```
┌──────────────────────────────────────────────────────────────────────┐
│                          • MobaXterm 10.4 •                          │
│                (SSH client, X-server and networking tools)           │
│                                                                      │
│ → SSH session to acadgild@192.168.56.2                               │
│   • SSH compression : v                                              │
│   • SSH-browser     : v                                              │
│   • X11-forwarding  : v  (remote display is forwarded through SSH)   │
│   • DISPLAY         : v  (automatically set on remote server)        │
│                                                                      │
│ → For more info, ctrl+click on help or visit our website             │
└──────────────────────────────────────────────────────────────────────┘

Last login: Sat Jul 14 08:55:14 2018 from 192.168.56.1
[acadgild.mmisra ~]$ vim hadoop_help.txt
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$ vim hadoop_help.txt
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$ cd
[acadgild.mmisra ~]$ pig Wordcount.pig
18/07/14 13:03:37 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/07/14 13:03:37 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
18/07/14 13:03:37 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType
2018-07-14 13:03:37,894 [main] INFO  org.apache.pig.Main - Apache Pig version 0.16.0
(r1746530) compiled Jun 01 2016, 23:10:49
2018-07-14 13:03:37,894 [main] INFO  org.apache.pig.Main - Logging error messages to:
/home/acadgild/pig_1531553617892.log
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log
4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/sl
f4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2018-07-14 13:03:38,391 [main] WARN  org.apache.hadoop.util.NativeCodeLoader - Unable to
load native-hadoop library for your platform... using builtin-java classes where applicable
2018-07-14 13:03:38,599 [main] INFO  org.apache.pig.impl.util.Utils - Default bootup file
/home/acadgild/.pigbootup not found
2018-07-14 13:03:38,791 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2018-07-14 13:03:38,791 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 13:03:38,792 [main] INFO
org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop
file system at: hdfs://localhost:8020
2018-07-14 13:03:39,415 [main] INFO  org.apache.pig.PigServer - Pig Script ID for the
session: PIG-Wordcount.pig-5379e963-697e-4285-86f4-d2f1e2c5fecf
2018-07-14 13:03:39,415 [main] WARN  org.apache.pig.PigServer - ATS is disabled since
yarn.timeline-service.enabled set to false
2018-07-14 13:03:40,175 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 13:03:40,282 [main] INFO  org.apache.pig.impl.util.SpillableMemoryManager -
Selected heap (PS Old Gen) of size 699400192 to monitor. collectionUsageThreshold =
```

```
489580128, usageThreshold = 489580128
2018-07-14 13:03:40,445 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 13:03:40,513 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 13:03:40,521 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 13:03:40,592 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 13:03:40,817 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 13:03:40,847 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 13:03:40,922 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 13:03:40,922 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 13:03:40,971 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 13:03:41,013 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:03:41,331 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 13:03:41,338 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 13:03:41,338 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 13:03:41,338 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.output.compress is deprecated. Instead, use
mapreduce.output.fileoutputformat.compress
2018-07-14 13:03:41,341 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 13:03:41,343 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 13:03:41,352 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=336
2018-07-14 13:03:41,352 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 13:03:41,352 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2018-07-14 13:03:41,353 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - This
job cannot be converted run in-process
2018-07-14 13:03:41,371 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.submit.replication is deprecated. Instead, use
mapreduce.client.submit.file.replication
2018-07-14 13:03:41,815 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added
jar file:/home/acadgild/install/pig/pig-0.16.0/pig-0.16.0-core-h2.jar to
DistributedCache through /tmp/temp937933567/tmp-1643576466/pig-0.16.0-core-h2.jar
2018-07-14 13:03:41,842 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added
jar file:/home/acadgild/install/pig/pig-0.16.0/lib/automaton-1.11-8.jar to
DistributedCache through /tmp/temp937933567/tmp-1160672871/automaton-1.11-8.jar
2018-07-14 13:03:41,878 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added
jar file:/home/acadgild/install/pig/pig-0.16.0/lib/antlr-runtime-3.4.jar to
DistributedCache through /tmp/temp937933567/tmp1808281738/antlr-runtime-3.4.jar
2018-07-14 13:03:41,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Added
jar file:/home/acadgild/install/pig/pig-0.16.0/lib/joda-time-2.9.3.jar to
DistributedCache through /tmp/temp937933567/tmp1932100799/joda-time-2.9.3.jar
2018-07-14 13:03:41,968 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 13:03:41,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 13:03:41,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 13:03:41,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Setting
key [pig.schematuple.classes] with classes to deserialize []
2018-07-14 13:03:42,152 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 13:03:42,153 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker.http.address is deprecated. Instead, use
mapreduce.jobtracker.http.address
2018-07-14 13:03:42,166 [JobControl] INFO  org.apache.hadoop.yarn.client.RMProxy -
Connecting to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:03:42,209 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.task.id is deprecated. Instead,
use mapreduce.task.attempt.id
2018-07-14 13:03:42,211 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 13:03:42,765 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 13:03:42,818 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 13:03:42,826 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 13:03:42,826 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 13:03:42,860 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 13:03:42,936 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 13:03:43,104 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_1531539606285_0007
2018-07-14 13:03:43,298 [JobControl] INFO  org.apache.hadoop.mapred.YARNRunner - Job jar
is not present. Not adding any jar to the list of resources.
2018-07-14 13:03:43,440 [JobControl] INFO
org.apache.hadoop.yarn.client.api.impl.YarnClientImpl - Submitted application
application_1531539606285_0007
2018-07-14 13:03:43,530 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8088/proxy/application_1531539606285_0007/
2018-07-14 13:03:43,530 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_1531539606285_0007
2018-07-14 13:03:43,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases A,B,C,D
2018-07-14 13:03:43,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: A[1,4],B[3,4],D[5,4],C[4,4] C: D[5,4],C[4,4] R: D[5,4]
2018-07-14 13:03:43,562 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 0%
complete
2018-07-14 13:03:43,563 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_1531539606285_0007]
2018-07-14 13:03:57,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 13:03:57,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_1531539606285_0007]
2018-07-14 13:04:02,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_1531539606285_0007]
2018-07-14 13:04:03,691 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:04:03,703 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate -
Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job
history server
2018-07-14 13:04:04,080 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:04:04,101 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate -
Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job
history server
2018-07-14 13:04:04,203 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:04:04,211 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate -
Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job
history server
2018-07-14 13:04:04,290 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 13:04:04,296 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 13:03:41     2018-07-14 13:04:04
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_1531539606285_0007  1       1       4       4       4       4       3       3
3       3       A,B,C,D GROUP_BY,COMBINER
hdfs://localhost:8020/tmp/temp937933567/tmp638482957,

Input(s):
Successfully read 21 records (699 bytes) from: "/files/test.txt"

Output(s):
Successfully stored 4 records (48 bytes) in:
```

```
"hdfs://localhost:8020/tmp/temp937933567/tmp638482957"

Counters:
Total records written : 4
Total bytes written : 48
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_1531539606285_0007


2018-07-14 13:04:04,298 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:04:04,302 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate -
Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job
history server
2018-07-14 13:04:04,356 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:04:04,362 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate -
Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job
history server
2018-07-14 13:04:04,402 [main] INFO  org.apache.hadoop.yarn.client.RMProxy - Connecting
to ResourceManager at localhost/127.0.0.1:8032
2018-07-14 13:04:04,413 [main] INFO  org.apache.hadoop.mapred.ClientServiceDelegate -
Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job
history server
2018-07-14 13:04:04,470 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 13:04:04,474 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 13:04:04,477 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 13:04:04,497 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 13:04:04,497 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(a,21)
(day,21)
(have,21)
(nice,21)
2018-07-14 13:04:04,603 [main] INFO  org.apache.pig.Main - Pig script completed in 27
seconds and 15 milliseconds (27015 ms)
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$ cd
[acadgild.mmisra ~]$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 8
Server version: 8.0.3-rc-log MySQL Community Server (GPL)

Copyright (c) 2000, 2017, Oracle and/or its affiliates. All rights reserved.
```

```
Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> show databases;
+--------------------+
| Database           |
+--------------------+
| information_schema |
| metastore          |
| mysql              |
| oozie              |
| performance_schema |
| simplidb           |
| sys                |
| testdb             |
+--------------------+
8 rows in set (0.01 sec)

mysql> ^C
mysql> ^C
mysql>
mysql>
mysql>
mysql>
mysql>
mysql>
mysql> quit;
Bye
You have new mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$ pig -x local
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log
4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/sl
f4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/07/14 18:44:28 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/07/14 18:44:28 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
2018-07-14 18:44:28,462 [main] INFO  org.apache.pig.Main - Apache Pig version 0.16.0
(r1746530) compiled Jun 01 2016, 23:10:49
2018-07-14 18:44:28,463 [main] INFO  org.apache.pig.Main - Logging error messages to:
/home/acadgild/pig_1531574068461.log
2018-07-14 18:44:28,484 [main] INFO  org.apache.pig.impl.util.Utils - Default bootup file
/home/acadgild/.pigbootup not found
2018-07-14 18:44:28,700 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2018-07-14 18:44:28,700 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:44:28,703 [main] INFO
org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop
```

```
file system at: file:///
2018-07-14 18:44:28,785 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:28,806 [main] INFO  org.apache.pig.PigServer - Pig Script ID for the
session: PIG-default-00d2584d-fb8c-4899-9fb7-c0b74b840e2b
2018-07-14 18:44:28,806 [main] WARN  org.apache.pig.PigServer - ATS is disabled since
yarn.timeline-service.enabled set to false
grunt>
grunt>
grunt>
grunt>
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int);
2018-07-14 18:44:49,177 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:49,178 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
(EmpId:Int,Expense:Int);
2018-07-14 18:44:49,318 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:49,318 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt>
grunt>
grunt> dump expenses;
2018-07-14 18:44:54,238 [main] WARN  org.apache.hadoop.util.NativeCodeLoader - Unable to
load native-hadoop library for your platform... using builtin-java classes where applicable
2018-07-14 18:44:54,289 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:44:54,336 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:54,336 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:44:54,389 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:44:54,477 [main] INFO  org.apache.pig.impl.util.SpillableMemoryManager -
Selected heap (PS Old Gen) of size 699400192 to monitor. collectionUsageThreshold =
489580128, usageThreshold = 489580128
2018-07-14 18:44:54,557 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:44:54,596 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:44:54,596 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:44:54,634 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:54,634 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:44:54,651 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
session.id is deprecated. Instead, use dfs.metrics.session-id
2018-07-14 18:44:54,653 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Initializing JVM Metrics with processName=JobTracker, sessionId=
2018-07-14 18:44:54,692 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
```

```
the job
2018-07-14 18:44:54,700 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:44:54,700 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:44:54,700 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.output.compress is deprecated. Instead, use
mapreduce.output.fileoutputformat.compress
2018-07-14 18:44:54,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:44:54,734 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:44:54,734 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:44:54,734 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574094733-0
2018-07-14 18:44:54,788 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:44:54,789 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker.http.address is deprecated. Instead, use
mapreduce.jobtracker.http.address
2018-07-14 18:44:54,810 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:44:54,848 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.task.id is deprecated. Instead,
use mapreduce.task.attempt.id
2018-07-14 18:44:55,255 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:44:55,290 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:44:55,297 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:44:55,297 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:44:55,319 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:44:55,421 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:44:55,590 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1902519347_0001
2018-07-14 18:44:55,780 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:44:55,781 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1902519347_0001
2018-07-14 18:44:55,781 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses
2018-07-14 18:44:55,781 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1] C:  R:
2018-07-14 18:44:55,789 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 0%
complete
2018-07-14 18:44:55,790 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1902519347_0001]
2018-07-14 18:44:55,790 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:44:55,808 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:44:55,808 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:44:55,809 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:55,809 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:44:55,812 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:44:55,887 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:44:55,889 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1902519347_0001_m_000000_0
2018-07-14 18:44:55,968 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:44:55,976 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:44:55,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:44:55,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 18:44:56,037 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:44:56,040 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:44:56,057 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: expenses[2,11],expenses[-1,-1]
C:  R:
2018-07-14 18:44:56,073 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:44:56,074 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1902519347_0001_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:44:56,091 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:44:56,091 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1902519347_0001_m_000000_0 is allowed
to commit now
2018-07-14 18:44:56,095 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1902519347_0001_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1892639912/_temporary/0/task_local1902519347_0001_m_0000
00
2018-07-14 18:44:56,099 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:44:56,099 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1902519347_0001_m_000000_0' done.
2018-07-14 18:44:56,099 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1902519347_0001_m_000000_0
2018-07-14 18:44:56,100 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:44:56,288 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:44:56,292 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:44:56,292 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2018-07-14 18:44:56,292 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2018-07-14 18:44:56,294 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:44:56,342 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:44:56,346 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:44:54     2018-07-14 18:44:56     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1902519347_0001        1       0       n/a     n/a     n/a     n/a     0
0       0       0       expenses        MAP_ONLY
file:/tmp/temp-1902294536/tmp-1892639912,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 9 records in: "file:/tmp/temp-1902294536/tmp-1892639912"

Counters:
Total records written : 9
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1902519347_0001


2018-07-14 18:44:56,348 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:44:56,350 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 18:44:56,351 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:44:56,358 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:44:56,364 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:44:56,365 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:44:56,365 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:44:56,378 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:44:56,378 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,200)
(102,100)
(110,400)
(114,200)
(119,200)
(105,100)
(101,100)
(104,300)
(102,400)
grunt>
grunt>
grunt>
grunt> describe expenses;
expenses: {EmpId: int,Expense: int}
grunt> describe details;;
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt> s= foreach details generate details.Name;
grunt> describe s;
s: {Name: chararray}
grunt> dump s;
2018-07-14 18:46:12,581 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:46:12,618 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:46:12,619 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:46:12,619 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:46:12,620 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:46:12,630 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:46:12,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 18:46:12,633 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 18:46:12,633 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 18:46:12,633 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 18:46:12,649 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:46:12,649 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:46:12,650 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:12,651 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:46:12,652 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:46:12,658 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 18:46:12,660 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:46:12,660 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:46:12,660 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574172659-0
2018-07-14 18:46:12,681 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:46:12,684 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:12,708 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:46:12,712 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:46:12,713 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:46:12,713 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:46:12,714 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:46:12,756 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:46:12,791 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local457537790_0002
2018-07-14 18:46:12,920 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:46:12,924 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:46:12,934 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:46:12,934 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:46:12,935 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:46:12,935 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 18:46:12,937 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:46:12,945 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:46:12,945 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local457537790_0002_m_000000_0
2018-07-14 18:46:12,962 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:46:12,966 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:46:12,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:46:12,976 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:46:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:46:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:46:12,992 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[1,10],details[-1,-1]
C:  R:
2018-07-14 18:46:13,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:46:13,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local457537790_0002_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:46:13,027 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:46:13,027 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local457537790_0002_m_000000_0 is allowed to
commit now
2018-07-14 18:46:13,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local457537790_0002_m_000000_0' to
file:/tmp/temp-1902294536/tmp-912642892/_temporary/0/task_local457537790_0002_m_000000
2018-07-14 18:46:13,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local457537790_0002_m_000000_0' to
file:/tmp/temp-1902294536/tmp-2131063769/_temporary/0/task_local457537790_0002_m_00000
0
2018-07-14 18:46:13,038 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:46:13,038 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local457537790_0002_m_000000_0' done.
2018-07-14 18:46:13,038 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local457537790_0002_m_000000_0
2018-07-14 18:46:13,038 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
```

```
2018-07-14 18:46:13,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local457537790_0002
2018-07-14 18:46:13,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 18:46:13,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[1,10],details[-1,-1] C:  R:
2018-07-14 18:46:13,184 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 18:46:13,185 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,187 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,188 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,191 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:46:13,192 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:46:13,193 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:46:13,205 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:46:13,207 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,209 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:46:13,209 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:46:13,219 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:46:13,236 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:46:13,236 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:46:13,236 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:46:13,265 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:46:13,315 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local993460676_0003
2018-07-14 18:46:13,422 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:46:13,423 [Thread-74] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:46:13,427 [Thread-74] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:46:13,427 [Thread-74] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
```

```
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:46:13,427 [Thread-74] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:46:13,428 [Thread-74] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:46:13,428 [Thread-74] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:46:13,432 [Thread-74] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:46:13,432 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local993460676_0003_m_000000_0
2018-07-14 18:46:13,441 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:46:13,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 280
Input split[0]:
   Length = 280
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:46:13,447 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-2131063769/part-m-00000:0+280
2018-07-14 18:46:13,458 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:46:13,459 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 18:46:13,461 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: s[3,3] C:  R:
2018-07-14 18:46:13,462 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 18:46:13,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:46:13,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:46:13,477 [Thread-74] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:46:13,480 [Thread-74] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local993460676_0003
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,Amitabh,20000,1), 2nd
:(102,Shahrukh,10000,2) (common cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should
be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:522)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,Amitabh,20000,1), 2nd :(102,Shahrukh,10000,2) (common
cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
```

```
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextString(POUserFunc.java:426)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:341)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.processPlan(POForEach.java:404)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:321)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.runPipe
line(PigGenericMapBase.java:280)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:275)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:65)
        at org.apache.hadoop.mapreduce.Mapper.run(Mapper.java:145)
        at org.apache.hadoop.mapred.MapTask.runNewMapper(MapTask.java:784)
        at org.apache.hadoop.mapred.MapTask.run(MapTask.java:341)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$MapTaskRunnable.run(LocalJobRunner.java:24
3)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
2018-07-14 18:46:13,623 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local993460676_0003
2018-07-14 18:46:13,623 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases s
2018-07-14 18:46:13,623 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: s[3,3] C:  R:
2018-07-14 18:46:13,626 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 18:46:13,626 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local993460676_0003 has failed! Stop running all dependent jobs
2018-07-14 18:46:13,626 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:46:13,628 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,628 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,629 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 18:46:13,629 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:46:12     2018-07-14 18:46:13     UNKNOWN
```

```
Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local457537790_0002 1       0       n/a     n/a     n/a     n/a     0       0
0       0       details MULTI_QUERY,MAP_ONLY

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local993460676_0003 s       MAP_ONLY        Message: Job failed!
file:/tmp/temp-1902294536/tmp-1647182853,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp-1647182853"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local457537790_0002 ->      job_local993460676_0003,
job_local993460676_0003


2018-07-14 18:46:13,631 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,632 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,632 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:46:13,641 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 18:46:13,649 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias s
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> 2018-07-14 18:46:19,448 [communication thread] INFO
org.apache.hadoop.mapred.LocalJobRunner - map > map

grunt> describe details;
2018-07-14 18:46:28,532 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:46:28,533 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(t:tuple(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int));
2018-07-14 18:47:56,821 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:47:56,821 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt> dump details;
2018-07-14 18:48:01,891 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
```

```
features used in the script: UNKNOWN
2018-07-14 18:48:01,945 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:01,945 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:48:01,945 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:48:01,946 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:48:01,949 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:48:01,950 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:48:01,950 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:48:01,966 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:01,967 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:48:01,967 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:01,969 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:48:01,969 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:48:01,978 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:48:01,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:48:01,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:48:01,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574281979-0
2018-07-14 18:48:02,009 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:48:02,020 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,050 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:48:02,059 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:48:02,060 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:48:02,060 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:48:02,064 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:48:02,105 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
```

```
number of splits:1
2018-07-14 18:48:02,147 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local56732004_0004
2018-07-14 18:48:02,336 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:48:02,337 [Thread-95] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:48:02,345 [Thread-95] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:48:02,346 [Thread-95] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:48:02,346 [Thread-95] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:02,346 [Thread-95] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:48:02,346 [Thread-95] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:48:02,354 [Thread-95] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:48:02,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local56732004_0004_m_000000_0
2018-07-14 18:48:02,375 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:48:02,380 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:48:02,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:48:02,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:48:02,417 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:48:02,418 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:48:02,424 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[4,10],details[-1,-1]
C:  R:
2018-07-14 18:48:02,426 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 49] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,431 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
```

```
Unable to interpret value [49, 48, 50] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,431 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 51] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,431 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 52] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,432 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 53] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,432 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 54] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,433 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 55] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,433 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 56] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,433 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 48, 57] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,434 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 49, 48] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,434 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 49, 49] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,434 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 49, 50] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,435 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 49, 51] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,435 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigHadoopLogger -
org.apache.pig.builtin.Utf8StorageConverter(FIELD_DISCARDED_TYPE_CONVERSION_FAILED):
Unable to interpret value [49, 49, 52] in field being converted to type tuple, caught
ParseException <Unexpect end of tuple> field discarded
2018-07-14 18:48:02,437 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:48:02,437 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local56732004_0004_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:48:02,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:48:02,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local56732004_0004_m_000000_0 is allowed to
commit now
2018-07-14 18:48:02,451 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local56732004_0004_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1666671358/_temporary/0/task_local56732004_0004_m_000000
2018-07-14 18:48:02,453 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:48:02,453 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local56732004_0004_m_000000_0' done.
