hive (testdb)> **SELECT avg(t1.d1) from text_test t1 join text_test t2 on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;**

WARNING: 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.

Query ID = hduser_20180514190733_a259fb80-0792-437a-8356-0dda74f6c31a

Total jobs = 4

Stage-7 is selected by condition resolver.

Stage-8 is filtered out by condition resolver.

Stage-1 is filtered out by condition resolver.

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in
[jar:file:/usr/local/hadoop/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class
]

SLF4J: Found binding in
[jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLo
ggerBinder.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]

2018-05-14 19:07:37   Starting to launch local task to process map join;   maximum memory = 477626368

2018-05-14 19:07:39   Dump the side-table for tag: 1 with group count: 45031 into file:
file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-07-33_425_3413
940510929744928-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile41--.hashtable

2018-05-14 19:07:39   Uploaded 1 File to:
file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-07-33_425_3413
940510929744928-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile41--.hashtable (1136439
bytes)

2018-05-14 19:07:39   End of local task; Time Taken: 1.667 sec.

Execution completed successfully

MapredLocal task succeeded

Launching Job 2 out of 4

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

Job running in-process (local Hadoop)

2018-05-14 19:07:41,696 Stage-4 map = 100%, reduce = 0%

Ended Job = job_local316622750_0008

Launching Job 3 out of 4

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>

Job running in-process (local Hadoop)
2018-05-14 19:07:43,029 Stage-2 map = 100%, reduce = 100%
Ended Job = job_local234432737_0009
MapReduce Jobs Launched:
Stage-Stage-4: HDFS Read: 106813041 HDFS Write: 56837410 SUCCESS
Stage-Stage-2: HDFS Read: 213626082 HDFS Write: 113674820 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
_c0
3.348246
Time taken: 9.907 seconds, Fetched: 1 row(s)
hive (testdb)>

hive (testdb)> SELECT avg(t1.d1) from parquet_test_no_compression t1 join parquet_test_no_compression t2 on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;
WARNING: 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.
Query ID = hduser_20180514190919_d74e4bc1-e4bd-4405-98c5-df38f8c67ee6
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 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>
Job running in-process (local Hadoop)
2018-05-14 19:09:20,914 Stage-1 map = 100%, reduce = 100%
Ended Job = job_local1558902760_0010
Launching Job 2 out of 2
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>
Job running in-process (local Hadoop)
2018-05-14 19:09:22,235 Stage-2 map = 100%, reduce = 100%
Ended Job = job_local167135923_0011

MapReduce Jobs Launched:

Stage-Stage-1: HDFS Read: 215647758 HDFS Write: 113674820 SUCCESS

Stage-Stage-2: HDFS Read: 215647758 HDFS Write: 113674820 SUCCESS

Total MapReduce CPU Time Spent: 0 msec

OK

_c0

3.348246

Time taken: 2.973 seconds, Fetched: 1 row(s)

hive (testdb)>

hive (testdb)> SELECT avg(t1.d1) from parquet_test_gzip_compression t1 join parquet_test_gzip_compression t2 on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;

WARNING: 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.

Query ID = hduser_20180514191007_ccc192d5-a4ed-4eca-9b6c-2af77a1e162c

Total jobs = 4

Stage-7 is selected by condition resolver.

Stage-8 is filtered out by condition resolver.

Stage-1 is filtered out by condition resolver.

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in

[jar:file:/usr/local/hadoop/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class
]

SLF4J: Found binding in

[jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.5.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]

2018-05-14 19:10:12   Starting to launch local task to process map join;   maximum memory = 477626368

SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".

SLF4J: Defaulting to no-operation (NOP) logger implementation

SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.

2018-05-14 19:10:14   Dump the side-table for tag: 1 with group count: 45031 into file:

file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-10-07_798_7579
289160966824516-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile61--.hashtable

2018-05-14 19:10:14   Uploaded 1 File to:

file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-10-07_798_7579
289160966824516-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile61--.hashtable (1136439
bytes)

2018-05-14 19:10:14   End of local task; Time Taken: 2.045 sec.

