

Hive Session ID = 254d3148-6128-405b-9f77-b2567dd82052

Logging initialized using configuration in jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/hive-common-3.1.2.jar!/hive-log4j2.properties Async: true

Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.

Hive Session ID = 94d47b2a-f539-409d-93ae-33d7eb514147

hive> show databases;

OK

default

```
Time taken: 1.845 seconds, Fetched: 1 row(s)
hive> create database db1;
OK
Time taken: 0.755 seconds
hive> use db1;
OK
Time taken: 0.139 seconds
hive> create table flight (fno int, year int, dest varchar(10),delay float);
OK
Time taken: 2.275 seconds
hive> alter table flight rename to air_flight;
OK
Time taken: 2.493 seconds
hive> alter table air_flight add columns (source varchar(10));
OK
Time taken: 0.634 seconds
hive> create table flight (fno int, year int, dest varchar(10),delay float);
OK
Time taken: 0.414 seconds
hive> insert into flight values (123, 2009, "Mumbai", 30.0);
Query ID = hdoop_20230504141134_16689223-bc74-4fce-95a2-804b352e56fa
Total jobs = 3
Launching Job 1 out of 3
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_1683187088322_0001, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0001/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0001

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

2023-05-04 14:12:25,108 Stage-1 map = 0%, reduce = 0%

2023-05-04 14:12:41,935 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.14 sec

2023-05-04 14:12:55,147 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.19 sec

MapReduce Total cumulative CPU time: 7 seconds 190 msec

Ended Job = job_1683187088322_0001

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/flight/.hive-staging_hive_2023-05-04_14-11-34_998_5346237385295105948-1/-ext-10000

Loading data to table db1.flight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.19 sec HDFS Read: 19909 HDFS Write: 329 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 190 msec

OK

Time taken: 84.798 seconds

hive> insert into flight values (342, 2008, "Nagpur", 13.0);

Query ID = hdoop_20230504141316_63930241-aea8-4d84-bad9-faed85f06897

Total jobs = 3

Launching Job 1 out of 3

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_1683187088322_0002, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0002/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0002

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

2023-05-04 14:13:40,889 Stage-1 map = 0%, reduce = 0%

2023-05-04 14:13:56,135 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.28 sec

2023-05-04 14:14:09,146 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.52 sec

MapReduce Total cumulative CPU time: 7 seconds 520 msec

Ended Job = job_1683187088322_0002

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/flight/.hive-staging_hive_2023-05-04_14-13-16_751_8156403126521870125-1/-ext-10000

Loading data to table db1.flight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.52 sec HDFS Read: 19959 HDFS Write: 329 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 520 msec

OK

Time taken: 56.166 seconds

hive> insert into flight values (232, 2008, "Pune", 0.0);

Query ID = hdoop_20230504142230_5abe653d-50ad-4329-a1a2-7168b9c3996c

Total jobs = 3

Launching Job 1 out of 3

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_1683187088322_0003, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0003/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0003

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

2023-05-04 14:22:58,880 Stage-1 map = 0%, reduce = 0%

2023-05-04 14:23:16,732 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.26 sec

2023-05-04 14:23:33,155 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.6 sec

MapReduce Total cumulative CPU time: 7 seconds 600 msec

Ended Job = job_1683187088322_0003

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/flight/.hive-staging_hive_2023-05-04_14-22-30_328_1584438224102631460-1/-ext-10000

Loading data to table db1.flight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.6 sec HDFS Read: 19953 HDFS Write: 324 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 600 msec

OK

Time taken: 66.269 seconds

hive> insert into flight values (103, 2009, "Nashik",10.0);

Query ID = hdoop_20230504142338_bb87b18e-8d6a-4ccd-b3db-37f1fd91965b

Total jobs = 3

Launching Job 1 out of 3

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_1683187088322_0004, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0004/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0004

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

2023-05-04 14:24:10,496 Stage-1 map = 0%, reduce = 0%

2023-05-04 14:24:25,686 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.95 sec

2023-05-04 14:24:39,851 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.43 sec

MapReduce Total cumulative CPU time: 7 seconds 430 msec

Ended Job = job_1683187088322_0004

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/flight/.hive-staging_hive_2023-05-04_14-23-38_278_3471742521069876157-1/-ext-10000

Loading data to table db1.flight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.43 sec HDFS Read: 19964 HDFS Write: 329 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 430 msec

OK

Time taken: 66.552 seconds

hive> select * from flight;

OK