2018-07-14 18:48:02,453 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local56732004_0004_m_000000_0
2018-07-14 18:48:02,453 [Thread-95] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:48:02,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local56732004_0004
2018-07-14 18:48:02,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 18:48:02,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[4,10],details[-1,-1] C:  R:
2018-07-14 18:48:02,519 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 18:48:02,519 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local56732004_0004]
2018-07-14 18:48:02,640 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,641 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,641 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,650 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:48:02,651 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:48:01     2018-07-14 18:48:02     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local56732004_0004  1       0       n/a     n/a     n/a     n/a     0       0
0       0       details MAP_ONLY        file:/tmp/temp-1902294536/tmp-1666671358,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"
```

```
Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp-1666671358"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local56732004_0004


2018-07-14 18:48:02,652 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,654 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,655 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:02,663 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Encountered Warning FIELD_DISCARDED_TYPE_CONVERSION_FAILED 14 time(s).
2018-07-14 18:48:02,663 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:48:02,663 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:02,664 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:48:02,664 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:48:02,674 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:48:02,674 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
()
()
()
()
()
()
()
()
()
()
()
()
()
()
grunt> details = load 'employee_details.txt' using PigStorage(',');
2018-07-14 18:48:50,627 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:50,627 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> dump details;
2018-07-14 18:48:55,119 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:48:55,140 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:55,140 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:48:55,140 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:48:55,142 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:48:55,143 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:48:55,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:48:55,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:48:55,158 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:55,159 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:48:55,159 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,160 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:48:55,160 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:48:55,162 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:48:55,166 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:48:55,168 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,188 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:48:55,195 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:48:55,196 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:48:55,196 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:48:55,196 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:48:55,221 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:48:55,239 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1168975028_0005
2018-07-14 18:48:55,353 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:48:55,354 [Thread-118] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:48:55,359 [Thread-118] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:48:55,359 [Thread-118] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:48:55,359 [Thread-118] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:55,359 [Thread-118] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:48:55,359 [Thread-118] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:48:55,364 [Thread-118] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:48:55,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1168975028_0005_m_000000_0
2018-07-14 18:48:55,375 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:48:55,384 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:48:55,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:48:55,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:48:55,413 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:48:55,415 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 18:48:55,417 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[5,10] C:  R:
2018-07-14 18:48:55,422 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:48:55,422 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1168975028_0005_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:48:55,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:48:55,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1168975028_0005_m_000000_0 is allowed
to commit now
2018-07-14 18:48:55,427 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1168975028_0005_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1934904221/_temporary/0/task_local1168975028_0005_m_0000
00
2018-07-14 18:48:55,431 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:48:55,432 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1168975028_0005_m_000000_0' done.
2018-07-14 18:48:55,432 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
```

```
attempt_local1168975028_0005_m_000000_0
2018-07-14 18:48:55,432 [Thread-118] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:48:55,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1168975028_0005
2018-07-14 18:48:55,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 18:48:55,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[5,10] C:  R:
2018-07-14 18:48:55,582 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,587 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,590 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,614 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:48:55,615 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:48:55     2018-07-14 18:48:55     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1168975028_0005        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details MAP_ONLY
file:/tmp/temp-1902294536/tmp-1934904221,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp-1934904221"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1168975028_0005


2018-07-14 18:48:55,619 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,621 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,622 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:48:55,628 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
```

```
2018-07-14 18:48:55,628 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:48:55,629 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:48:55,629 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:48:55,642 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:48:55,642 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1)
(102,Shahrukh,10000,2)
(103,Akshay,11000,3)
(104,Anubhav,5000,4)
(105,Pawan,2500,5)
(106,Aamir,25000,1)
(107,Salman,17500,2)
(108,Ranbir,14000,3)
(109,Katrina,1000,4)
(110,Priyanka,2000,5)
(111,Tushar,500,1)
(112,Ajay,5000,2)
(113,Jubeen,1000,1)
(114,Madhuri,2000,2)
grunt> details = load 'employee_details.txt' using PigStorage(',') as
EmpId:Int,Name:CharArray,Salary:Int,Dept:Int;
2018-07-14 18:49:24,287 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
6, column 72>  mismatched input ',' expecting SEMI_COLON
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int);
2018-07-14 18:49:33,479 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:49:33,479 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> dump details;
2018-07-14 18:49:38,636 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:49:38,659 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:49:38,660 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:49:38,660 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:49:38,660 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:49:38,664 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:49:38,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:49:38,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:49:38,676 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:49:38,676 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:49:38,677 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:38,678 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:49:38,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:49:38,681 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:49:38,681 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:49:38,681 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:49:38,681 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574378681-0
2018-07-14 18:49:38,686 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:49:38,688 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:38,704 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:49:38,709 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:49:38,710 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:49:38,711 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:49:38,711 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:49:38,741 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:49:38,764 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local777440431_0006
2018-07-14 18:49:38,869 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:49:38,870 [Thread-141] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:49:38,875 [Thread-141] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:49:38,876 [Thread-141] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:49:38,876 [Thread-141] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:49:38,876 [Thread-141] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:49:38,876 [Thread-141] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:49:38,884 [Thread-141] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
```

```
2018-07-14 18:49:38,885 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local777440431_0006_m_000000_0
2018-07-14 18:49:38,894 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:49:38,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:49:38,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:49:38,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:49:38,918 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:49:38,918 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:49:38,922 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[6,10],details[-1,-1]
C:  R:
2018-07-14 18:49:38,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:49:38,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local777440431_0006_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:49:38,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:49:38,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local777440431_0006_m_000000_0 is allowed to
commit now
2018-07-14 18:49:38,940 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local777440431_0006_m_000000_0' to
file:/tmp/temp-1902294536/tmp1227851900/_temporary/0/task_local777440431_0006_m_000000
2018-07-14 18:49:38,941 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:49:38,941 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local777440431_0006_m_000000_0' done.
2018-07-14 18:49:38,941 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local777440431_0006_m_000000_0
2018-07-14 18:49:38,941 [Thread-141] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:49:39,075 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local777440431_0006
2018-07-14 18:49:39,075 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 18:49:39,075 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[6,10],details[-1,-1] C:  R:
2018-07-14 18:49:39,077 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:39,077 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:39,077 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:39,080 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:49:39,080 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:49:38     2018-07-14 18:49:39     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local777440431_0006 1       0       n/a     n/a     n/a     n/a     0       0
0       0       details MAP_ONLY        file:/tmp/temp-1902294536/tmp1227851900,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp1227851900"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local777440431_0006


2018-07-14 18:49:39,081 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:39,082 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:39,082 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:49:39,084 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:49:39,085 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:49:39,085 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:49:39,085 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:49:39,093 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:49:39,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1)
(102,Shahrukh,10000,2)
(103,Akshay,11000,3)
```

```
(104,Anubhav,5000,4)
(105,Pawan,2500,5)
(106,Aamir,25000,1)
(107,Salman,17500,2)
(108,Ranbir,14000,3)
(109,Katrina,1000,4)
(110,Priyanka,2000,5)
(111,Tushar,500,1)
(112,Ajay,5000,2)
(113,Jubeen,1000,1)
(114,Madhuri,2000,2)
grunt> describe details;
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt> details = load 'employee_details.txt' using PigStorage(',');
2018-07-14 18:49:57,728 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:49:57,728 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> describe details;
Schema for details unknown.
grunt> dump details;
2018-07-14 18:50:08,960 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:50:08,981 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:50:08,981 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:50:08,981 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:50:08,981 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:50:08,982 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:50:08,983 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:50:08,983 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:50:08,998 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:50:08,998 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:50:08,999 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,000 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:50:09,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:50:09,002 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:50:09,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:50:09,008 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
```

```
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,019 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:50:09,027 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:50:09,028 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:50:09,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:50:09,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:50:09,054 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:50:09,074 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1992327078_0007
2018-07-14 18:50:09,166 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:50:09,172 [Thread-164] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:50:09,180 [Thread-164] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:50:09,180 [Thread-164] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:50:09,180 [Thread-164] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:50:09,180 [Thread-164] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:50:09,180 [Thread-164] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:50:09,185 [Thread-164] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:50:09,186 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1992327078_0007_m_000000_0
2018-07-14 18:50:09,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:50:09,195 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:50:09,203 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:50:09,204 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:50:09,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 18:50:09,219 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 18:50:09,221 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[7,10] C:  R:
2018-07-14 18:50:09,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:50:09,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1992327078_0007_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:50:09,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:50:09,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1992327078_0007_m_000000_0 is allowed
to commit now
2018-07-14 18:50:09,226 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1992327078_0007_m_000000_0' to
file:/tmp/temp-1902294536/tmp248748980/_temporary/0/task_local1992327078_0007_m_000000
2018-07-14 18:50:09,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:50:09,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1992327078_0007_m_000000_0' done.
2018-07-14 18:50:09,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1992327078_0007_m_000000_0
2018-07-14 18:50:09,228 [Thread-164] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:50:09,368 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1992327078_0007
2018-07-14 18:50:09,368 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 18:50:09,369 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[7,10] C:  R:
2018-07-14 18:50:09,386 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,389 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,391 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,396 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:50:09,397 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:50:09     2018-07-14 18:50:09     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1992327078_0007        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details MAP_ONLY        file:/tmp/temp-1902294536/tmp248748980,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"
```

```
Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp248748980"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1992327078_0007


2018-07-14 18:50:09,404 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,406 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,410 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:50:09,412 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:50:09,413 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:50:09,413 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:50:09,413 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:50:09,424 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:50:09,424 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1)
(102,Shahrukh,10000,2)
(103,Akshay,11000,3)
(104,Anubhav,5000,4)
(105,Pawan,2500,5)
(106,Aamir,25000,1)
(107,Salman,17500,2)
(108,Ranbir,14000,3)
(109,Katrina,1000,4)
(110,Priyanka,2000,5)
(111,Tushar,500,1)
(112,Ajay,5000,2)
(113,Jubeen,1000,1)
(114,Madhuri,2000,2)
grunt> describe details;
Schema for details unknown.
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int);
2018-07-14 18:50:19,094 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:50:19,094 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt> s= foreach details generate Salary+1;
grunt> dump s;
2018-07-14 18:52:05,193 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:52:05,218 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:52:05,218 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:52:05,218 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:52:05,219 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:52:05,229 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $0,
$1, $3
2018-07-14 18:52:05,234 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:52:05,235 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:52:05,235 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:52:05,247 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:52:05,248 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:52:05,248 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,249 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:52:05,250 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:52:05,251 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:52:05,258 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:52:05,258 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:52:05,258 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574525258-0
2018-07-14 18:52:05,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:52:05,275 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,290 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:52:05,296 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:52:05,297 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:52:05,297 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:52:05,297 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
```

```
2018-07-14 18:52:05,333 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:52:05,357 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local161381131_0008
2018-07-14 18:52:05,447 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:52:05,447 [Thread-187] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:52:05,452 [Thread-187] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:52:05,453 [Thread-187] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:52:05,453 [Thread-187] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:52:05,453 [Thread-187] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:52:05,453 [Thread-187] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:52:05,456 [Thread-187] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:52:05,456 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local161381131_0008_m_000000_0
2018-07-14 18:52:05,461 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:52:05,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:52:05,464 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:52:05,464 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:52:05,472 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:52:05,472 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:52:05,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],s[-1,-1] C:  R:
2018-07-14 18:52:05,476 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:52:05,476 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local161381131_0008_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:52:05,478 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:52:05,478 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task attempt_local161381131_0008_m_000000_0 is allowed to
commit now
2018-07-14 18:52:05,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local161381131_0008_m_000000_0' to
file:/tmp/temp-1902294536/tmp1753743572/_temporary/0/task_local161381131_0008_m_000000
2018-07-14 18:52:05,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:52:05,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local161381131_0008_m_000000_0' done.
2018-07-14 18:52:05,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local161381131_0008_m_000000_0
2018-07-14 18:52:05,481 [Thread-187] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:52:05,650 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local161381131_0008
2018-07-14 18:52:05,650 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details,s
2018-07-14 18:52:05,650 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],s[-1,-1] C:  R:
2018-07-14 18:52:05,664 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,666 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,667 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:52:05,674 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:52:05     2018-07-14 18:52:05     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local161381131_0008 1       0       n/a     n/a     n/a     n/a     0       0
0       0       details,s       MAP_ONLY
file:/tmp/temp-1902294536/tmp1753743572,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp1753743572"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
```

```
job_local161381131_0008


2018-07-14 18:52:05,676 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,677 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,679 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:52:05,681 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:52:05,682 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:52:05,682 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:52:05,683 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:52:05,692 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:52:05,692 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(20001)
(10001)
(11001)
(5001)
(2501)
(25001)
(17501)
(14001)
(1001)
(2001)
(501)
(5001)
(1001)
(2001)
grunt> describe s;
s: {int}
grunt> describe details;
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt>
grunt>
grunt>
grunt> s= foreach details generate EmpId;
grunt> dump s;
2018-07-14 18:53:40,462 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:53:40,484 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:53:40,484 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:53:40,484 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:53:40,484 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:53:40,485 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $1,
$2, $3
```

```
2018-07-14 18:53:40,486 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:53:40,487 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:53:40,487 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:53:40,499 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:53:40,499 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:53:40,500 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,501 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:53:40,501 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:53:40,502 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:53:40,503 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:53:40,503 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:53:40,503 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574620503-0
2018-07-14 18:53:40,508 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:53:40,512 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,530 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:53:40,534 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:53:40,534 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:53:40,534 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:53:40,535 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:53:40,568 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:53:40,586 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local990758863_0009
2018-07-14 18:53:40,675 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:53:40,676 [Thread-210] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:53:40,679 [Thread-210] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:53:40,680 [Thread-210] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
```

```
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:53:40,680 [Thread-210] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:53:40,680 [Thread-210] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:53:40,680 [Thread-210] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:53:40,683 [Thread-210] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:53:40,683 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local990758863_0009_m_000000_0
2018-07-14 18:53:40,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:53:40,690 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:53:40,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:53:40,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:53:40,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:53:40,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:53:40,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],s[-1,-1] C:  R:
2018-07-14 18:53:40,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:53:40,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local990758863_0009_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:53:40,713 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:53:40,713 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local990758863_0009_m_000000_0 is allowed to
commit now
2018-07-14 18:53:40,714 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local990758863_0009_m_000000_0' to
file:/tmp/temp-1902294536/tmp1422510216/_temporary/0/task_local990758863_0009_m_000000
2018-07-14 18:53:40,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:53:40,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local990758863_0009_m_000000_0' done.
2018-07-14 18:53:40,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local990758863_0009_m_000000_0
2018-07-14 18:53:40,716 [Thread-210] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
```

```
task executor complete.
2018-07-14 18:53:40,877 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local990758863_0009
2018-07-14 18:53:40,877 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details,s
2018-07-14 18:53:40,877 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],s[-1,-1] C:  R:
2018-07-14 18:53:40,889 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,892 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,894 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,901 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:53:40,901 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:53:40     2018-07-14 18:53:40     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local990758863_0009 1       0       n/a     n/a     n/a     n/a     0       0
0       0       details,s       MAP_ONLY
file:/tmp/temp-1902294536/tmp1422510216,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp1422510216"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local990758863_0009


2018-07-14 18:53:40,903 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,904 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,905 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:53:40,908 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:53:40,909 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 18:53:40,909 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:53:40,909 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:53:40,922 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:53:40,922 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101)
(102)
(103)
(104)
(105)
(106)
(107)
(108)
(109)
(110)
(111)
(112)
(113)
(114)
grunt> describe details;
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt> s= foreach details generate EmpId+Dept;
grunt> dump s;
2018-07-14 18:54:07,104 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:54:07,127 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:54:07,127 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:54:07,127 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:54:07,128 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:54:07,131 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $1,
$2
2018-07-14 18:54:07,137 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:54:07,138 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:54:07,138 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:54:07,150 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:54:07,150 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:54:07,151 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,156 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:54:07,156 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:54:07,157 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:54:07,158 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:54:07,158 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:54:07,158 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574647157-0
2018-07-14 18:54:07,162 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:54:07,164 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,178 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:54:07,182 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:54:07,182 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:54:07,182 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:54:07,182 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:54:07,206 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:54:07,225 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1776839666_0010
2018-07-14 18:54:07,337 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:54:07,337 [Thread-233] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:54:07,343 [Thread-233] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:54:07,343 [Thread-233] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:54:07,343 [Thread-233] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:54:07,343 [Thread-233] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:54:07,343 [Thread-233] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:54:07,347 [Thread-233] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:54:07,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1776839666_0010_m_000000_0
2018-07-14 18:54:07,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:54:07,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
```

```
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:54:07,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:54:07,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:54:07,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:54:07,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:54:07,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],s[-1,-1] C:  R:
2018-07-14 18:54:07,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:54:07,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1776839666_0010_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:54:07,385 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:54:07,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1776839666_0010_m_000000_0 is allowed
to commit now
2018-07-14 18:54:07,387 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1776839666_0010_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1700168907/_temporary/0/task_local1776839666_0010_m_0000
00
2018-07-14 18:54:07,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:54:07,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1776839666_0010_m_000000_0' done.