Execution completed successfully

MapredLocal task succeeded

Launching Job 2 out of 4

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

Job running in-process (local Hadoop)

2018-05-14 19:10:16,496 Stage-4 map = 100%, reduce = 0%

Ended Job = job_local200617737_0012

Launching Job 3 out of 4

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>

Job running in-process (local Hadoop)

2018-05-14 19:10:17,810 Stage-2 map = 100%, reduce = 100%

Ended Job = job_local1117950550_0013

MapReduce Jobs Launched:

Stage-Stage-4: HDFS Read: 108192150 HDFS Write: 56837410 SUCCESS

Stage-Stage-2: HDFS Read: 216384300 HDFS Write: 113674820 SUCCESS

Total MapReduce CPU Time Spent: 0 msec

OK

_c0

3.348246

Time taken: 10.015 seconds, Fetched: 1 row(s)

hive (testdb)> SELECT avg(t1.d1) from parquet_test_snappy_compression t1 join parquet_test_snappy_compression t2 on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;

WARNING: 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.

Query ID = hduser_20180514191108_b2803875-a7c6-49c3-8429-a2fa694e63dc

Total jobs = 4

Stage-7 is selected by condition resolver.

Stage-8 is filtered out by condition resolver.

Stage-1 is filtered out by condition resolver.

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in

[jar:file:/usr/local/hadoop/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class

]

SLF4J: Found binding in

[jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.5.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]

2018-05-14 19:11:13   Starting to launch local task to process map join;   maximum memory = 477626368

SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".

SLF4J: Defaulting to no-operation (NOP) logger implementation

SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.

2018-05-14 19:11:15   Dump the side-table for tag: 1 with group count: 45031 into file: file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-11-08_920_9158 055368150850368-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile81--.hashtable

2018-05-14 19:11:15   Uploaded 1 File to: file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-11-08_920_9158 055368150850368-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile81--.hashtable (1136439 bytes)

2018-05-14 19:11:15   End of local task; Time Taken: 1.977 sec.

Execution completed successfully

MapredLocal task succeeded

Launching Job 2 out of 4

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

Job running in-process (local Hadoop)

2018-05-14 19:11:17,539 Stage-4 map = 100%, reduce = 0%

Ended Job = job_local693032220_0014

Launching Job 3 out of 4

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>

Job running in-process (local Hadoop)

2018-05-14 19:11:18,854 Stage-2 map = 100%, reduce = 100%

Ended Job = job_local979330050_0015

MapReduce Jobs Launched:

Stage-Stage-4: HDFS Read: 108714320 HDFS Write: 56837410 SUCCESS

Stage-Stage-2: HDFS Read: 217428640 HDFS Write: 113674820 SUCCESS

Total MapReduce CPU Time Spent: 0 msec

OK

_c0

3.348246

Time taken: 9.938 seconds, Fetched: 1 row(s)

hive (testdb)>

hive (testdb)> SELECT avg(t1.d1) from avro_test t1 join avro_test t2 on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;

WARNING: 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.

Query ID = hduser_20180514191308_45b62c17-0cdf-4101-b75f-fdd431951b78

Total jobs = 4

Stage-7 is selected by condition resolver.

Stage-8 is filtered out by condition resolver.

Stage-1 is filtered out by condition resolver.

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/usr/local/hadoop/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.5.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]

2018-05-14 19:13:13   Starting to launch local task to process map join;   maximum memory = 477626368

2018-05-14 19:13:18   Dump the side-table for tag: 1 with group count: 45031 into file: file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-13-08_703_1954 938412461675610-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile121--.hashtable

2018-05-14 19:13:18   Uploaded 1 File to: file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-13-08_703_1954 938412461675610-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile121--.hashtable (1136439 bytes)

2018-05-14 19:13:18   End of local task; Time Taken: 4.942 sec.