```
123     2009    Mumbai          30.0
342     2008    Nagpur 13.0
232     2008    Pune    0.0
103     2009    Nashik  10.0
Time taken: 0.735 seconds, Fetched: 4 row(s)
hive> create table nflight (fno int, year int,source varchar(10))
    > row format delimited
    > fields terminated by ','
    > lines terminated by '\n'
    > stored as textfile;
OK
Time taken: 0.417 seconds
hive> select * from nflight;
OK
Time taken: 0.684 seconds
hive> insert into nflight values (121, 2010, "Mumbai");
Query ID = hdoop_20230505171251_5a63fd46-c6b4-47ed-b291-a02f7ccbe56e
Total jobs = 3
Launching Job 1 out of 3
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_1683187088322_0005, Tracking URL =
http://hadoop:8088/proxy/application_1683187088322_0005/
Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job  -kill job_1683187088322_0005
```

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

2023-05-05 17:13:20,980 Stage-1 map = 0%, reduce = 0%

2023-05-05 17:13:36,087 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.98 sec

2023-05-05 17:13:49,014 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.03 sec

MapReduce Total cumulative CPU time: 7 seconds 30 msec

Ended Job = job_1683187088322_0005

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/nflight/.hive-staging_hive_2023-05-05_17-12-51_107_7146661429602371659-1/-ext-10000

Loading data to table db1.nflight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 7.03 sec   HDFS Read: 17105 HDFS Write: 286 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 30 msec

OK

Time taken: 61.548 seconds

hive> insert into nflight values (133, 2003, "Pen");

Query ID = hdoop_20230505171411_a5388182-27f3-47d6-9472-bf0ee55d067a

Total jobs = 3

Launching Job 1 out of 3

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_1683187088322_0006, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0006/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0006

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

2023-05-05 17:15:00,378 Stage-1 map = 0%, reduce = 0%

2023-05-05 17:15:36,458 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.12 sec

2023-05-05 17:15:59,814 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.26 sec

MapReduce Total cumulative CPU time: 7 seconds 260 msec

Ended Job = job_1683187088322_0006

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/nflight/.hive-staging_hive_2023-05-05_17-14-11_516_7950959861939931640-1/-ext-10000

Loading data to table db1.nflight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.26 sec HDFS Read: 17095 HDFS Write: 283 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 260 msec

OK

Time taken: 115.31 seconds

hive> insert into nflight values (191, 2003, "Delhi");

Query ID = hdoop_20230505171620_5826e07e-f89f-48fa-9429-cbe87b1faeeb

Total jobs = 3

Launching Job 1 out of 3

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_1683187088322_0007, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0007/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0007

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

2023-05-05 17:17:10,702 Stage-1 map = 0%, reduce = 0%

2023-05-05 17:17:48,481 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.11 sec

2023-05-05 17:18:11,865 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.25 sec

MapReduce Total cumulative CPU time: 7 seconds 250 msec

Ended Job = job_1683187088322_0007

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/nflight/.hive-staging_hive_2023-05-05_17-16-20_329_6148931622642876724-1/-ext-10000

Loading data to table db1.nflight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.25 sec HDFS Read: 17105 HDFS Write: 285 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 250 msec

OK

Time taken: 116.458 seconds

hive> insert into nflight values (103, 2008, "Pune");

Query ID = hdoop_20230505171846_0ac0a2ba-ce66-4498-bc91-25fcf9cc9737

Total jobs = 3

Launching Job 1 out of 3

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_1683187088322_0008, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0008/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0008

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

2023-05-05 17:19:15,664 Stage-1 map = 0%, reduce = 0%

2023-05-05 17:19:34,199 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.46 sec

2023-05-05 17:19:51,634 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.73 sec

MapReduce Total cumulative CPU time: 7 seconds 730 msec

Ended Job = job_1683187088322_0008

Stage-4 is selected by condition resolver.

Stage-3 is filtered out by condition resolver.

Stage-5 is filtered out by condition resolver.

Moving data to directory hdfs://localhost:9000/user/hive/warehouse/db1.db/nflight/.hive-staging_hive_2023-05-05_17-18-47_005_2789327012668292123-1/-ext-10000

Loading data to table db1.nflight

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.73 sec HDFS Read: 17105 HDFS Write: 284 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 730 msec

OK

Time taken: 69.443 seconds

hive> select * from flight;

OK

123 2009 Mumbai 30.0

342 2008 Nagpur 13.0

232 2008 Pune 0.0

103 2009 Nashik 10.0

Time taken: 0.572 seconds, Fetched: 4 row(s)

hive> select a.fno, a.year, a.dest, a.delay, b.source

> from flight a join nflight b

> on (a.fno = b.fno) ;

Query ID = hdoop_20230505172131_b830904b-0599-41fc-9ca3-b5fddd7a3b3a

Total jobs = 1

SLF4J: Found binding in [jar:file:/home/hdoop/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]

Execution completed successfully

MapredLocal task succeeded

Launching Job 1 out of 1

Number of reduce tasks is set to 0 since there's no reduce operator

Starting Job = job_1683187088322_0009, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0009/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0009

Hadoop job information for Stage-3: number of mappers: 1; number of reducers: 0

2023-05-05 17:22:38,387 Stage-3 map = 0%, reduce = 0%

2023-05-05 17:22:57,829 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 4.29 sec

MapReduce Total cumulative CPU time: 4 seconds 290 msec

Ended Job = job_1683187088322_0009

MapReduce Jobs Launched:

Stage-Stage-3: Map: 1 Cumulative CPU: 4.29 sec HDFS Read: 10487 HDFS Write: 125 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 290 msec

OK

103 2009 Nashik 10.0 Pune

Time taken: 89.034 seconds, Fetched: 1 row(s)

hive> select avg(delay) from flight where year = 2008;

Query ID = hdoop_20230505172311_233e13fd-de47-4dd5-a2d0-34eb011e432f

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_1683187088322_0010, Tracking URL = http://hadoop:8088/proxy/application_1683187088322_0010/

Kill Command = /home/hdoop/hadoop-3.2.3/bin/mapred job -kill job_1683187088322_0010

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

2023-05-05 17:23:33,609 Stage-1 map = 0%, reduce = 0%

2023-05-05 17:23:52,353 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.19 sec

2023-05-05 17:24:13,057 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.46 sec

MapReduce Total cumulative CPU time: 8 seconds 460 msec

Ended Job = job_1683187088322_0010

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 8.46 sec   HDFS Read: 16854 HDFS Write: 103 SUCCESS

Total MapReduce CPU Time Spent: 8 seconds 460 msec

OK

6.5

Time taken: 64.766 seconds, Fetched: 1 row(s)

hive>