2018-07-14 18:54:07,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1776839666_0010_m_000000_0
2018-07-14 18:54:07,388 [Thread-233] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:54:07,541 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1776839666_0010
2018-07-14 18:54:07,541 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details,s
2018-07-14 18:54:07,541 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],s[-1,-1] C:  R:
2018-07-14 18:54:07,552 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,556 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,560 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,575 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
```

```
complete
2018-07-14 18:54:07,577 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:54:07     2018-07-14 18:54:07     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1776839666_0010        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details,s       MAP_ONLY
file:/tmp/temp-1902294536/tmp-1700168907,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp-1700168907"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1776839666_0010


2018-07-14 18:54:07,579 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,580 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,581 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:54:07,586 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:54:07,586 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:54:07,587 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:54:07,587 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:54:07,606 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:54:07,606 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102)
(104)
(106)
(108)
(110)
(107)
(109)
(111)
(113)
```

```
(115)
(112)
(114)
(114)
(116)
grunt> describe details;
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt> dump details;
2018-07-14 18:55:48,506 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:55:48,532 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:55:48,532 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:55:48,532 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:55:48,532 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:55:48,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:55:48,535 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:55:48,535 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:55:48,546 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:55:48,546 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:55:48,546 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,547 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:55:48,548 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:55:48,548 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:55:48,549 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:55:48,549 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:55:48,549 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574748549-0
2018-07-14 18:55:48,553 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:55:48,555 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,566 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:55:48,572 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
```

```
2018-07-14 18:55:48,573 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:55:48,573 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:55:48,574 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:55:48,602 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:55:48,618 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local366579684_0011
2018-07-14 18:55:48,711 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:55:48,711 [Thread-256] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:55:48,715 [Thread-256] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:55:48,715 [Thread-256] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:55:48,715 [Thread-256] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:55:48,715 [Thread-256] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:55:48,716 [Thread-256] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:55:48,721 [Thread-256] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:55:48,721 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local366579684_0011_m_000000_0
2018-07-14 18:55:48,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:55:48,730 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:55:48,735 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:55:48,735 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 18:55:48,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:55:48,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:55:48,752 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],details[-1,-1]
```

```
C:  R:
2018-07-14 18:55:48,757 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:55:48,757 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local366579684_0011_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:55:48,759 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:55:48,759 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local366579684_0011_m_000000_0 is allowed to
commit now
2018-07-14 18:55:48,760 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local366579684_0011_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1319883061/_temporary/0/task_local366579684_0011_m_00000
0
2018-07-14 18:55:48,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:55:48,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local366579684_0011_m_000000_0' done.
2018-07-14 18:55:48,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local366579684_0011_m_000000_0
2018-07-14 18:55:48,761 [Thread-256] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:55:48,914 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local366579684_0011
2018-07-14 18:55:48,914 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 18:55:48,914 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1] C:  R:
2018-07-14 18:55:48,922 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,923 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,924 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,928 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:55:48,928 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:55:48     2018-07-14 18:55:48     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local366579684_0011 1       0       n/a     n/a     n/a     n/a     0       0
0       0       details MAP_ONLY        file:/tmp/temp-1902294536/tmp-1319883061,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp-1319883061"
```

```
Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local366579684_0011


2018-07-14 18:55:48,929 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,931 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,931 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:55:48,934 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:55:48,935 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:55:48,935 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:55:48,935 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:55:48,945 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:55:48,945 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1)
(102,Shahrukh,10000,2)
(103,Akshay,11000,3)
(104,Anubhav,5000,4)
(105,Pawan,2500,5)
(106,Aamir,25000,1)
(107,Salman,17500,2)
(108,Ranbir,14000,3)
(109,Katrina,1000,4)
(110,Priyanka,2000,5)
(111,Tushar,500,1)
(112,Ajay,5000,2)
(113,Jubeen,1000,1)
(114,Madhuri,2000,2)
grunt> dump expenses;
2018-07-14 18:56:16,377 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 18:56:16,394 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:16,395 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:56:16,395 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:56:16,395 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:56:16,396 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
```

```
concatenation threshold: 100 optimistic? false
2018-07-14 18:56:16,397 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:56:16,397 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:56:16,406 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:16,406 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:56:16,406 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,407 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:56:16,407 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:56:16,408 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:56:16,408 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:56:16,408 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:56:16,408 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574776408-0
2018-07-14 18:56:16,412 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:56:16,413 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,426 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:56:16,431 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:56:16,432 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:56:16,432 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:56:16,432 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:56:16,464 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:56:16,480 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local610159054_0012
2018-07-14 18:56:16,563 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:56:16,564 [Thread-279] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:56:16,568 [Thread-279] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:56:16,568 [Thread-279] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:56:16,568 [Thread-279] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:16,568 [Thread-279] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:56:16,568 [Thread-279] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:56:16,573 [Thread-279] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:56:16,573 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local610159054_0012_m_000000_0
2018-07-14 18:56:16,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:56:16,586 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:56:16,594 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:56:16,595 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 18:56:16,608 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:56:16,608 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:56:16,613 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: expenses[2,11],expenses[-1,-1]
C:  R:
2018-07-14 18:56:16,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:56:16,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local610159054_0012_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:56:16,615 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:56:16,618 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local610159054_0012_m_000000_0 is allowed to
commit now
2018-07-14 18:56:16,619 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local610159054_0012_m_000000_0' to
file:/tmp/temp-1902294536/tmp886971256/_temporary/0/task_local610159054_0012_m_000000
2018-07-14 18:56:16,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:56:16,624 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local610159054_0012_m_000000_0' done.
2018-07-14 18:56:16,624 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local610159054_0012_m_000000_0
2018-07-14 18:56:16,624 [Thread-279] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
```

```
2018-07-14 18:56:16,766 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local610159054_0012
2018-07-14 18:56:16,766 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses
2018-07-14 18:56:16,766 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1] C:  R:
2018-07-14 18:56:16,778 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,781 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,785 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,807 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:56:16,809 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:56:16     2018-07-14 18:56:16     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local610159054_0012 1       0       n/a     n/a     n/a     n/a     0       0
0       0       expenses        MAP_ONLY
file:/tmp/temp-1902294536/tmp886971256,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 9 records in: "file:/tmp/temp-1902294536/tmp886971256"

Counters:
Total records written : 9
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local610159054_0012


2018-07-14 18:56:16,811 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,812 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,814 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:16,825 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:56:16,825 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:16,827 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:56:16,827 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:56:16,842 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:56:16,842 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,200)
(102,100)
(110,400)
(114,200)
(119,200)
(105,100)
(101,100)
(104,300)
(102,400)
grunt> describe expenses;
expenses: {EmpId: int,Expense: int}
grunt> s=group expenses by EmpId;
2018-07-14 18:56:36,081 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1000: Error
during parsing. Encountered " <PATH> "s=group "" at line 26, column 1.
Was expecting one of:
    <EOF>
    "cat" ...
    "clear" ...
    "fs" ...
    "sh" ...
    "cd" ...
    "cp" ...
    "copyFromLocal" ...
    "copyToLocal" ...
    "dump" ...
    "\\d" ...
    "describe" ...
    "\\de" ...
    "aliases" ...
    "explain" ...
    "\\e" ...
    "help" ...
    "history" ...
    "kill" ...
    "ls" ...
    "mv" ...
    "mkdir" ...
    "pwd" ...
    "quit" ...
    "\\q" ...
    "register" ...
    "rm" ...
    "rmf" ...
    "set" ...
    "illustrate" ...
    "\\i" ...
    "run" ...
    "exec" ...
    "%default" ...
    "%declare" ...
    "scriptDone" ...
    "" ...
    "" ...
    <EOL> ...
    ";" ...
```

```
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> s= group expenses by EmpId;
grunt>
grunt> dump s;
2018-07-14 18:56:41,955 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 18:56:41,975 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:41,976 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:56:41,976 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:56:41,976 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:56:41,982 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:56:41,987 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:56:41,987 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:56:42,005 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:42,006 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:56:42,006 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,007 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:56:42,008 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:56:42,008 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 18:56:42,010 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 18:56:42,011 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 18:56:42,012 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 18:56:42,014 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:56:42,014 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:56:42,014 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:56:42,014 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574802014-0
```

```
2018-07-14 18:56:42,064 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:56:42,066 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,083 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:56:42,088 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:56:42,088 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:56:42,088 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:56:42,089 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:56:42,114 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:56:42,135 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1288919948_0013
2018-07-14 18:56:42,229 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:56:42,230 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:56:42,235 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:56:42,235 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:56:42,235 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:42,235 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 18:56:42,235 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:56:42,235 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:56:42,241 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:56:42,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1288919948_0013_m_000000_0
2018-07-14 18:56:42,249 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:56:42,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:56:42,257 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:56:42,257 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 18:56:42,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 18:56:42,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 18:56:42,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 18:56:42,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 18:56:42,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 18:56:42,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 18:56:42,298 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:56:42,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:56:42,302 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],s[12,3] C:  R:
2018-07-14 18:56:42,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:56:42,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 18:56:42,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 18:56:42,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 18:56:42,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 18:56:42,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 18:56:42,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1288919948_0013_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:56:42,321 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:56:42,321 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1288919948_0013_m_000000_0' done.
2018-07-14 18:56:42,321 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1288919948_0013_m_000000_0
2018-07-14 18:56:42,321 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:56:42,324 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 18:56:42,327 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1288919948_0013_r_000000_0
2018-07-14 18:56:42,341 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 18:56:42,346 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1b44f729
2018-07-14 18:56:42,371 [pool-27-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
```

```
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 18:56:42,386 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1288919948_0013_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 18:56:42,428 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#1 about to shuffle
output of map attempt_local1288919948_0013_m_000000_0 decomp: 125 len: 129 to MEMORY
2018-07-14 18:56:42,430 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local1288919948_0013_m_000000_0
2018-07-14 18:56:42,432 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 18:56:42,435 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 18:56:42,436 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:56:42,436 [pool-27-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 18:56:42,445 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 18:56:42,445 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 18:56:42,450 [pool-27-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 18:56:42,450 [pool-27-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 18:56:42,451 [pool-27-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 18:56:42,451 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 18:56:42,451 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 18:56:42,451 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:56:42,467 [pool-27-thread-1] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.skip.on is deprecated. Instead,
use mapreduce.job.skiprecords
2018-07-14 18:56:42,473 [pool-27-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:56:42,473 [pool-27-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 18:56:42,483 [pool-27-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],s[12,3] C:  R:
2018-07-14 18:56:42,497 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1288919948_0013_r_000000_0 is done. And is in the process of committing
2018-07-14 18:56:42,500 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:56:42,500 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1288919948_0013_r_000000_0 is allowed to commit now
2018-07-14 18:56:42,506 [pool-27-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1288919948_0013_r_000000_0' to
file:/tmp/temp-1902294536/tmp816634581/_temporary/0/task_local1288919948_0013_r_000000
```

```
2018-07-14 18:56:42,508 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 18:56:42,508 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1288919948_0013_r_000000_0' done.
2018-07-14 18:56:42,509 [pool-27-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1288919948_0013_r_000000_0
2018-07-14 18:56:42,513 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 18:56:42,564 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1288919948_0013
2018-07-14 18:56:42,565 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,s
2018-07-14 18:56:42,565 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],s[12,3] C:  R:
2018-07-14 18:56:42,569 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1288919948_0013]
2018-07-14 18:56:42,735 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,735 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,736 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,748 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:56:42,752 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:56:42     2018-07-14 18:56:42
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1288919948_0013        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,s      GROUP_BY
file:/tmp/temp-1902294536/tmp816634581,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp816634581"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1288919948_0013
```

```
2018-07-14 18:56:42,752 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,754 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,755 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:56:42,759 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:56:42,759 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:56:42,760 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:56:42,760 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:56:42,774 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:56:42,774 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> describe s;
s: {group: int,expenses: {(EmpId: int,Expense: int)}}
grunt> p= foreach s generate group, SUM(Expense);
2018-07-14 18:57:37,343 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 13, column 33> Invalid field projection. Projected field [Expense] does not exist
in schema: group:int,expenses:bag{:tuple(EmpId:int,Expense:int)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> p= foreach s generate group, SUM(expenses.Expense);
grunt> dump p;
2018-07-14 18:57:52,105 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 18:57:52,126 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:57:52,126 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:57:52,126 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 18:57:52,126 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:57:52,128 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:57:52,130 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 18:57:52,136 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:57:52,136 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:57:52,145 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:57:52,145 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:57:52,146 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,147 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:57:52,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:57:52,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 18:57:52,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 18:57:52,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 18:57:52,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 18:57:52,152 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:57:52,153 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:57:52,153 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:57:52,153 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574872153-0
2018-07-14 18:57:52,172 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:57:52,174 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,187 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:57:52,189 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:57:52,189 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:57:52,189 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:57:52,189 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:57:52,211 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:57:52,238 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1973883662_0014
2018-07-14 18:57:52,331 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:57:52,331 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:57:52,337 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
```

```
2018-07-14 18:57:52,337 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:57:52,337 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:57:52,337 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 18:57:52,337 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:57:52,337 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:57:52,339 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:57:52,340 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1973883662_0014_m_000000_0
2018-07-14 18:57:52,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:57:52,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:57:52,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:57:52,348 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 18:57:52,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 18:57:52,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 18:57:52,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 18:57:52,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 18:57:52,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 18:57:52,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 18:57:52,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:57:52,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:57:52,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[13,3],s[12,3] C: p[13,3],s[12,3] R: p[13,3]
2018-07-14 18:57:52,382 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
```

```
2018-07-14 18:57:52,382 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 18:57:52,382 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 18:57:52,382 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 18:57:52,382 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 18:57:52,391 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigCombiner$Combine -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[13,3],s[12,3] C: p[13,3],s[12,3] R: p[13,3]
2018-07-14 18:57:52,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 18:57:52,396 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1973883662_0014_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:57:52,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:57:52,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1973883662_0014_m_000000_0' done.
2018-07-14 18:57:52,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1973883662_0014_m_000000_0
2018-07-14 18:57:52,398 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:57:52,399 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 18:57:52,399 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1973883662_0014_r_000000_0
2018-07-14 18:57:52,411 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 18:57:52,412 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3bd35db3
2018-07-14 18:57:52,415 [pool-30-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 18:57:52,420 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1973883662_0014_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 18:57:52,427 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#2 about to shuffle
output of map attempt_local1973883662_0014_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 18:57:52,427 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1973883662_0014_m_000000_0
2018-07-14 18:57:52,427 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 18:57:52,428 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 18:57:52,429 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:57:52,429 [pool-30-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 18:57:52,430 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 18:57:52,431 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
```

```
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 18:57:52,433 [pool-30-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 18:57:52,433 [pool-30-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 18:57:52,433 [pool-30-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 18:57:52,433 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 18:57:52,433 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 18:57:52,434 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:57:52,441 [pool-30-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:57:52,441 [pool-30-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 18:57:52,444 [pool-30-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[13,3],s[12,3] C: p[13,3],s[12,3] R: p[13,3]
2018-07-14 18:57:52,446 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1973883662_0014_r_000000_0 is done. And is in the process of committing
2018-07-14 18:57:52,448 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:57:52,448 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1973883662_0014_r_000000_0 is allowed to commit now
2018-07-14 18:57:52,450 [pool-30-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1973883662_0014_r_000000_0' to
file:/tmp/temp-1902294536/tmp768773754/_temporary/0/task_local1973883662_0014_r_000000
2018-07-14 18:57:52,450 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 18:57:52,450 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1973883662_0014_r_000000_0' done.
2018-07-14 18:57:52,451 [pool-30-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1973883662_0014_r_000000_0
2018-07-14 18:57:52,451 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 18:57:52,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1973883662_0014
2018-07-14 18:57:52,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 18:57:52,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[13,3],s[12,3] C: p[13,3],s[12,3]
R: p[13,3]
2018-07-14 18:57:52,647 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,650 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,659 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,682 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:57:52,683 [main] INFO
```

```
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:57:52     2018-07-14 18:57:52
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1973883662_0014        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp768773754,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp768773754"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1973883662_0014


2018-07-14 18:57:52,685 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,687 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,688 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:57:52,692 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:57:52,692 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:57:52,693 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:57:52,693 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:57:52,706 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:57:52,706 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> describe s;
s: {group: int,expenses: {(EmpId: int,Expense: int)}}
grunt> dump s;
```

```
2018-07-14 18:58:16,042 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 18:58:16,061 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:58:16,062 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:58:16,062 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 18:58:16,062 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 18:58:16,063 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 18:58:16,064 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 18:58:16,064 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 18:58:16,076 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:58:16,076 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:58:16,078 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,078 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 18:58:16,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 18:58:16,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 18:58:16,079 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 18:58:16,079 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 18:58:16,079 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 18:58:16,080 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 18:58:16,080 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 18:58:16,080 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 18:58:16,080 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531574896080-0
2018-07-14 18:58:16,085 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 18:58:16,088 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 18:58:16,096 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 18:58:16,098 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 18:58:16,098 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:58:16,098 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 18:58:16,098 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 18:58:16,126 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 18:58:16,144 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local487449593_0015
2018-07-14 18:58:16,235 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 18:58:16,235 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 18:58:16,239 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 18:58:16,239 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 18:58:16,239 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 18:58:16,239 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 18:58:16,239 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 18:58:16,239 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 18:58:16,242 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 18:58:16,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local487449593_0015_m_000000_0
2018-07-14 18:58:16,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 18:58:16,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 18:58:16,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 18:58:16,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 18:58:16,268 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 18:58:16,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 18:58:16,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 18:58:16,269 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 18:58:16,269 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 18:58:16,269 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 18:58:16,271 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:58:16,271 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 18:58:16,278 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],s[12,3] C:  R:
2018-07-14 18:58:16,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 18:58:16,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 18:58:16,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 18:58:16,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 18:58:16,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 18:58:16,282 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 18:58:16,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local487449593_0015_m_000000_0 is done. And
is in the process of committing
2018-07-14 18:58:16,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 18:58:16,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local487449593_0015_m_000000_0' done.
2018-07-14 18:58:16,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local487449593_0015_m_000000_0
2018-07-14 18:58:16,286 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 18:58:16,287 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 18:58:16,289 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local487449593_0015_r_000000_0
2018-07-14 18:58:16,294 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 18:58:16,294 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5fded3a9
2018-07-14 18:58:16,295 [pool-33-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 18:58:16,296 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local487449593_0015_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
```

```
2018-07-14 18:58:16,297 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#3 about to shuffle
output of map attempt_local487449593_0015_m_000000_0 decomp: 125 len: 129 to MEMORY
2018-07-14 18:58:16,297 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local487449593_0015_m_000000_0
2018-07-14 18:58:16,297 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 18:58:16,297 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 18:58:16,298 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:58:16,298 [pool-33-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 18:58:16,300 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 18:58:16,300 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 18:58:16,301 [pool-33-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 18:58:16,302 [pool-33-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 18:58:16,302 [pool-33-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 18:58:16,302 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 18:58:16,302 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 18:58:16,302 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:58:16,310 [pool-33-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 18:58:16,310 [pool-33-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 18:58:16,311 [pool-33-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],s[12,3] C:  R:
2018-07-14 18:58:16,312 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local487449593_0015_r_000000_0 is done. And is in the process of committing
2018-07-14 18:58:16,314 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 18:58:16,314 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local487449593_0015_r_000000_0 is allowed to commit now
2018-07-14 18:58:16,315 [pool-33-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local487449593_0015_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1673270977/_temporary/0/task_local487449593_0015_r_00000
0
2018-07-14 18:58:16,316 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 18:58:16,316 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local487449593_0015_r_000000_0' done.