Execution completed successfully

MapredLocal task succeeded

Launching Job 2 out of 4

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

Job running in-process (local Hadoop)

2018-05-14 19:13:21,126 Stage-4 map = 0%, reduce = 0%

2018-05-14 19:13:24,129 Stage-4 map = 100%, reduce = 0%

Ended Job = job_local1399178643_0018

Launching Job 3 out of 4

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>
Job running in-process (local Hadoop)
2018-05-14 19:13:25,437 Stage-2 map = 100%, reduce = 100%
Ended Job = job_local162800837_0019
MapReduce Jobs Launched:
Stage-Stage-4: HDFS Read: 151089187 HDFS Write: 56837410 SUCCESS
Stage-Stage-2: HDFS Read: 302178374 HDFS Write: 113674820 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
_c0
3.348246
Time taken: 16.737 seconds, Fetched: 1 row(s)
hive (testdb)>

**hive (testdb)> SELECT avg(t1.d1) from parquet_test_no_compression t1 join text_test t2
on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;**
WARNING: 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.
Query ID = hduser_20180514191418_11c71842-9010-49a9-aba3-00ff392eafc2
Total jobs = 3
Stage-6 is selected by condition resolver.
Stage-1 is filtered out by condition resolver.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/usr/local/hadoop/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class
]
SLF4J: Found binding in
[jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLo
ggerBinder.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]
2018-05-14 19:14:22   Starting to launch local task to process map join;   maximum memory =
477626368
2018-05-14 19:14:24   Dump the side-table for tag: 1 with group count: 45031 into file:
file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-14-18_414_3843
146482035812019-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile141--.hashtable
2018-05-14 19:14:24   Uploaded 1 File to:
file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-14-18_414_3843
146482035812019-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile141--.hashtable
(1136439 bytes)
2018-05-14 19:14:24   End of local task; Time Taken: 1.619 sec.

Execution completed successfully

MapredLocal task succeeded

Launching Job 2 out of 3

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

Job running in-process (local Hadoop)

2018-05-14 19:14:26,628 Stage-4 map = 100%, reduce = 0%

Ended Job = job_local726055397_0020

Launching Job 3 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>

Job running in-process (local Hadoop)

2018-05-14 19:14:27,933 Stage-2 map = 100%, reduce = 100%

Ended Job = job_local419620656_0021

MapReduce Jobs Launched:

Stage-Stage-4: HDFS Read: 152100025 HDFS Write: 56837410 SUCCESS

Stage-Stage-2: HDFS Read: 304200050 HDFS Write: 113674820 SUCCESS

Total MapReduce CPU Time Spent: 0 msec

OK

_c0

3.348246

Time taken: 9.521 seconds, Fetched: 1 row(s)

hive (testdb)>

hive (testdb)> SELECT avg(t1.d1) from parquet_test_no_compression t1 join avro_test t2 on t1.d3 = t2.d3 and t1.d4 = t2.d4 where t1.d5<3;

WARNING: 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.

Query ID = hduser_20180514191505_03a85a72-5b96-40a6-b54c-1306fad14ce5

Total jobs = 3

Stage-6 is selected by condition resolver.

Stage-1 is filtered out by condition resolver.

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in

[jar:file:/usr/local/hadoop/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class

]

SLF4J: Found binding in
[jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.5.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]
2018-05-14 19:15:10   Starting to launch local task to process map join;   maximum memory = 477626368
2018-05-14 19:15:14   Dump the side-table for tag: 1 with group count: 45031 into file:
file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-15-05_892_2547
463534737559274-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile151--.hashtable
2018-05-14 19:15:15   Uploaded 1 File to:
file:/tmp/hduser/435e9e81-aff5-47a0-bd60-d780b99f1bf8/hive_2018-05-14_19-15-05_892_2547
463534737559274-1/-local-10005/HashTable-Stage-4/MapJoin-mapfile151--.hashtable
(1136439 bytes)
2018-05-14 19:15:15   End of local task; Time Taken: 4.759 sec.
Execution completed successfully
MapredLocal task succeeded
Launching Job 2 out of 3
Number of reduce tasks is set to 0 since there's no reduce operator
Job running in-process (local Hadoop)
2018-05-14 19:15:17,282 Stage-4 map = 100%, reduce = 0%
Ended Job = job_local1291659805_0022
Launching Job 3 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>
Job running in-process (local Hadoop)
2018-05-14 19:15:18,601 Stage-2 map = 100%, reduce = 100%
Ended Job = job_local112893897_0023
MapReduce Jobs Launched:
Stage-Stage-4: HDFS Read: 153110863 HDFS Write: 56837410 SUCCESS
Stage-Stage-2: HDFS Read: 306221726 HDFS Write: 113674820 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
_c0
3.348246
Time taken: 12.713 seconds, Fetched: 1 row(s)