2018-07-14 18:58:16,316 [pool-33-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local487449593_0015_r_000000_0
2018-07-14 18:58:16,319 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
reduce task executor complete.
2018-07-14 18:58:16,438 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local487449593_0015
2018-07-14 18:58:16,439 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,s
2018-07-14 18:58:16,439 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],s[12,3] C:  R:
2018-07-14 18:58:16,450 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,453 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,457 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,477 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 18:58:16,479 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 18:58:16     2018-07-14 18:58:16
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local487449593_0015 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,s      GROUP_BY
file:/tmp/temp-1902294536/tmp-1673270977,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1673270977"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local487449593_0015


2018-07-14 18:58:16,483 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,486 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,487 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 18:58:16,494 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 18:58:16,494 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 18:58:16,495 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 18:58:16,495 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 18:58:16,510 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 18:58:16,510 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> p= foreach s generate group, SUM(expenses.Expense);
grunt>
grunt>
grunt> describe p;
p: {group: int,long}
grunt> describe s;
s: {group: int,expenses: {(EmpId: int,Expense: int)}}
grunt> p= foreach s generate group as EmpId, SUM(expenses.Expense);
grunt> dump p;
2018-07-14 19:07:00,501 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:07:00,520 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:00,520 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:07:00,520 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:07:00,520 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:07:00,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:07:00,533 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:07:00,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 19:07:00,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 19:07:00,547 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:00,547 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:07:00,547 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:00,549 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:07:00,549 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:07:00,549 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:07:00,549 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:07:00,549 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:07:00,549 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:07:00,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:07:00,550 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:07:00,550 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:07:00,550 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575420550-0
2018-07-14 19:07:00,559 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:07:00,561 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:00,572 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:07:00,574 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:07:00,575 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:07:00,575 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:07:00,575 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:07:00,597 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:07:00,612 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1620462377_0016
2018-07-14 19:07:00,692 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:07:00,694 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:07:00,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:07:00,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:07:00,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:00,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
```

```
2018-07-14 19:07:00,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:07:00,701 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:07:00,705 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:07:00,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1620462377_0016_m_000000_0
2018-07-14 19:07:00,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:07:00,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:07:00,713 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:07:00,713 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:07:00,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:07:00,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:07:00,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:07:00,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:07:00,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:07:00,730 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:07:00,731 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:07:00,731 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:07:00,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[15,3],s[12,3] C: p[15,3],s[12,3] R: p[15,3]
2018-07-14 19:07:00,737 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:07:00,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:07:00,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:07:00,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:07:00,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:07:00,745 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:07:00,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1620462377_0016_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:07:00,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:07:00,748 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1620462377_0016_m_000000_0' done.
2018-07-14 19:07:00,748 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1620462377_0016_m_000000_0
2018-07-14 19:07:00,748 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:07:00,748 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:07:00,752 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1620462377_0016_r_000000_0
2018-07-14 19:07:00,758 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:07:00,758 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7362eb9e
2018-07-14 19:07:00,759 [pool-36-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:07:00,760 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1620462377_0016_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:07:00,761 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#4 about to shuffle
output of map attempt_local1620462377_0016_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:07:00,761 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1620462377_0016_m_000000_0
2018-07-14 19:07:00,761 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:07:00,766 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:07:00,766 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:00,766 [pool-36-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:07:00,767 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:07:00,767 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:07:00,769 [pool-36-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:07:00,769 [pool-36-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:07:00,769 [pool-36-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:07:00,769 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:07:00,769 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
```

```
2018-07-14 19:07:00,769 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:00,776 [pool-36-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:07:00,776 [pool-36-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:07:00,780 [pool-36-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[15,3],s[12,3] C: p[15,3],s[12,3] R: p[15,3]
2018-07-14 19:07:00,783 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1620462377_0016_r_000000_0 is done. And is in the process of committing
2018-07-14 19:07:00,784 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:00,784 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1620462377_0016_r_000000_0 is allowed to commit now
2018-07-14 19:07:00,785 [pool-36-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1620462377_0016_r_000000_0' to
file:/tmp/temp-1902294536/tmp258691143/_temporary/0/task_local1620462377_0016_r_000000
2018-07-14 19:07:00,787 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:07:00,787 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1620462377_0016_r_000000_0' done.
2018-07-14 19:07:00,787 [pool-36-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1620462377_0016_r_000000_0
2018-07-14 19:07:00,788 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:07:00,996 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1620462377_0016
2018-07-14 19:07:00,996 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:07:00,996 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[15,3],s[12,3] C: p[15,3],s[12,3]
R: p[15,3]
2018-07-14 19:07:01,007 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:01,011 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:01,013 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:01,029 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:07:01,031 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:07:00     2018-07-14 19:07:01
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1620462377_0016        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
```

```
file:/tmp/temp-1902294536/tmp258691143,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp258691143"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1620462377_0016


2018-07-14 19:07:01,034 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:01,035 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:01,036 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:01,041 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:07:01,041 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:01,042 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:07:01,042 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:07:01,056 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:07:01,056 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> describe p;
p: {EmpId: int,long}
grunt> q =group p all;
grunt> dump q;
2018-07-14 19:07:33,060 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:07:33,080 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:33,081 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:07:33,081 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:07:33,081 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
```

```
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:07:33,092 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:07:33,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:07:33,094 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:07:33,094 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:07:33,103 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:33,104 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:07:33,104 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:33,105 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:07:33,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:07:33,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:07:33,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:07:33,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:07:33,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:07:33,107 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:07:33,107 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:07:33,107 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:07:33,107 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575453107-0
2018-07-14 19:07:33,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:07:33,121 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:33,129 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:07:33,131 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:07:33,131 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:07:33,131 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
```

```
2018-07-14 19:07:33,131 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:07:33,155 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:07:33,172 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1934374961_0017
2018-07-14 19:07:33,243 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:07:33,244 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:07:33,247 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:07:33,248 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:07:33,248 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:33,248 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:07:33,248 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:07:33,248 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:07:33,252 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:07:33,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1934374961_0017_m_000000_0
2018-07-14 19:07:33,257 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:07:33,259 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:07:33,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:07:33,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:07:33,367 [Service Thread] INFO
org.apache.pig.impl.util.SpillableMemoryManager - first memory handler call- Usage
threshold init = 88080384(86016K) used = 543542648(530803K) committed = 583532544(569856K)
max = 699400192(683008K), toFree = 298752584
2018-07-14 19:07:33,439 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:07:33,439 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:07:33,439 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:07:33,439 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:07:33,439 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:07:33,439 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:07:33,441 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:07:33,441 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:07:33,444 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[15,3],s[12,3] C: p[15,3],s[12,3] R: p[15,3]
2018-07-14 19:07:33,457 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:07:33,458 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:07:33,458 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:07:33,458 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:07:33,458 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:07:33,467 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:07:33,469 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1934374961_0017_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:07:33,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:07:33,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1934374961_0017_m_000000_0' done.
2018-07-14 19:07:33,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1934374961_0017_m_000000_0
2018-07-14 19:07:33,471 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:07:33,472 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:07:33,476 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1934374961_0017_r_000000_0
2018-07-14 19:07:33,482 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:07:33,482 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@fd9e69b
2018-07-14 19:07:33,483 [pool-39-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:07:33,484 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1934374961_0017_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:07:33,486 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#5 about to shuffle
output of map attempt_local1934374961_0017_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:07:33,486 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1934374961_0017_m_000000_0
```

```
2018-07-14 19:07:33,486 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:07:33,487 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:07:33,488 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:33,488 [pool-39-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:07:33,491 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:07:33,493 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:07:33,494 [pool-39-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:07:33,494 [pool-39-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:07:33,494 [pool-39-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:07:33,494 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:07:33,494 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:07:33,494 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:33,501 [pool-39-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:07:33,502 [pool-39-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:07:33,508 [pool-39-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[15,3],s[12,3] C: p[15,3],s[12,3] R: p[15,3]
2018-07-14 19:07:33,512 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1934374961_0017_r_000000_0 is done. And is in the process of committing
2018-07-14 19:07:33,514 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:33,515 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1934374961_0017_r_000000_0 is allowed to commit now
2018-07-14 19:07:33,516 [pool-39-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1934374961_0017_r_000000_0' to
file:/tmp/temp-1902294536/tmp419123922/_temporary/0/task_local1934374961_0017_r_000000
2018-07-14 19:07:33,517 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:07:33,517 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1934374961_0017_r_000000_0' done.
2018-07-14 19:07:33,517 [pool-39-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1934374961_0017_r_000000_0
2018-07-14 19:07:33,517 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:07:33,617 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1934374961_0017
2018-07-14 19:07:33,617 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
```

```
2018-07-14 19:07:33,617 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[15,3],s[12,3] C: p[15,3],s[12,3]
R: p[15,3]
2018-07-14 19:07:33,618 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:07:33,619 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:33,620 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:33,620 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:33,625 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:07:33,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:07:33,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:07:33,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:07:33,626 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:07:33,631 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:07:33,633 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:33,635 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:33,635 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:07:33,641 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:07:33,657 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:07:33,657 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:07:33,657 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:07:33,690 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:07:33,710 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local866527147_0018
2018-07-14 19:07:33,798 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:07:33,799 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:07:33,803 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:07:33,803 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation -
```

```
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:07:33,803 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:33,803 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:07:33,803 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:07:33,803 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:07:33,806 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:07:33,807 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local866527147_0018_m_000000_0
2018-07-14 19:07:33,811 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:07:33,813 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:07:33,815 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp419123922/part-r-00000:0+62
2018-07-14 19:07:33,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:07:33,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:07:33,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:07:33,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:07:33,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:07:33,992 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:07:33,993 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:07:33,993 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:07:33,994 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[16,3] C:  R:
2018-07-14 19:07:33,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:07:33,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:07:33,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:07:33,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
```

```
2018-07-14 19:07:33,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:07:33,998 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:07:33,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local866527147_0018_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:07:34,000 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:07:34,000 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local866527147_0018_m_000000_0' done.
2018-07-14 19:07:34,000 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local866527147_0018_m_000000_0
2018-07-14 19:07:34,000 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:07:34,001 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:07:34,004 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local866527147_0018_r_000000_0
2018-07-14 19:07:34,009 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:07:34,009 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@60e7e21e
2018-07-14 19:07:34,010 [pool-42-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:07:34,011 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local866527147_0018_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:07:34,012 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#6 about to shuffle
output of map attempt_local866527147_0018_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 19:07:34,012 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local866527147_0018_m_000000_0
2018-07-14 19:07:34,012 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 19:07:34,013 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:07:34,013 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:34,013 [pool-42-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:07:34,014 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:07:34,014 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:07:34,017 [pool-42-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:07:34,017 [pool-42-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 19:07:34,017 [pool-42-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
```

```
2018-07-14 19:07:34,017 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:07:34,018 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:07:34,018 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:34,026 [pool-42-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:07:34,026 [pool-42-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:07:34,027 [pool-42-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[16,3] C:  R:
2018-07-14 19:07:34,029 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local866527147_0018_r_000000_0 is done. And is in the process of committing
2018-07-14 19:07:34,031 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:07:34,031 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local866527147_0018_r_000000_0 is allowed to commit now
2018-07-14 19:07:34,033 [pool-42-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local866527147_0018_r_000000_0' to
file:/tmp/temp-1902294536/tmp-443803467/_temporary/0/task_local866527147_0018_r_000000
2018-07-14 19:07:34,033 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:07:34,033 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local866527147_0018_r_000000_0' done.
2018-07-14 19:07:34,033 [pool-42-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local866527147_0018_r_000000_0
2018-07-14 19:07:34,034 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:07:34,133 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local866527147_0018
2018-07-14 19:07:34,133 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 19:07:34,133 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[16,3] C:  R:
2018-07-14 19:07:34,136 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,137 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,137 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,142 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:07:34,143 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:07:33     2018-07-14 19:07:34
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
```

```
job_local1934374961_0017        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
job_local866527147_0018 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     q       GROUP_BY        file:/tmp/temp-1902294536/tmp-443803467,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-443803467"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1934374961_0017        ->      job_local866527147_0018,
job_local866527147_0018


2018-07-14 19:07:34,145 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,145 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,146 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,150 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,150 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,151 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:07:34,152 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:07:34,152 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:07:34,152 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:07:34,153 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:07:34,159 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:07:34,159 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,long)}}
grunt> pp = foreach q generate MAX($1);
2018-07-14 19:08:29,165 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1045:
<line 17, column 24> Could not infer the matching function for org.apache.pig.builtin.MAX
as multiple or none of them fit. Please use an explicit cast.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,long)}}
grunt>
grunt>
grunt>
grunt> p= foreach s generate group as EmpId, SUM(expenses.Expense) as sum;
```

```
grunt>
grunt>
grunt> dump p;
2018-07-14 19:08:59,541 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:08:59,559 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:08:59,559 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:08:59,559 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:08:59,559 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:08:59,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:08:59,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:08:59,563 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 19:08:59,563 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 19:08:59,574 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:08:59,574 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:08:59,575 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:08:59,576 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:08:59,576 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:08:59,577 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:08:59,577 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:08:59,577 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:08:59,577 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:08:59,578 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:08:59,579 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:08:59,579 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:08:59,579 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
```

```
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575539578-0
2018-07-14 19:08:59,586 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:08:59,587 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:08:59,598 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:08:59,600 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:08:59,600 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:08:59,600 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:08:59,601 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:08:59,625 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:08:59,641 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local777898185_0019
2018-07-14 19:08:59,744 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:08:59,745 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:08:59,748 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:08:59,748 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:08:59,748 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:08:59,748 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:08:59,748 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:08:59,748 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:08:59,750 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:08:59,750 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local777898185_0019_m_000000_0
2018-07-14 19:08:59,754 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:08:59,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------
```

```
2018-07-14 19:08:59,758 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:08:59,758 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:08:59,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:08:59,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:08:59,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:08:59,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:08:59,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:08:59,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:08:59,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:08:59,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:08:59,878 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:08:59,879 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:08:59,879 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:08:59,879 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:08:59,879 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:08:59,879 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:08:59,887 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:08:59,889 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local777898185_0019_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:08:59,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:08:59,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local777898185_0019_m_000000_0' done.
2018-07-14 19:08:59,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local777898185_0019_m_000000_0
2018-07-14 19:08:59,891 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:08:59,891 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:08:59,895 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local777898185_0019_r_000000_0
2018-07-14 19:08:59,901 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:08:59,901 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7a99f98d
2018-07-14 19:08:59,903 [pool-45-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
```

```
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:08:59,904 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local777898185_0019_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:08:59,905 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#7 about to shuffle
output of map attempt_local777898185_0019_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:08:59,905 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local777898185_0019_m_000000_0
2018-07-14 19:08:59,905 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:08:59,906 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:08:59,908 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:08:59,908 [pool-45-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:08:59,909 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:08:59,909 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:08:59,912 [pool-45-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:08:59,912 [pool-45-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:08:59,912 [pool-45-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:08:59,912 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:08:59,912 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:08:59,912 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:08:59,921 [pool-45-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:08:59,921 [pool-45-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:08:59,925 [pool-45-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:08:59,926 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local777898185_0019_r_000000_0 is done. And is in the process of committing
2018-07-14 19:08:59,927 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:08:59,927 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local777898185_0019_r_000000_0 is allowed to commit now
2018-07-14 19:08:59,928 [pool-45-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local777898185_0019_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1323623703/_temporary/0/task_local777898185_0019_r_00000
0
2018-07-14 19:08:59,932 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- reduce > reduce
2018-07-14 19:08:59,932 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local777898185_0019_r_000000_0' done.
2018-07-14 19:08:59,932 [pool-45-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local777898185_0019_r_000000_0
2018-07-14 19:08:59,932 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:09:00,053 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local777898185_0019
2018-07-14 19:09:00,053 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:09:00,053 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:09:00,063 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:00,066 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:00,069 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:00,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:09:00,096 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:08:59     2018-07-14 19:09:00
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local777898185_0019 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-1323623703,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1323623703"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local777898185_0019


2018-07-14 19:09:00,098 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:00,107 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:00,110 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:00,120 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:09:00,120 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:00,121 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:09:00,122 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:09:00,133 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:09:00,133 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> describe p;
p: {EmpId: int,sum: long}
grunt> q = group p all;
grunt> dump q;
2018-07-14 19:09:36,269 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:09:36,285 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:36,285 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:09:36,285 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:09:36,286 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:09:36,289 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:09:36,290 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:09:36,292 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:09:36,292 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:09:36,301 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:36,301 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:09:36,301 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:36,302 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
```

```
2018-07-14 19:09:36,302 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:09:36,302 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:09:36,302 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:09:36,302 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:09:36,302 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:09:36,304 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:09:36,304 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:09:36,304 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:09:36,304 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575576304-0
2018-07-14 19:09:36,313 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:09:36,314 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:36,322 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:09:36,324 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:09:36,324 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:09:36,324 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:09:36,324 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:09:36,347 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:09:36,365 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local289936611_0020
2018-07-14 19:09:36,442 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:09:36,442 [Thread-505] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:09:36,445 [Thread-505] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:09:36,445 [Thread-505] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:09:36,446 [Thread-505] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:36,446 [Thread-505] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:09:36,446 [Thread-505] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:09:36,446 [Thread-505] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:09:36,448 [Thread-505] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:09:36,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local289936611_0020_m_000000_0
2018-07-14 19:09:36,452 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:09:36,452 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:09:36,454 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:09:36,454 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:09:36,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:09:36,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:09:36,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:09:36,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:09:36,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:09:36,471 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:09:36,472 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:09:36,472 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:09:36,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:09:36,478 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:09:36,478 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:09:36,478 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:09:36,478 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:09:36,478 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
```

```
26214364(104857456); length = 33/6553600
2018-07-14 19:09:36,486 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:09:36,488 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local289936611_0020_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:09:36,489 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:09:36,489 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local289936611_0020_m_000000_0' done.
2018-07-14 19:09:36,489 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local289936611_0020_m_000000_0
2018-07-14 19:09:36,490 [Thread-505] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:09:36,491 [Thread-505] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:09:36,491 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local289936611_0020_r_000000_0
2018-07-14 19:09:36,497 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:09:36,497 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@13582efb
2018-07-14 19:09:36,498 [pool-48-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:09:36,499 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local289936611_0020_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:09:36,500 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#8 about to shuffle
output of map attempt_local289936611_0020_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:09:36,500 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local289936611_0020_m_000000_0
2018-07-14 19:09:36,500 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:09:36,501 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:09:36,501 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:09:36,501 [pool-48-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:09:36,502 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:09:36,502 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:09:36,503 [pool-48-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:09:36,503 [pool-48-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:09:36,503 [pool-48-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:09:36,503 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
```

```
2018-07-14 19:09:36,503 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:09:36,503 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:09:36,511 [pool-48-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:09:36,511 [pool-48-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:09:36,514 [pool-48-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:09:36,515 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local289936611_0020_r_000000_0 is done. And is in the process of committing
2018-07-14 19:09:36,516 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:09:36,516 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local289936611_0020_r_000000_0 is allowed to commit now
2018-07-14 19:09:36,517 [pool-48-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local289936611_0020_r_000000_0' to
file:/tmp/temp-1902294536/tmp-891250152/_temporary/0/task_local289936611_0020_r_000000
2018-07-14 19:09:36,518 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:09:36,518 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local289936611_0020_r_000000_0' done.
2018-07-14 19:09:36,518 [pool-48-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local289936611_0020_r_000000_0
2018-07-14 19:09:36,518 [Thread-505] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:09:36,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local289936611_0020
2018-07-14 19:09:36,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:09:36,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:09:36,646 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:09:36,649 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:36,650 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:36,651 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:36,660 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:09:36,661 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:09:36,661 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:09:36,661 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:09:36,664 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:09:36,675 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:09:36,678 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:36,680 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:36,681 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:09:36,687 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:09:36,697 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:09:36,697 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:09:36,697 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:09:36,718 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:09:36,733 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1651910536_0021
2018-07-14 19:09:36,807 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:09:36,808 [Thread-534] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:09:36,810 [Thread-534] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:09:36,810 [Thread-534] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:09:36,811 [Thread-534] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:36,811 [Thread-534] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:09:36,811 [Thread-534] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:09:36,811 [Thread-534] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:09:36,813 [Thread-534] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:09:36,814 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1651910536_0021_m_000000_0
2018-07-14 19:09:36,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:09:36,818 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
```

```
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:09:36,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-891250152/part-r-00000:0+62
2018-07-14 19:09:36,836 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:09:36,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:09:36,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:09:36,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:09:36,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:09:36,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:09:36,838 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:09:36,838 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:09:36,839 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[18,4] C:  R:
2018-07-14 19:09:36,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:09:36,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:09:36,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:09:36,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 19:09:36,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:09:36,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:09:36,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1651910536_0021_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:09:36,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:09:36,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1651910536_0021_m_000000_0' done.
2018-07-14 19:09:36,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1651910536_0021_m_000000_0
2018-07-14 19:09:36,844 [Thread-534] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:09:36,845 [Thread-534] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:09:36,845 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1651910536_0021_r_000000_0
2018-07-14 19:09:36,852 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:09:36,852 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4d7a4bc2
2018-07-14 19:09:36,852 [pool-51-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:09:36,854 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1651910536_0021_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:09:36,857 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#9 about to shuffle
output of map attempt_local1651910536_0021_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 19:09:36,857 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1651910536_0021_m_000000_0
2018-07-14 19:09:36,857 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 19:09:36,858 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:09:36,859 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:09:36,859 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:09:36,860 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:09:36,860 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:09:36,863 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:09:36,863 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 19:09:36,863 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:09:36,863 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:09:36,863 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:09:36,863 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:09:36,870 [pool-51-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:09:36,870 [pool-51-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:09:36,870 [pool-51-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[18,4] C:  R:
2018-07-14 19:09:36,871 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1651910536_0021_r_000000_0 is done. And is in the process of committing
2018-07-14 19:09:36,873 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:09:36,873 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1651910536_0021_r_000000_0 is allowed to commit now
2018-07-14 19:09:36,874 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1651910536_0021_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1223680871/_temporary/0/task_local1651910536_0021_r_0000
00
2018-07-14 19:09:36,875 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- reduce > reduce
2018-07-14 19:09:36,875 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1651910536_0021_r_000000_0' done.
2018-07-14 19:09:36,875 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1651910536_0021_r_000000_0
2018-07-14 19:09:36,876 [Thread-534] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:09:37,014 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1651910536_0021
2018-07-14 19:09:37,014 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 19:09:37,014 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[18,4] C:  R:
2018-07-14 19:09:37,029 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,032 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,034 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,044 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:09:37,045 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:09:36     2018-07-14 19:09:37
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1651910536_0021        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q       GROUP_BY
file:/tmp/temp-1902294536/tmp-1223680871,
job_local289936611_0020 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1223680871"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local289936611_0020 ->      job_local1651910536_0021,
job_local1651910536_0021


2018-07-14 19:09:37,046 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,047 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,048 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,051 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,052 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,052 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:09:37,055 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:09:37,055 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:09:37,056 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:09:37,056 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:09:37,068 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:09:37,068 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt> w = foreach q generate MAX(sum);
2018-07-14 19:10:10,405 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 19, column 27> Invalid field projection. Projected field [sum] does not exist in
schema: group:chararray,p:bag{:tuple(EmpId:int,sum:double)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,sum: long)}}
grunt> w = foreach q generate MAX(p.sum);
grunt> dump w;
2018-07-14 19:10:34,711 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:10:34,729 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:34,730 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:10:34,730 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:10:34,730 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:10:34,734 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:10:34,735 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:10:34,736 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:10:34,738 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:10:34,738 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
```

```
plan size after optimization: 2
2018-07-14 19:10:34,748 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:34,748 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:10:34,749 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:34,750 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:10:34,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:10:34,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:10:34,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:10:34,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:10:34,751 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:10:34,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:10:34,753 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:10:34,753 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:10:34,753 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575634752-0
2018-07-14 19:10:34,763 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:10:34,764 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:34,774 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:10:34,776 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:10:34,776 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:10:34,776 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:10:34,777 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:10:34,797 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:10:34,815 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1302891387_0022
2018-07-14 19:10:34,884 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:10:34,884 [Thread-563] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:10:34,887 [Thread-563] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:10:34,887 [Thread-563] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:10:34,887 [Thread-563] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:34,887 [Thread-563] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:10:34,887 [Thread-563] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:10:34,888 [Thread-563] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:10:34,890 [Thread-563] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:10:34,890 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1302891387_0022_m_000000_0
2018-07-14 19:10:34,894 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:34,894 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:10:34,897 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:10:34,897 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:10:34,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:10:34,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:10:34,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:10:34,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:10:34,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:10:34,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:10:34,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:34,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:10:34,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
```

```
2018-07-14 19:10:34,921 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:10:34,921 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:10:34,921 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:10:34,921 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:10:34,921 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:10:34,931 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:10:34,933 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1302891387_0022_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:10:34,934 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:10:34,934 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1302891387_0022_m_000000_0' done.
2018-07-14 19:10:34,934 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1302891387_0022_m_000000_0
2018-07-14 19:10:34,934 [Thread-563] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:10:34,935 [Thread-563] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:10:34,935 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1302891387_0022_r_000000_0
2018-07-14 19:10:34,941 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:34,941 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@11ec8829
2018-07-14 19:10:34,941 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:10:34,942 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1302891387_0022_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:10:34,944 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#10 about to shuffle
output of map attempt_local1302891387_0022_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:10:34,944 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1302891387_0022_m_000000_0
2018-07-14 19:10:34,945 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:10:34,945 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:10:34,945 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:34,946 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:10:34,950 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:34,950 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:10:34,950 [pool-54-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:10:34,951 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:10:34,951 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:10:34,951 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:34,951 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:10:34,952 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:34,959 [pool-54-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:34,959 [pool-54-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:10:34,961 [pool-54-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:10:34,964 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1302891387_0022_r_000000_0 is done. And is in the process of committing
2018-07-14 19:10:34,966 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:34,966 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1302891387_0022_r_000000_0 is allowed to commit now
2018-07-14 19:10:34,968 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1302891387_0022_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1258256709/_temporary/0/task_local1302891387_0022_r_0000
00
2018-07-14 19:10:34,971 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:10:34,971 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1302891387_0022_r_000000_0' done.
2018-07-14 19:10:34,971 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1302891387_0022_r_000000_0
2018-07-14 19:10:34,974 [Thread-563] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:10:35,087 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1302891387_0022
2018-07-14 19:10:35,087 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:10:35,087 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:10:35,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:10:35,096 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,100 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,105 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,118 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
```

```
the job
2018-07-14 19:10:35,119 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:10:35,119 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:10:35,119 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:10:35,122 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:10:35,134 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:10:35,135 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,138 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:35,138 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:10:35,144 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:10:35,162 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:10:35,162 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:10:35,162 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:10:35,183 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:10:35,198 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local28809691_0023
2018-07-14 19:10:35,264 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:10:35,264 [Thread-592] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:10:35,267 [Thread-592] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:10:35,267 [Thread-592] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:10:35,267 [Thread-592] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:35,267 [Thread-592] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:10:35,267 [Thread-592] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:10:35,267 [Thread-592] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:10:35,270 [Thread-592] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for map tasks
2018-07-14 19:10:35,270 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local28809691_0023_m_000000_0
2018-07-14 19:10:35,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:35,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:10:35,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1258256709/part-r-00000:0+62
2018-07-14 19:10:35,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:10:35,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:10:35,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:10:35,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:10:35,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:10:35,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:10:35,297 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:35,297 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:10:35,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: w[19,4],q[18,4] C:
w[19,4],q[18,4] R: w[19,4]
2018-07-14 19:10:35,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:10:35,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:10:35,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:10:35,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 19:10:35,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:10:35,311 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:10:35,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local28809691_0023_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:10:35,313 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:10:35,313 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local28809691_0023_m_000000_0' done.
2018-07-14 19:10:35,313 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
```

```
attempt_local28809691_0023_m_000000_0
2018-07-14 19:10:35,313 [Thread-592] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:10:35,314 [Thread-592] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:10:35,316 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local28809691_0023_r_000000_0
2018-07-14 19:10:35,325 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:35,326 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3396bbf9
2018-07-14 19:10:35,330 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:10:35,331 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local28809691_0023_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:10:35,333 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#11 about to shuffle
output of map attempt_local28809691_0023_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 19:10:35,333 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local28809691_0023_m_000000_0
2018-07-14 19:10:35,333 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 19:10:35,333 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:10:35,334 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:35,334 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:10:35,335 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:35,335 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 19:10:35,336 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 19:10:35,336 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 19:10:35,337 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:10:35,337 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:35,338 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 19:10:35,339 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:35,346 [pool-57-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:35,346 [pool-57-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:10:35,348 [pool-57-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: w[19,4],q[18,4] C:
```

```
w[19,4],q[18,4] R: w[19,4]
2018-07-14 19:10:35,349 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local28809691_0023_r_000000_0 is done. And is in the process of committing
2018-07-14 19:10:35,352 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:35,352 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local28809691_0023_r_000000_0 is allowed to commit now
2018-07-14 19:10:35,355 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local28809691_0023_r_000000_0' to
file:/tmp/temp-1902294536/tmp784907936/_temporary/0/task_local28809691_0023_r_000000
2018-07-14 19:10:35,356 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:10:35,356 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local28809691_0023_r_000000_0' done.
2018-07-14 19:10:35,356 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local28809691_0023_r_000000_0
2018-07-14 19:10:35,356 [Thread-592] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:10:35,468 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local28809691_0023
2018-07-14 19:10:35,468 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 19:10:35,468 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: w[19,4],q[18,4] C: w[19,4],q[18,4] R: w[19,4]
2018-07-14 19:10:35,487 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,489 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,491 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:10:35,499 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:10:34     2018-07-14 19:10:35
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1302891387_0022        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
job_local28809691_0023  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     q,w     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp784907936,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp784907936"

Counters:
```

```
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1302891387_0022        ->      job_local28809691_0023,
job_local28809691_0023


2018-07-14 19:10:35,500 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,505 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,506 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,508 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,508 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,509 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:35,511 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:10:35,511 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:35,512 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:10:35,512 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:10:35,520 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:10:35,520 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(500)
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,sum: long)}}
grunt> w = foreach q generate p.EmpId,MAX(p.sum);
grunt> dump w;
2018-07-14 19:10:56,500 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:10:56,516 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:56,517 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:10:56,517 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:10:56,517 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:10:56,521 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:10:56,522 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:10:56,522 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:10:56,522 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:10:56,531 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:56,531 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:10:56,531 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:56,532 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:10:56,532 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:10:56,532 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:10:56,532 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:10:56,533 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:10:56,533 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:10:56,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:10:56,534 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:10:56,534 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:10:56,534 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575656534-0
2018-07-14 19:10:56,541 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:10:56,547 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:56,556 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:10:56,557 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:10:56,558 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:10:56,558 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:10:56,559 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:10:56,581 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:10:56,610 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1694242624_0024
2018-07-14 19:10:56,674 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
```

```
track the job: http://localhost:8080/
2018-07-14 19:10:56,674 [Thread-621] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:10:56,685 [Thread-621] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:10:56,685 [Thread-621] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:10:56,685 [Thread-621] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:56,685 [Thread-621] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:10:56,685 [Thread-621] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:10:56,685 [Thread-621] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:10:56,688 [Thread-621] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:10:56,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1694242624_0024_m_000000_0
2018-07-14 19:10:56,691 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:56,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:10:56,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:10:56,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:10:56,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:10:56,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:10:56,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:10:56,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:10:56,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:10:56,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:10:56,719 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:56,719 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
```

```
2018-07-14 19:10:56,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:10:56,721 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:10:56,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:10:56,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:10:56,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:10:56,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:10:56,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:10:56,731 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1694242624_0024_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:10:56,736 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:10:56,736 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1694242624_0024_m_000000_0' done.
2018-07-14 19:10:56,736 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1694242624_0024_m_000000_0
2018-07-14 19:10:56,736 [Thread-621] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:10:56,736 [Thread-621] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:10:56,740 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1694242624_0024_r_000000_0
2018-07-14 19:10:56,746 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:56,746 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1942a1ca
2018-07-14 19:10:56,746 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:10:56,748 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1694242624_0024_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:10:56,749 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#12 about to shuffle
output of map attempt_local1694242624_0024_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:10:56,749 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1694242624_0024_m_000000_0
2018-07-14 19:10:56,749 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:10:56,750 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:10:56,750 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:56,750 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:10:56,751 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger -
```

```
Merging 1 sorted segments
2018-07-14 19:10:56,751 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:10:56,752 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:10:56,752 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:10:56,752 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:10:56,752 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:56,752 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:10:56,752 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:56,762 [pool-60-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:56,762 [pool-60-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:10:56,764 [pool-60-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:10:56,781 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1694242624_0024_r_000000_0 is done. And is in the process of committing
2018-07-14 19:10:56,782 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:56,782 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1694242624_0024_r_000000_0 is allowed to commit now
2018-07-14 19:10:56,786 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1694242624_0024_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1402384200/_temporary/0/task_local1694242624_0024_r_0000
00
2018-07-14 19:10:56,787 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:10:56,787 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1694242624_0024_r_000000_0' done.
2018-07-14 19:10:56,787 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1694242624_0024_r_000000_0
2018-07-14 19:10:56,787 [Thread-621] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:10:56,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1694242624_0024
2018-07-14 19:10:56,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:10:56,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:10:56,982 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:10:56,984 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:56,986 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 19:10:56,988 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:56,997 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:10:56,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:10:56,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:10:56,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:10:57,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:10:57,013 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:10:57,015 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,018 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:57,018 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:10:57,026 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:10:57,034 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:10:57,034 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:10:57,034 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:10:57,054 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:10:57,068 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1535979548_0025
2018-07-14 19:10:57,142 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:10:57,143 [Thread-650] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:10:57,147 [Thread-650] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:10:57,148 [Thread-650] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:10:57,148 [Thread-650] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:57,148 [Thread-650] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:10:57,148 [Thread-650] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 19:10:57,148 [Thread-650] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:10:57,150 [Thread-650] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:10:57,150 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1535979548_0025_m_000000_0
2018-07-14 19:10:57,155 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:57,156 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:10:57,158 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1402384200/part-r-00000:0+62
2018-07-14 19:10:57,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:10:57,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:10:57,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:10:57,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:10:57,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:10:57,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:10:57,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:57,277 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:10:57,278 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[18,4] C:  R: w[20,4]
2018-07-14 19:10:57,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:10:57,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:10:57,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:10:57,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 19:10:57,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:10:57,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:10:57,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1535979548_0025_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:10:57,282 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:10:57,282 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task 'attempt_local1535979548_0025_m_000000_0' done.
2018-07-14 19:10:57,282 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1535979548_0025_m_000000_0
2018-07-14 19:10:57,284 [Thread-650] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:10:57,284 [Thread-650] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:10:57,285 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1535979548_0025_r_000000_0
2018-07-14 19:10:57,289 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:10:57,289 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@39330d5b
2018-07-14 19:10:57,289 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:10:57,290 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1535979548_0025_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:10:57,291 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#13 about to shuffle
output of map attempt_local1535979548_0025_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 19:10:57,291 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1535979548_0025_m_000000_0
2018-07-14 19:10:57,291 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 19:10:57,292 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:10:57,293 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:57,293 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:10:57,294 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:57,294 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:10:57,295 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:10:57,295 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 19:10:57,295 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:10:57,295 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:10:57,295 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:10:57,295 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:57,302 [pool-63-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:10:57,302 [pool-63-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
```

```
2018-07-14 19:10:57,305 [pool-63-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[18,4] C:  R: w[20,4]
2018-07-14 19:10:57,305 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1535979548_0025_r_000000_0 is done. And is in the process of committing
2018-07-14 19:10:57,307 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:10:57,307 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1535979548_0025_r_000000_0 is allowed to commit now
2018-07-14 19:10:57,308 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1535979548_0025_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1217095578/_temporary/0/task_local1535979548_0025_r_0000
00
2018-07-14 19:10:57,309 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:10:57,309 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1535979548_0025_r_000000_0' done.
2018-07-14 19:10:57,309 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1535979548_0025_r_000000_0
2018-07-14 19:10:57,309 [Thread-650] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:10:57,446 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1535979548_0025
2018-07-14 19:10:57,447 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 19:10:57,447 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[18,4] C:  R: w[20,4]
2018-07-14 19:10:57,450 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,452 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,453 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,471 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:10:57,473 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:10:56     2018-07-14 19:10:57
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1535979548_0025        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q,w     GROUP_BY
file:/tmp/temp-1902294536/tmp-1217095578,
job_local1694242624_0024        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
```

```
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1217095578"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1694242624_0024        ->      job_local1535979548_0025,
job_local1535979548_0025


2018-07-14 19:10:57,474 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,476 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,478 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,483 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,483 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,484 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:10:57,487 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:10:57,487 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:10:57,488 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:10:57,488 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:10:57,497 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:10:57,497 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
({(119),(114),(110),(105),(104),(102),(101)},500)
grunt> q = group p BY EmpId;
grunt> dump q;
2018-07-14 19:11:40,000 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:11:40,016 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:40,017 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:11:40,017 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:11:40,017 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:11:40,019 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:11:40,020 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
```

```
2018-07-14 19:11:40,022 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:11:40,022 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:11:40,033 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:40,034 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:11:40,034 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,036 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:11:40,037 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:11:40,037 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:11:40,037 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:11:40,038 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:11:40,038 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:11:40,039 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:11:40,039 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:11:40,039 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:11:40,039 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575700039-0
2018-07-14 19:11:40,049 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:11:40,050 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,062 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:11:40,064 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:11:40,065 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:11:40,065 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:11:40,065 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:11:40,088 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:11:40,101 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1299515036_0026
```

```
2018-07-14 19:11:40,187 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:11:40,187 [Thread-679] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:11:40,193 [Thread-679] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:11:40,193 [Thread-679] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:11:40,194 [Thread-679] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:40,194 [Thread-679] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:11:40,194 [Thread-679] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:11:40,194 [Thread-679] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:11:40,196 [Thread-679] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:11:40,197 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1299515036_0026_m_000000_0
2018-07-14 19:11:40,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:40,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:11:40,204 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:11:40,204 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:11:40,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:11:40,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:11:40,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:11:40,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:11:40,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:11:40,220 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:11:40,221 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:40,221 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
```

```
generate code.
2018-07-14 19:11:40,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:11:40,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:11:40,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:11:40,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:11:40,226 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:11:40,226 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:11:40,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:11:40,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1299515036_0026_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:11:40,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:11:40,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1299515036_0026_m_000000_0' done.
2018-07-14 19:11:40,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1299515036_0026_m_000000_0
2018-07-14 19:11:40,235 [Thread-679] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:11:40,235 [Thread-679] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:11:40,236 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1299515036_0026_r_000000_0
2018-07-14 19:11:40,241 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:40,241 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3501e892
2018-07-14 19:11:40,242 [pool-66-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:11:40,243 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1299515036_0026_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:11:40,243 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#14 about to shuffle
output of map attempt_local1299515036_0026_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:11:40,243 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1299515036_0026_m_000000_0
2018-07-14 19:11:40,244 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:11:40,244 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:11:40,245 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:40,245 [pool-66-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
```

```
2018-07-14 19:11:40,246 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:40,246 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:11:40,246 [pool-66-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:11:40,247 [pool-66-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:11:40,247 [pool-66-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:11:40,247 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:40,247 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:11:40,247 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:40,254 [pool-66-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:40,254 [pool-66-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:11:40,256 [pool-66-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:11:40,257 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1299515036_0026_r_000000_0 is done. And is in the process of committing
2018-07-14 19:11:40,259 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:40,259 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1299515036_0026_r_000000_0 is allowed to commit now
2018-07-14 19:11:40,260 [pool-66-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1299515036_0026_r_000000_0' to
file:/tmp/temp-1902294536/tmp1721366441/_temporary/0/task_local1299515036_0026_r_00000
0
2018-07-14 19:11:40,261 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:11:40,261 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1299515036_0026_r_000000_0' done.
2018-07-14 19:11:40,261 [pool-66-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1299515036_0026_r_000000_0
2018-07-14 19:11:40,261 [Thread-679] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:11:40,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1299515036_0026
2018-07-14 19:11:40,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:11:40,389 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:11:40,392 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:11:40,395 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,397 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,398 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,407 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:11:40,407 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:11:40,407 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:11:40,407 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:11:40,408 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=62
2018-07-14 19:11:40,408 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:11:40,410 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:11:40,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:11:40,417 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,420 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:40,420 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:11:40,426 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:11:40,438 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:11:40,438 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:11:40,438 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:11:40,460 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:11:40,484 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1894076151_0027
2018-07-14 19:11:40,599 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:11:40,600 [Thread-708] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:11:40,606 [Thread-708] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:11:40,606 [Thread-708] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:11:40,606 [Thread-708] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:40,606 [Thread-708] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:11:40,606 [Thread-708] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:11:40,606 [Thread-708] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:11:40,608 [Thread-708] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:11:40,609 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1894076151_0027_m_000000_0
2018-07-14 19:11:40,615 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:40,616 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:11:40,617 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1721366441/part-r-00000:0+62
2018-07-14 19:11:40,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:11:40,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:11:40,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:11:40,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:11:40,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:11:40,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:11:40,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:40,656 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:11:40,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[21,4] C:  R:
2018-07-14 19:11:40,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:11:40,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:11:40,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:11:40,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 19:11:40,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
```

```
2018-07-14 19:11:40,659 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:11:40,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1894076151_0027_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:11:40,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:11:40,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1894076151_0027_m_000000_0' done.
2018-07-14 19:11:40,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1894076151_0027_m_000000_0
2018-07-14 19:11:40,661 [Thread-708] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:11:40,662 [Thread-708] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:11:40,662 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1894076151_0027_r_000000_0
2018-07-14 19:11:40,666 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:40,666 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@15bf1127
2018-07-14 19:11:40,667 [pool-69-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:11:40,668 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1894076151_0027_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:11:40,670 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#15 about to shuffle
output of map attempt_local1894076151_0027_m_000000_0 decomp: 99 len: 103 to MEMORY
2018-07-14 19:11:40,670 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1894076151_0027_m_000000_0
2018-07-14 19:11:40,670 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 19:11:40,672 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:11:40,672 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:40,672 [pool-69-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 91 bytes
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 99 bytes to
disk to satisfy reduce memory limit
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 103 bytes from
disk
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:40,674 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
```

```
to the last merge-pass, with 1 segments left of total size: 91 bytes
2018-07-14 19:11:40,675 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:40,681 [pool-69-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:40,681 [pool-69-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:11:40,681 [pool-69-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[21,4] C:  R:
2018-07-14 19:11:40,682 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1894076151_0027_r_000000_0 is done. And is in the process of committing
2018-07-14 19:11:40,683 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:40,683 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1894076151_0027_r_000000_0 is allowed to commit now
2018-07-14 19:11:40,684 [pool-69-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1894076151_0027_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1662451390/_temporary/0/task_local1894076151_0027_r_0000
00
2018-07-14 19:11:40,688 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:11:40,688 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1894076151_0027_r_000000_0' done.
2018-07-14 19:11:40,688 [pool-69-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1894076151_0027_r_000000_0
2018-07-14 19:11:40,688 [Thread-708] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:11:40,802 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1894076151_0027
2018-07-14 19:11:40,802 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 19:11:40,802 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[21,4] C:  R:
2018-07-14 19:11:40,812 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,815 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,818 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,834 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:11:40,836 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:11:40     2018-07-14 19:11:40
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1299515036_0026        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
```

```
job_local1894076151_0027        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q       GROUP_BY
file:/tmp/temp-1902294536/tmp-1662451390,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1662451390"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1299515036_0026        ->      job_local1894076151_0027,
job_local1894076151_0027


2018-07-14 19:11:40,838 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,839 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,840 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,844 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,845 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,846 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:40,848 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:11:40,848 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:40,848 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:11:40,848 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:11:40,858 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:11:40,858 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,300)})
(102,{(102,500)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> w = foreach q generate MAX(p.sum);
grunt> dump w;
2018-07-14 19:11:54,881 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:11:54,898 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:54,898 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2018-07-14 19:11:54,898 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:11:54,898 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:11:54,902 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:11:54,903 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:11:54,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:11:54,905 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:11:54,905 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:11:54,913 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:54,913 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:11:54,914 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:54,915 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:11:54,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:11:54,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:11:54,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:11:54,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:11:54,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:11:54,916 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:11:54,917 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:11:54,917 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:11:54,917 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575714916-0
2018-07-14 19:11:54,924 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:11:54,925 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 19:11:54,935 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:11:54,936 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:11:54,937 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:11:54,937 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:11:54,937 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:11:54,960 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:11:54,975 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1146407435_0028
2018-07-14 19:11:55,072 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:11:55,072 [Thread-737] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:11:55,075 [Thread-737] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:11:55,075 [Thread-737] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:11:55,075 [Thread-737] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:55,075 [Thread-737] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:11:55,075 [Thread-737] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:11:55,075 [Thread-737] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:11:55,077 [Thread-737] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:11:55,077 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1146407435_0028_m_000000_0
2018-07-14 19:11:55,081 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:55,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:11:55,084 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:11:55,084 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:11:55,098 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:11:55,098 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:11:55,098 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:11:55,098 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:11:55,098 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:11:55,099 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:11:55,100 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:55,100 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:11:55,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:11:55,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:11:55,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:11:55,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:11:55,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:11:55,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:11:55,111 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:11:55,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1146407435_0028_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:11:55,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:11:55,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1146407435_0028_m_000000_0' done.
2018-07-14 19:11:55,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1146407435_0028_m_000000_0
2018-07-14 19:11:55,114 [Thread-737] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:11:55,114 [Thread-737] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:11:55,116 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1146407435_0028_r_000000_0
2018-07-14 19:11:55,124 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:55,124 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@238b899b
2018-07-14 19:11:55,125 [pool-72-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=663905856, maxSingleShuffleLimit=165976464, mergeThreshold=438177888,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:11:55,126 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1146407435_0028_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
```

```
2018-07-14 19:11:55,127 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#16 about to shuffle
output of map attempt_local1146407435_0028_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:11:55,127 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1146407435_0028_m_000000_0
2018-07-14 19:11:55,128 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:11:55,129 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:11:55,129 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:55,129 [pool-72-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:11:55,131 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:55,131 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:11:55,132 [pool-72-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:11:55,132 [pool-72-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:11:55,132 [pool-72-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:11:55,132 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:55,132 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:11:55,132 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:55,142 [pool-72-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:55,142 [pool-72-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:11:55,144 [pool-72-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:11:55,145 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1146407435_0028_r_000000_0 is done. And is in the process of committing
2018-07-14 19:11:55,146 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:55,146 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1146407435_0028_r_000000_0 is allowed to commit now
2018-07-14 19:11:55,147 [pool-72-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1146407435_0028_r_000000_0' to
file:/tmp/temp-1902294536/tmp986182920/_temporary/0/task_local1146407435_0028_r_000000
2018-07-14 19:11:55,149 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:11:55,149 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1146407435_0028_r_000000_0' done.
2018-07-14 19:11:55,149 [pool-72-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1146407435_0028_r_000000_0
2018-07-14 19:11:55,149 [Thread-737] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
```

```
2018-07-14 19:11:55,274 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1146407435_0028
2018-07-14 19:11:55,274 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:11:55,274 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:11:55,275 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:11:55,277 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,279 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,281 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,297 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:11:55,298 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:11:55,298 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:11:55,298 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:11:55,300 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=62
2018-07-14 19:11:55,300 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:11:55,302 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:11:55,318 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:11:55,320 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,327 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:55,327 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:11:55,333 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:11:55,343 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:11:55,343 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:11:55,343 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
```

```
2018-07-14 19:11:55,364 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:11:55,381 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local912167681_0029
2018-07-14 19:11:55,467 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:11:55,468 [Thread-766] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:11:55,471 [Thread-766] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:11:55,471 [Thread-766] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:11:55,471 [Thread-766] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:55,471 [Thread-766] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:11:55,471 [Thread-766] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:11:55,471 [Thread-766] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:11:55,475 [Thread-766] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:11:55,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local912167681_0029_m_000000_0
2018-07-14 19:11:55,479 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:55,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:11:55,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp986182920/part-r-00000:0+62
2018-07-14 19:11:55,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:11:55,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:11:55,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:11:55,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:11:55,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:11:55,527 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:11:55,528 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 19:11:55,528 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:11:55,529 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: w[22,4],q[21,4] C:
w[22,4],q[21,4] R: w[22,4]
2018-07-14 19:11:55,532 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:11:55,532 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:11:55,532 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:11:55,532 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 19:11:55,532 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:11:55,541 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:11:55,542 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local912167681_0029_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:11:55,544 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:11:55,544 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local912167681_0029_m_000000_0' done.
2018-07-14 19:11:55,544 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local912167681_0029_m_000000_0
2018-07-14 19:11:55,544 [Thread-766] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:11:55,544 [Thread-766] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:11:55,547 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local912167681_0029_r_000000_0
2018-07-14 19:11:55,555 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:11:55,555 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1da52eb5
2018-07-14 19:11:55,556 [pool-75-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=664272896, maxSingleShuffleLimit=166068224, mergeThreshold=438420128,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:11:55,557 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local912167681_0029_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:11:55,559 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#17 about to shuffle
output of map attempt_local912167681_0029_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:11:55,559 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local912167681_0029_m_000000_0
2018-07-14 19:11:55,559 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:11:55,559 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:11:55,560 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:55,560 [pool-75-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
```

```
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:11:55,561 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:55,561 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:11:55,564 [pool-75-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:11:55,564 [pool-75-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:11:55,564 [pool-75-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:11:55,564 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:11:55,565 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:11:55,565 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:55,575 [pool-75-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:11:55,575 [pool-75-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:11:55,577 [pool-75-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: w[22,4],q[21,4] C:
w[22,4],q[21,4] R: w[22,4]
2018-07-14 19:11:55,579 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local912167681_0029_r_000000_0 is done. And is in the process of committing
2018-07-14 19:11:55,582 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:11:55,582 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local912167681_0029_r_000000_0 is allowed to commit now
2018-07-14 19:11:55,583 [pool-75-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local912167681_0029_r_000000_0' to
file:/tmp/temp-1902294536/tmp339476510/_temporary/0/task_local912167681_0029_r_000000
2018-07-14 19:11:55,583 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:11:55,583 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local912167681_0029_r_000000_0' done.
2018-07-14 19:11:55,583 [pool-75-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local912167681_0029_r_000000_0
2018-07-14 19:11:55,583 [Thread-766] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:11:55,771 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local912167681_0029
2018-07-14 19:11:55,771 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 19:11:55,771 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: w[22,4],q[21,4] C: w[22,4],q[21,4] R: w[22,4]
2018-07-14 19:11:55,780 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,785 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,789 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,810 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:11:55,811 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:11:54     2018-07-14 19:11:55
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1146407435_0028        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
job_local912167681_0029 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     q,w     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp339476510,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp339476510"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1146407435_0028        ->      job_local912167681_0029,
job_local912167681_0029


2018-07-14 19:11:55,812 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,813 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,814 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,817 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,818 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,819 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:11:55,821 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:11:55,822 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:11:55,823 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:11:55,823 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:11:55,832 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:11:55,832 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(300)
(500)
(300)
(100)
(400)
(200)
(200)
grunt> dump q;
2018-07-14 19:12:03,732 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:12:03,749 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:03,749 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:12:03,749 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:12:03,749 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:12:03,751 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:12:03,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:12:03,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:12:03,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:12:03,763 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:03,763 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:12:03,764 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:03,767 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:12:03,768 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:12:03,768 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:12:03,768 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:12:03,768 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:12:03,768 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:12:03,769 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Setting up single store job
2018-07-14 19:12:03,769 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:12:03,769 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:12:03,769 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575723769-0
2018-07-14 19:12:03,776 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:12:03,778 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:03,787 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:12:03,789 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:12:03,789 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:12:03,789 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:12:03,789 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:12:03,812 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:12:03,826 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local109928207_0030
2018-07-14 19:12:03,898 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:12:03,898 [Thread-795] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:12:03,901 [Thread-795] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:12:03,901 [Thread-795] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:12:03,901 [Thread-795] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:03,901 [Thread-795] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:12:03,901 [Thread-795] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:12:03,901 [Thread-795] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:12:03,904 [Thread-795] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:12:03,904 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local109928207_0030_m_000000_0
2018-07-14 19:12:03,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:03,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
```

```
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:12:03,912 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:12:03,912 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:12:03,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:12:03,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:12:03,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:12:03,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:12:03,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:12:03,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:12:03,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:03,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:12:03,956 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:12:03,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:12:03,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:12:03,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:12:03,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:12:03,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:12:03,968 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:12:03,969 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local109928207_0030_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:12:03,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:12:03,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local109928207_0030_m_000000_0' done.
2018-07-14 19:12:03,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local109928207_0030_m_000000_0
2018-07-14 19:12:03,979 [Thread-795] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:12:03,980 [Thread-795] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:12:03,984 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- Starting task: attempt_local109928207_0030_r_000000_0
2018-07-14 19:12:03,990 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:03,990 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@22dc72a0
2018-07-14 19:12:03,991 [pool-78-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=704643072, maxSingleShuffleLimit=176160768, mergeThreshold=465064448,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:12:03,995 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local109928207_0030_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:12:03,996 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#18 about to shuffle
output of map attempt_local109928207_0030_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:12:03,997 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local109928207_0030_m_000000_0
2018-07-14 19:12:03,997 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:12:03,998 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:12:03,999 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:03,999 [pool-78-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:12:04,000 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:04,000 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:12:04,001 [pool-78-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:12:04,001 [pool-78-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:12:04,001 [pool-78-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:12:04,001 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:04,001 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:12:04,001 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:04,011 [pool-78-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:04,011 [pool-78-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:12:04,013 [pool-78-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:12:04,015 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local109928207_0030_r_000000_0 is done. And is in the process of committing
2018-07-14 19:12:04,016 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:04,016 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local109928207_0030_r_000000_0 is allowed to commit now
2018-07-14 19:12:04,017 [pool-78-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local109928207_0030_r_000000_0' to
file:/tmp/temp-1902294536/tmp-635329554/_temporary/0/task_local109928207_0030_r_000000
2018-07-14 19:12:04,018 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:12:04,018 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local109928207_0030_r_000000_0' done.
2018-07-14 19:12:04,018 [pool-78-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local109928207_0030_r_000000_0
2018-07-14 19:12:04,018 [Thread-795] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:12:04,202 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local109928207_0030
2018-07-14 19:12:04,202 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:12:04,202 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:12:04,203 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:12:04,204 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,205 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,206 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,213 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:12:04,215 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:12:04,216 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:12:04,216 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:12:04,217 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=62
2018-07-14 19:12:04,217 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:12:04,221 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:12:04,225 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:12:04,227 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,229 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:04,229 [JobControl] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:12:04,234 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:12:04,243 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:12:04,243 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:12:04,243 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:12:04,265 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:12:04,282 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1864229347_0031
2018-07-14 19:12:04,347 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:12:04,347 [Thread-824] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:12:04,350 [Thread-824] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:12:04,350 [Thread-824] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:12:04,350 [Thread-824] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:04,350 [Thread-824] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:12:04,350 [Thread-824] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:12:04,350 [Thread-824] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:12:04,352 [Thread-824] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:12:04,352 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1864229347_0031_m_000000_0
2018-07-14 19:12:04,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:04,357 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:12:04,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-635329554/part-r-00000:0+62
2018-07-14 19:12:04,490 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:12:04,490 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:12:04,490 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:12:04,490 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:12:04,491 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:12:04,491 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:12:04,492 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:04,492 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:12:04,492 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[21,4] C:  R:
2018-07-14 19:12:04,493 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:12:04,493 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:12:04,493 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:12:04,493 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 19:12:04,493 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:12:04,494 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:12:04,495 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1864229347_0031_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:12:04,495 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:12:04,495 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1864229347_0031_m_000000_0' done.
2018-07-14 19:12:04,495 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1864229347_0031_m_000000_0
2018-07-14 19:12:04,496 [Thread-824] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:12:04,496 [Thread-824] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:12:04,496 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1864229347_0031_r_000000_0
2018-07-14 19:12:04,503 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:04,503 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@48fd39a3
2018-07-14 19:12:04,510 [pool-81-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=704643072, maxSingleShuffleLimit=176160768, mergeThreshold=465064448,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:12:04,511 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1864229347_0031_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:12:04,515 [localfetcher#19] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#19 about to shuffle
output of map attempt_local1864229347_0031_m_000000_0 decomp: 99 len: 103 to MEMORY
2018-07-14 19:12:04,515 [localfetcher#19] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1864229347_0031_m_000000_0
2018-07-14 19:12:04,516 [localfetcher#19] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 19:12:04,516 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:12:04,516 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:04,516 [pool-81-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:12:04,517 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:04,517 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 91 bytes
2018-07-14 19:12:04,646 [pool-81-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 99 bytes to
disk to satisfy reduce memory limit
2018-07-14 19:12:04,646 [pool-81-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 103 bytes from
disk
2018-07-14 19:12:04,646 [pool-81-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:12:04,646 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:04,646 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 91 bytes
2018-07-14 19:12:04,646 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:04,654 [pool-81-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:04,654 [pool-81-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:12:04,654 [pool-81-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[21,4] C:  R:
2018-07-14 19:12:04,655 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1864229347_0031_r_000000_0 is done. And is in the process of committing
2018-07-14 19:12:04,656 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:04,656 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1864229347_0031_r_000000_0 is allowed to commit now
2018-07-14 19:12:04,657 [pool-81-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1864229347_0031_r_000000_0' to
file:/tmp/temp-1902294536/tmp1731440988/_temporary/0/task_local1864229347_0031_r_00000
0
2018-07-14 19:12:04,658 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:12:04,658 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1864229347_0031_r_000000_0' done.
2018-07-14 19:12:04,658 [pool-81-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1864229347_0031_r_000000_0
2018-07-14 19:12:04,658 [Thread-824] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:12:04,728 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1864229347_0031
2018-07-14 19:12:04,728 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 19:12:04,728 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[21,4] C:  R:
2018-07-14 19:12:04,736 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1864229347_0031]
2018-07-14 19:12:04,849 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,850 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,850 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,852 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:12:04,853 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:12:03     2018-07-14 19:12:04
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local109928207_0030 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
job_local1864229347_0031        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q       GROUP_BY
file:/tmp/temp-1902294536/tmp1731440988,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp1731440988"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local109928207_0030 ->      job_local1864229347_0031,
job_local1864229347_0031


2018-07-14 19:12:04,853 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,854 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,854 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,863 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,864 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,865 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:04,870 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:12:04,871 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:04,872 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:12:04,872 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:12:04,880 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:12:04,880 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,300)})
(102,{(102,500)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> q = group p all;
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,sum: long)}}
grunt> w = foreach q generate MAX(p.sum);
grunt> dump w;
2018-07-14 19:12:30,944 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:12:30,962 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:30,962 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:12:30,962 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:12:30,962 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:12:30,965 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:12:30,965 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:12:30,966 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:12:30,966 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:12:30,966 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:12:30,975 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:30,975 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:12:30,975 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:30,976 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:12:30,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:12:30,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:12:30,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:12:30,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:12:30,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:12:30,978 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:12:30,978 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:12:30,978 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:12:30,978 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575750978-0
2018-07-14 19:12:30,985 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:12:30,987 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:30,995 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:12:30,996 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:12:30,996 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:12:30,996 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:12:30,998 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:12:31,020 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:12:31,035 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local489844286_0032
2018-07-14 19:12:31,101 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:12:31,102 [Thread-853] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:12:31,105 [Thread-853] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:12:31,105 [Thread-853] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
```

```
2018-07-14 19:12:31,105 [Thread-853] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:31,105 [Thread-853] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:12:31,105 [Thread-853] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:12:31,105 [Thread-853] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:12:31,107 [Thread-853] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:12:31,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local489844286_0032_m_000000_0
2018-07-14 19:12:31,111 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:31,111 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:12:31,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:12:31,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:12:31,129 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:12:31,129 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:12:31,129 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:12:31,129 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:12:31,129 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:12:31,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:12:31,131 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:31,131 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:12:31,132 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:12:31,133 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:12:31,133 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:12:31,133 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
```

```
2018-07-14 19:12:31,133 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:12:31,133 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:12:31,140 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:12:31,141 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local489844286_0032_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:12:31,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:12:31,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local489844286_0032_m_000000_0' done.
2018-07-14 19:12:31,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local489844286_0032_m_000000_0
2018-07-14 19:12:31,142 [Thread-853] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:12:31,142 [Thread-853] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:12:31,143 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local489844286_0032_r_000000_0
2018-07-14 19:12:31,146 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:31,146 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@66c781d9
2018-07-14 19:12:31,147 [pool-84-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:12:31,151 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local489844286_0032_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:12:31,152 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#20 about to shuffle
output of map attempt_local489844286_0032_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:12:31,152 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local489844286_0032_m_000000_0
2018-07-14 19:12:31,152 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:12:31,155 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:12:31,157 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:31,157 [pool-84-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:12:31,157 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:31,157 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:12:31,158 [pool-84-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:12:31,158 [pool-84-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:12:31,158 [pool-84-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:12:31,158 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:31,159 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:12:31,159 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:31,166 [pool-84-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:31,166 [pool-84-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:12:31,168 [pool-84-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:12:31,169 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local489844286_0032_r_000000_0 is done. And is in the process of committing
2018-07-14 19:12:31,170 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:31,170 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local489844286_0032_r_000000_0 is allowed to commit now
2018-07-14 19:12:31,172 [pool-84-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local489844286_0032_r_000000_0' to
file:/tmp/temp-1902294536/tmp-395502175/_temporary/0/task_local489844286_0032_r_000000
2018-07-14 19:12:31,173 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:12:31,173 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local489844286_0032_r_000000_0' done.
2018-07-14 19:12:31,173 [pool-84-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local489844286_0032_r_000000_0
2018-07-14 19:12:31,173 [Thread-853] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:12:31,303 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local489844286_0032
2018-07-14 19:12:31,303 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:12:31,303 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:12:31,307 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:12:31,313 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,317 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,320 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,329 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:12:31,329 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:12:31,329 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
```

```
2018-07-14 19:12:31,329 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:12:31,331 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:12:31,344 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:12:31,346 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,348 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:31,348 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:12:31,358 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:12:31,365 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:12:31,365 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:12:31,365 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:12:31,385 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:12:31,400 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1603296706_0033
2018-07-14 19:12:31,474 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:12:31,474 [Thread-882] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:12:31,477 [Thread-882] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:12:31,477 [Thread-882] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:12:31,477 [Thread-882] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:31,477 [Thread-882] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:12:31,477 [Thread-882] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:12:31,477 [Thread-882] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:12:31,479 [Thread-882] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:12:31,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1603296706_0033_m_000000_0
2018-07-14 19:12:31,483 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:31,484 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:12:31,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-395502175/part-r-00000:0+62
2018-07-14 19:12:31,501 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:12:31,501 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:12:31,501 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:12:31,501 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:12:31,501 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:12:31,502 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:12:31,503 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:31,503 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:12:31,504 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: w[24,4],q[23,4] C:
w[24,4],q[23,4] R: w[24,4]
2018-07-14 19:12:31,506 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:12:31,506 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:12:31,506 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:12:31,506 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 19:12:31,506 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:12:31,513 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:12:31,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1603296706_0033_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:12:31,516 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:12:31,516 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1603296706_0033_m_000000_0' done.
2018-07-14 19:12:31,516 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1603296706_0033_m_000000_0
2018-07-14 19:12:31,516 [Thread-882] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:12:31,516 [Thread-882] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:12:31,517 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1603296706_0033_r_000000_0
```

```
2018-07-14 19:12:31,523 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:12:31,523 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@cff1422
2018-07-14 19:12:31,524 [pool-87-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:12:31,526 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1603296706_0033_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:12:31,527 [localfetcher#21] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#21 about to shuffle
output of map attempt_local1603296706_0033_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 19:12:31,528 [localfetcher#21] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1603296706_0033_m_000000_0
2018-07-14 19:12:31,528 [localfetcher#21] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 19:12:31,529 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:12:31,529 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:31,529 [pool-87-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:12:31,533 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 19:12:31,534 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:31,550 [pool-87-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:12:31,550 [pool-87-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:12:31,552 [pool-87-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: w[24,4],q[23,4] C:
w[24,4],q[23,4] R: w[24,4]
2018-07-14 19:12:31,553 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1603296706_0033_r_000000_0 is done. And is in the process of committing
2018-07-14 19:12:31,554 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:12:31,554 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1603296706_0033_r_000000_0 is allowed to commit now
```

```
2018-07-14 19:12:31,556 [pool-87-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1603296706_0033_r_000000_0' to
file:/tmp/temp-1902294536/tmp1482174766/_temporary/0/task_local1603296706_0033_r_00000
0
2018-07-14 19:12:31,557 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:12:31,557 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1603296706_0033_r_000000_0' done.
2018-07-14 19:12:31,557 [pool-87-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1603296706_0033_r_000000_0
2018-07-14 19:12:31,557 [Thread-882] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:12:31,675 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1603296706_0033
2018-07-14 19:12:31,676 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 19:12:31,676 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: w[24,4],q[23,4] C: w[24,4],q[23,4] R: w[24,4]
2018-07-14 19:12:31,677 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,677 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,677 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,679 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:12:31,679 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:12:30     2018-07-14 19:12:31
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1603296706_0033        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q,w     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp1482174766,
job_local489844286_0032 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp1482174766"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
```

```
Job DAG:
job_local489844286_0032 ->      job_local1603296706_0033,
job_local1603296706_0033


2018-07-14 19:12:31,680 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,681 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,682 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,685 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,686 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,686 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:12:31,690 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:12:31,690 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:12:31,690 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:12:31,690 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:12:31,697 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:12:31,697 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(500)
grunt>
grunt>
grunt> w = foreach q generate MAX(p.sum);
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,sum: long)}}
grunt> w = foreach q generate p.(EmpId,MAX(p.sum));
2018-07-14 19:13:49,574 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
26, column 35>  mismatched input '(' expecting RIGHT_PAREN
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> w = foreach q generate (EmpId,MAX(p.sum));
2018-07-14 19:14:51,161 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 26, column 24> Invalid field projection. Projected field [EmpId] does not exist in
schema: group:chararray,p:bag{:tuple(EmpId:int,sum:double)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> w = foreach q generate (p.EmpId,MAX(p.sum));
grunt> dump w;
2018-07-14 19:14:58,262 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:14:58,282 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:14:58,282 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:14:58,282 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:14:58,282 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:14:58,285 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:14:58,286 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:14:58,288 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:14:58,288 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:14:58,297 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:14:58,297 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:14:58,298 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:58,299 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:14:58,299 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:14:58,299 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:14:58,299 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:14:58,299 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:14:58,299 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:14:58,300 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:14:58,300 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 19:14:58,300 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:14:58,300 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531575898300-0
2018-07-14 19:14:58,311 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:14:58,312 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:58,321 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:14:58,323 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:14:58,324 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:14:58,324 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:14:58,324 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
```

```
(combined) to process : 1
2018-07-14 19:14:58,346 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:14:58,364 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1473711459_0034
2018-07-14 19:14:58,439 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:14:58,439 [Thread-911] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:14:58,442 [Thread-911] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:14:58,442 [Thread-911] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:14:58,442 [Thread-911] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:14:58,442 [Thread-911] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:14:58,442 [Thread-911] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:14:58,442 [Thread-911] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:14:58,444 [Thread-911] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:14:58,444 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1473711459_0034_m_000000_0
2018-07-14 19:14:58,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:14:58,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:14:58,450 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:14:58,450 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:14:58,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:14:58,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:14:58,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:14:58,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:14:58,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:14:58,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
```

```
2018-07-14 19:14:58,477 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:14:58,477 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:14:58,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:14:58,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:14:58,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:14:58,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:14:58,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:14:58,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:14:58,493 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:14:58,494 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1473711459_0034_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:14:58,495 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:14:58,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1473711459_0034_m_000000_0' done.
2018-07-14 19:14:58,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1473711459_0034_m_000000_0
2018-07-14 19:14:58,496 [Thread-911] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:14:58,496 [Thread-911] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:14:58,496 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1473711459_0034_r_000000_0
2018-07-14 19:14:58,501 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:14:58,501 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@33abb258
2018-07-14 19:14:58,502 [pool-90-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:14:58,504 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1473711459_0034_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:14:58,506 [localfetcher#22] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#22 about to shuffle
output of map attempt_local1473711459_0034_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:14:58,506 [localfetcher#22] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1473711459_0034_m_000000_0
2018-07-14 19:14:58,506 [localfetcher#22] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:14:58,507 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
```

```
2018-07-14 19:14:58,507 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:14:58,507 [pool-90-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:14:58,508 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:14:58,508 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:14:58,509 [pool-90-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:14:58,509 [pool-90-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:14:58,509 [pool-90-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:14:58,509 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:14:58,509 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:14:58,509 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:14:58,516 [pool-90-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:14:58,517 [pool-90-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:14:58,518 [pool-90-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:14:58,519 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1473711459_0034_r_000000_0 is done. And is in the process of committing
2018-07-14 19:14:58,521 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:14:58,521 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1473711459_0034_r_000000_0 is allowed to commit now
2018-07-14 19:14:58,522 [pool-90-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1473711459_0034_r_000000_0' to
file:/tmp/temp-1902294536/tmp1210713651/_temporary/0/task_local1473711459_0034_r_00000
0
2018-07-14 19:14:58,526 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:14:58,526 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1473711459_0034_r_000000_0' done.
2018-07-14 19:14:58,526 [pool-90-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1473711459_0034_r_000000_0
2018-07-14 19:14:58,527 [Thread-911] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:14:58,641 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1473711459_0034
2018-07-14 19:14:58,641 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:14:58,641 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:14:58,644 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:14:58,648 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:58,652 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:58,659 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:58,667 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:14:58,668 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:14:58,668 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:14:58,669 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:14:58,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:14:58,699 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:14:58,703 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:58,707 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:14:58,707 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:14:58,713 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:14:58,721 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:14:58,721 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:14:58,721 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:14:58,741 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:14:58,756 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local747774298_0035
2018-07-14 19:14:58,825 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:14:58,825 [Thread-940] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:14:58,829 [Thread-940] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:14:58,829 [Thread-940] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:14:58,829 [Thread-940] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 19:14:58,829 [Thread-940] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:14:58,829 [Thread-940] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:14:58,829 [Thread-940] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:14:58,832 [Thread-940] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:14:58,832 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local747774298_0035_m_000000_0
2018-07-14 19:14:58,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:14:58,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:14:58,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1210713651/part-r-00000:0+62
2018-07-14 19:14:58,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:14:58,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:14:58,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:14:58,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:14:58,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:14:58,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:14:58,862 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:14:58,862 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:14:58,865 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R: w[26,4]
2018-07-14 19:14:58,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:14:58,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:14:58,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:14:58,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 19:14:58,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:14:58,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
```

```
2018-07-14 19:14:58,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local747774298_0035_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:14:58,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:14:58,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local747774298_0035_m_000000_0' done.
2018-07-14 19:14:58,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local747774298_0035_m_000000_0
2018-07-14 19:14:58,872 [Thread-940] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:14:58,873 [Thread-940] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:14:58,874 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local747774298_0035_r_000000_0
2018-07-14 19:14:58,886 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:14:58,887 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@55e5dd22
2018-07-14 19:14:58,889 [pool-93-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:14:58,891 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local747774298_0035_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:14:58,895 [localfetcher#23] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#23 about to shuffle
output of map attempt_local747774298_0035_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 19:14:58,895 [localfetcher#23] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local747774298_0035_m_000000_0
2018-07-14 19:14:58,895 [localfetcher#23] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 19:14:58,896 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:14:58,897 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:14:58,897 [pool-93-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:14:58,897 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:14:58,897 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:14:58,899 [pool-93-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:14:58,899 [pool-93-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 19:14:58,899 [pool-93-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:14:58,899 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:14:58,899 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:14:58,899 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- 1 / 1 copied.
2018-07-14 19:14:58,908 [pool-93-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:14:58,908 [pool-93-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:14:58,911 [pool-93-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R: w[26,4]
2018-07-14 19:14:58,914 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local747774298_0035_r_000000_0 is done. And is in the process of committing
2018-07-14 19:14:58,917 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:14:58,917 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local747774298_0035_r_000000_0 is allowed to commit now
2018-07-14 19:14:58,918 [pool-93-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local747774298_0035_r_000000_0' to
file:/tmp/temp-1902294536/tmp-945805193/_temporary/0/task_local747774298_0035_r_000000
2018-07-14 19:14:58,919 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:14:58,919 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local747774298_0035_r_000000_0' done.
2018-07-14 19:14:58,919 [pool-93-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local747774298_0035_r_000000_0
2018-07-14 19:14:58,919 [Thread-940] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:14:59,129 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local747774298_0035
2018-07-14 19:14:59,129 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 19:14:59,129 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[23,4] C:  R: w[26,4]
2018-07-14 19:14:59,146 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,151 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,158 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,164 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:14:59,167 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:14:58     2018-07-14 19:14:59
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1473711459_0034        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
job_local747774298_0035 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     q,w     GROUP_BY        file:/tmp/temp-1902294536/tmp-945805193,
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-945805193"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1473711459_0034        ->      job_local747774298_0035,
job_local747774298_0035


2018-07-14 19:14:59,168 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,170 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,172 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,175 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,176 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,176 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:14:59,178 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:14:59,181 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:14:59,181 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:14:59,181 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:14:59,193 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:14:59,193 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(({(119),(114),(110),(105),(104),(102),(101)},500))
grunt> describe w;
w: {org.apache.pig.builtin.totuple_18: ({(EmpId: int)},long)}
grunt> w = foreach q generate (p.EmpId,MAX(p.sum)) as rr;
grunt> describe w;
w: {rr: ({(EmpId: int)},long)}
grunt> x = JOIN details BY EmpId, w BY EmpId
>> ;
2018-07-14 19:16:10,669 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 28, column 32> Invalid field projection. Projected field [EmpId] does not exist in
schema: rr:tuple(:bag{:tuple(EmpId:int)},:double).
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> x = JOIN details BY EmpId, w BY rr.EmpId
>> ;
2018-07-14 19:16:27,196 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1128: Cannot
find field EmpId in :bag{:tuple(EmpId:int)},:double
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
```

```
grunt>
grunt> describe x;
2018-07-14 19:16:38,322 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1003: Unable
to find an operator for alias x
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe w;
w: {rr: ({(EmpId: int)},long)}
grunt> dump w;
2018-07-14 19:16:48,515 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 19:16:48,529 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:16:48,530 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:16:48,530 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 19:16:48,530 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 19:16:48,533 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 19:16:48,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 19:16:48,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 19:16:48,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 19:16:48,543 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:16:48,543 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:16:48,544 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:48,545 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:16:48,546 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:16:48,546 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:16:48,546 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 19:16:48,547 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 19:16:48,547 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:16:48,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:16:48,550 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
```

```
[pig.schematuple] is false, will not generate code.
2018-07-14 19:16:48,550 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 19:16:48,550 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531576008550-0
2018-07-14 19:16:48,558 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:16:48,560 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:48,567 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:16:48,569 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 19:16:48,569 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:16:48,569 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 19:16:48,569 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:16:48,590 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:16:48,605 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local839029817_0036
2018-07-14 19:16:48,666 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:16:48,667 [Thread-969] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:16:48,670 [Thread-969] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:16:48,670 [Thread-969] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:16:48,670 [Thread-969] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:16:48,670 [Thread-969] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:16:48,670 [Thread-969] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:16:48,670 [Thread-969] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:16:48,676 [Thread-969] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:16:48,676 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local839029817_0036_m_000000_0
2018-07-14 19:16:48,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:16:48,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
```

```
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:16:48,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 19:16:48,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 19:16:48,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:16:48,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:16:48,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:16:48,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:16:48,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:16:48,704 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:16:48,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:16:48,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 19:16:48,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:16:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:16:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:16:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:16:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 19:16:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 19:16:48,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:16:48,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local839029817_0036_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:16:48,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:16:48,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local839029817_0036_m_000000_0' done.
2018-07-14 19:16:48,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local839029817_0036_m_000000_0
2018-07-14 19:16:48,718 [Thread-969] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:16:48,718 [Thread-969] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:16:48,719 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local839029817_0036_r_000000_0
2018-07-14 19:16:48,723 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
```

```
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:16:48,723 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3b42f111
2018-07-14 19:16:48,724 [pool-96-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:16:48,726 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local839029817_0036_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:16:48,730 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#24 about to shuffle
output of map attempt_local839029817_0036_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 19:16:48,730 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local839029817_0036_m_000000_0
2018-07-14 19:16:48,730 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 19:16:48,731 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:16:48,732 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:16:48,732 [pool-96-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:16:48,733 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:16:48,733 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:16:48,734 [pool-96-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:16:48,734 [pool-96-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 19:16:48,734 [pool-96-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:16:48,734 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:16:48,734 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 19:16:48,735 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:16:48,746 [pool-96-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:16:48,746 [pool-96-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:16:48,748 [pool-96-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 19:16:48,750 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local839029817_0036_r_000000_0 is done. And is in the process of committing
2018-07-14 19:16:48,751 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:16:48,751 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local839029817_0036_r_000000_0 is allowed to commit now
2018-07-14 19:16:48,752 [pool-96-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local839029817_0036_r_000000_0' to
file:/tmp/temp-1902294536/tmp1607424121/_temporary/0/task_local839029817_0036_r_000000
2018-07-14 19:16:48,753 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:16:48,753 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local839029817_0036_r_000000_0' done.
2018-07-14 19:16:48,753 [pool-96-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local839029817_0036_r_000000_0
2018-07-14 19:16:48,753 [Thread-969] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:16:48,867 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local839029817_0036
2018-07-14 19:16:48,868 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 19:16:48,868 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 19:16:48,872 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 19:16:48,878 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:48,883 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:48,885 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:48,889 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 19:16:48,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 19:16:48,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 19:16:48,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 19:16:48,893 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 19:16:48,903 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 19:16:48,904 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:48,906 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:16:48,906 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:16:48,912 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 19:16:48,922 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:16:48,922 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
2018-07-14 19:16:48,922 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 19:16:48,944 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 19:16:48,959 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local323255575_0037
2018-07-14 19:16:49,020 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 19:16:49,020 [Thread-998] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 19:16:49,023 [Thread-998] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 19:16:49,024 [Thread-998] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 19:16:49,024 [Thread-998] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 19:16:49,024 [Thread-998] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 19:16:49,024 [Thread-998] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 19:16:49,024 [Thread-998] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 19:16:49,026 [Thread-998] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 19:16:49,026 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local323255575_0037_m_000000_0
2018-07-14 19:16:49,030 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 19:16:49,031 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 19:16:49,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1607424121/part-r-00000:0+62
2018-07-14 19:16:49,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 19:16:49,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 19:16:49,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 19:16:49,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 19:16:49,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 19:16:49,056 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
```

```
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 19:16:49,056 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:16:49,056 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 19:16:49,057 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R: w[27,4]
2018-07-14 19:16:49,058 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 19:16:49,058 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 19:16:49,059 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 19:16:49,059 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 19:16:49,059 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 19:16:49,061 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 19:16:49,064 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local323255575_0037_m_000000_0 is done. And
is in the process of committing
2018-07-14 19:16:49,065 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 19:16:49,065 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local323255575_0037_m_000000_0' done.
2018-07-14 19:16:49,065 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local323255575_0037_m_000000_0
2018-07-14 19:16:49,065 [Thread-998] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 19:16:49,067 [Thread-998] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 19:16:49,067 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local323255575_0037_r_000000_0
2018-07-14 19:16:49,072 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 19:16:49,072 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2cd2aaff
2018-07-14 19:16:49,072 [pool-99-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 19:16:49,076 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local323255575_0037_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 19:16:49,077 [localfetcher#25] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#25 about to shuffle
output of map attempt_local323255575_0037_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 19:16:49,077 [localfetcher#25] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local323255575_0037_m_000000_0
2018-07-14 19:16:49,077 [localfetcher#25] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 19:16:49,078 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 19:16:49,078 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- 1 / 1 copied.
2018-07-14 19:16:49,078 [pool-99-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 19:16:49,080 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:16:49,080 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:16:49,081 [pool-99-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 19:16:49,081 [pool-99-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 19:16:49,081 [pool-99-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 19:16:49,081 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 19:16:49,081 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 19:16:49,081 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:16:49,092 [pool-99-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 19:16:49,092 [pool-99-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 19:16:49,096 [pool-99-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R: w[27,4]
2018-07-14 19:16:49,096 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local323255575_0037_r_000000_0 is done. And is in the process of committing
2018-07-14 19:16:49,098 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 19:16:49,098 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local323255575_0037_r_000000_0 is allowed to commit now
2018-07-14 19:16:49,098 [pool-99-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local323255575_0037_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1718108302/_temporary/0/task_local323255575_0037_r_00000
0
2018-07-14 19:16:49,099 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 19:16:49,099 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local323255575_0037_r_000000_0' done.
2018-07-14 19:16:49,099 [pool-99-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local323255575_0037_r_000000_0
2018-07-14 19:16:49,099 [Thread-998] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 19:16:49,221 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local323255575_0037
2018-07-14 19:16:49,221 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 19:16:49,221 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[23,4] C:  R: w[27,4]
2018-07-14 19:16:49,231 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,233 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 19:16:49,235 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,238 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 19:16:49,239 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 19:16:48     2018-07-14 19:16:49
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local323255575_0037 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     q,w     GROUP_BY        file:/tmp/temp-1902294536/tmp-1718108302,
job_local839029817_0036 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1718108302"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local839029817_0036 ->      job_local323255575_0037,
job_local323255575_0037


2018-07-14 19:16:49,240 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,242 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,242 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,244 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,245 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,245 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 19:16:49,247 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 19:16:49,247 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 19:16:49,247 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 19:16:49,247 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 19:16:49,256 [main] INFO
```

```
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 19:16:49,256 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(({(119),(114),(110),(105),(104),(102),(101)},500))
grunt> w = foreach q generate MAX(p.sum);
grunt> dump w;
2018-07-14 20:18:46,666 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:18:46,688 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:18:46,689 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:18:46,689 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:18:46,689 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:18:46,696 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:18:46,698 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:18:46,699 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:18:46,701 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:18:46,701 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:18:46,714 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:18:46,714 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:18:46,715 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:46,716 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:18:46,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:18:46,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:18:46,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:18:46,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:18:46,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:18:46,718 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Setting up single store job
2018-07-14 20:18:46,719 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:18:46,719 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:18:46,719 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531579726718-0
2018-07-14 20:18:46,730 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:18:46,738 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:46,748 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:18:46,749 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:18:46,750 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:18:46,750 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:18:46,750 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:18:46,772 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:18:46,789 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local208826023_0038
2018-07-14 20:18:46,853 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:18:46,853 [Thread-1027] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:18:46,857 [Thread-1027] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:18:46,857 [Thread-1027] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:18:46,857 [Thread-1027] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:18:46,857 [Thread-1027] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:18:46,857 [Thread-1027] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:18:46,857 [Thread-1027] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:18:46,859 [Thread-1027] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:18:46,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local208826023_0038_m_000000_0
2018-07-14 20:18:46,863 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:18:46,864 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
```

```
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:18:46,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:18:46,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:18:46,894 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:18:46,895 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:18:46,895 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:18:46,895 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:18:46,895 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:18:46,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:18:46,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:18:46,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:18:46,900 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:18:46,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:18:46,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:18:46,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:18:46,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:18:46,902 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:18:46,912 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:18:46,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local208826023_0038_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:18:46,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:18:46,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local208826023_0038_m_000000_0' done.
2018-07-14 20:18:46,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local208826023_0038_m_000000_0
2018-07-14 20:18:46,914 [Thread-1027] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:18:46,914 [Thread-1027] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:18:46,915 [pool-102-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local208826023_0038_r_000000_0
2018-07-14 20:18:46,920 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:18:46,920 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5aae79ce
2018-07-14 20:18:46,921 [pool-102-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:18:46,926 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local208826023_0038_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:18:46,928 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#26 about to shuffle
output of map attempt_local208826023_0038_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:18:46,928 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local208826023_0038_m_000000_0
2018-07-14 20:18:46,928 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:18:46,929 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:18:46,929 [pool-102-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:18:46,929 [pool-102-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:18:46,930 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:18:46,930 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:18:46,932 [pool-102-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:18:46,932 [pool-102-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:18:46,932 [pool-102-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:18:46,932 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:18:46,932 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:18:46,932 [pool-102-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:18:46,941 [pool-102-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:18:46,942 [pool-102-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:18:46,946 [pool-102-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:18:46,947 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local208826023_0038_r_000000_0 is done. And is in the process of committing
2018-07-14 20:18:46,948 [pool-102-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
```

```
2018-07-14 20:18:46,948 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local208826023_0038_r_000000_0 is allowed to commit now
2018-07-14 20:18:46,949 [pool-102-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local208826023_0038_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1016576859/_temporary/0/task_local208826023_0038_r_00000
0
2018-07-14 20:18:46,949 [pool-102-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:18:46,949 [pool-102-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local208826023_0038_r_000000_0' done.
2018-07-14 20:18:46,949 [pool-102-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local208826023_0038_r_000000_0
2018-07-14 20:18:46,949 [Thread-1027] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:18:47,156 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local208826023_0038
2018-07-14 20:18:47,156 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:18:47,156 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:18:47,160 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:18:47,166 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,169 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,174 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,185 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:18:47,185 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:18:47,186 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:18:47,186 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:18:47,189 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:18:47,206 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:18:47,208 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,212 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:18:47,212 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:18:47,218 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
```

```
not be found. See Job or Job#setJar(String).
2018-07-14 20:18:47,227 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:18:47,227 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:18:47,227 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:18:47,252 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:18:47,267 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1246547637_0039
2018-07-14 20:18:47,332 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:18:47,332 [Thread-1056] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:18:47,337 [Thread-1056] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:18:47,337 [Thread-1056] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:18:47,337 [Thread-1056] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:18:47,337 [Thread-1056] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:18:47,337 [Thread-1056] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:18:47,338 [Thread-1056] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:18:47,340 [Thread-1056] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:18:47,340 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1246547637_0039_m_000000_0
2018-07-14 20:18:47,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:18:47,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:18:47,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1016576859/part-r-00000:0+62
2018-07-14 20:18:47,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:18:47,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:18:47,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:18:47,364 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:18:47,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:18:47,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:18:47,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:18:47,365 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:18:47,367 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: w[28,4],q[23,4] C:
w[28,4],q[23,4] R: w[28,4]
2018-07-14 20:18:47,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:18:47,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:18:47,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:18:47,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:18:47,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:18:47,379 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:18:47,380 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1246547637_0039_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:18:47,381 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:18:47,381 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1246547637_0039_m_000000_0' done.
2018-07-14 20:18:47,381 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1246547637_0039_m_000000_0
2018-07-14 20:18:47,381 [Thread-1056] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:18:47,382 [Thread-1056] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:18:47,382 [pool-105-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1246547637_0039_r_000000_0
2018-07-14 20:18:47,386 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:18:47,386 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@13ac80e3
2018-07-14 20:18:47,387 [pool-105-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:18:47,387 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1246547637_0039_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:18:47,388 [localfetcher#27] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#27 about to shuffle
output of map attempt_local1246547637_0039_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:18:47,388 [localfetcher#27] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1246547637_0039_m_000000_0
```

```
2018-07-14 20:18:47,388 [localfetcher#27] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:18:47,389 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:18:47,389 [pool-105-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:18:47,389 [pool-105-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:18:47,390 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:18:47,390 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:18:47,390 [pool-105-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:18:47,390 [pool-105-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:18:47,390 [pool-105-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:18:47,390 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:18:47,391 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:18:47,391 [pool-105-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:18:47,397 [pool-105-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:18:47,397 [pool-105-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:18:47,398 [pool-105-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: w[28,4],q[23,4] C:
w[28,4],q[23,4] R: w[28,4]
2018-07-14 20:18:47,399 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1246547637_0039_r_000000_0 is done. And is in the process of committing
2018-07-14 20:18:47,400 [pool-105-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:18:47,400 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1246547637_0039_r_000000_0 is allowed to commit now
2018-07-14 20:18:47,400 [pool-105-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1246547637_0039_r_000000_0' to
file:/tmp/temp-1902294536/tmp-194319927/_temporary/0/task_local1246547637_0039_r_00000
0
2018-07-14 20:18:47,401 [pool-105-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:18:47,401 [pool-105-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1246547637_0039_r_000000_0' done.
2018-07-14 20:18:47,401 [pool-105-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1246547637_0039_r_000000_0
2018-07-14 20:18:47,401 [Thread-1056] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:18:47,535 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1246547637_0039
2018-07-14 20:18:47,536 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 20:18:47,536 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: w[28,4],q[23,4] C: w[28,4],q[23,4] R: w[28,4]
2018-07-14 20:18:47,548 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,551 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,557 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,565 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:18:47,568 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:18:46     2018-07-14 20:18:47
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1246547637_0039        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q,w     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-194319927,
job_local208826023_0038 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-194319927"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local208826023_0038 ->      job_local1246547637_0039,
job_local1246547637_0039


2018-07-14 20:18:47,570 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,571 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,572 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,575 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,577 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:18:47,578 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:18:47,580 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:18:47,580 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:18:47,581 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:18:47,581 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:18:47,592 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:18:47,593 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(500)
grunt> w = foreach q generate MAX(p.sum);
grunt> w = foreach q generate p.EmpId,MAX(p.sum);
grunt> dump w;
2018-07-14 20:19:19,562 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:19:19,578 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:19,578 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:19:19,578 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:19:19,578 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:19:19,580 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:19:19,582 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:19:19,582 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:19:19,583 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:19:19,596 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:19,597 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:19:19,597 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:19,598 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:19:19,598 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:19:19,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:19:19,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
```

```
2018-07-14 20:19:19,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:19:19,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:19:19,600 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:19:19,600 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:19:19,600 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:19:19,600 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531579759600-0
2018-07-14 20:19:19,608 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:19:19,611 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:19,618 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:19:19,619 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:19:19,620 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:19:19,620 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:19:19,620 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:19:19,645 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:19:19,658 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1564095694_0040
2018-07-14 20:19:19,725 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:19:19,725 [Thread-1085] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:19:19,730 [Thread-1085] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:19:19,730 [Thread-1085] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:19:19,730 [Thread-1085] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:19,730 [Thread-1085] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:19:19,730 [Thread-1085] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:19:19,730 [Thread-1085] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:19:19,732 [Thread-1085] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
```

```
2018-07-14 20:19:19,732 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1564095694_0040_m_000000_0
2018-07-14 20:19:19,737 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:19,737 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:19:19,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:19:19,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:19:19,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:19:19,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:19:19,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:19:19,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:19:19,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:19:19,756 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:19:19,757 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:19,757 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:19:19,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:19:19,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:19:19,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:19:19,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:19:19,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:19:19,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:19:19,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:19:19,772 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1564095694_0040_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:19:19,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:19:19,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1564095694_0040_m_000000_0' done.
```

```
2018-07-14 20:19:19,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1564095694_0040_m_000000_0
2018-07-14 20:19:19,774 [Thread-1085] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:19:19,774 [Thread-1085] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:19:19,774 [pool-108-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1564095694_0040_r_000000_0
2018-07-14 20:19:19,780 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:19,780 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@74c27366
2018-07-14 20:19:19,780 [pool-108-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:19:19,781 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1564095694_0040_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:19:19,782 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#28 about to shuffle
output of map attempt_local1564095694_0040_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:19:19,782 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1564095694_0040_m_000000_0
2018-07-14 20:19:19,782 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:19:19,783 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:19:19,783 [pool-108-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:19,783 [pool-108-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:19:19,784 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:19,784 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:19:19,786 [pool-108-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:19:19,786 [pool-108-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:19:19,786 [pool-108-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:19:19,786 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:19,786 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:19:19,786 [pool-108-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:19,792 [pool-108-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:19,792 [pool-108-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
```

```
2018-07-14 20:19:19,794 [pool-108-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:19:19,795 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1564095694_0040_r_000000_0 is done. And is in the process of committing
2018-07-14 20:19:19,796 [pool-108-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:19,796 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1564095694_0040_r_000000_0 is allowed to commit now
2018-07-14 20:19:19,797 [pool-108-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1564095694_0040_r_000000_0' to
file:/tmp/temp-1902294536/tmp-702606949/_temporary/0/task_local1564095694_0040_r_00000
0
2018-07-14 20:19:19,797 [pool-108-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:19:19,797 [pool-108-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1564095694_0040_r_000000_0' done.
2018-07-14 20:19:19,797 [pool-108-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1564095694_0040_r_000000_0
2018-07-14 20:19:19,797 [Thread-1085] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:19:19,928 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1564095694_0040
2018-07-14 20:19:19,928 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:19:19,928 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:19:19,929 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:19:19,931 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:19,932 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:19,933 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:19,935 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:19:19,936 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:19:19,936 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:19:19,936 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:19:19,938 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:19:19,949 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:19:19,952 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:19:19,954 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:19,954 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:19:19,960 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:19:19,969 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:19:19,969 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:19:19,969 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:19:19,989 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:19:20,007 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1072259984_0041
2018-07-14 20:19:20,069 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:19:20,070 [Thread-1114] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:19:20,073 [Thread-1114] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:19:20,073 [Thread-1114] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:19:20,074 [Thread-1114] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:20,074 [Thread-1114] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:19:20,074 [Thread-1114] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:19:20,074 [Thread-1114] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:19:20,076 [Thread-1114] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:19:20,076 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1072259984_0041_m_000000_0
2018-07-14 20:19:20,081 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:20,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:19:20,083 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
```

```
split being processed file:/tmp/temp-1902294536/tmp-702606949/part-r-00000:0+62
2018-07-14 20:19:20,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:19:20,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:19:20,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:19:20,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:19:20,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:19:20,108 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:19:20,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:20,110 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:19:20,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R: w[30,4]
2018-07-14 20:19:20,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:19:20,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:19:20,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:19:20,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:19:20,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:19:20,113 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:19:20,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1072259984_0041_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:19:20,115 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:19:20,115 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1072259984_0041_m_000000_0' done.
2018-07-14 20:19:20,115 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1072259984_0041_m_000000_0
2018-07-14 20:19:20,115 [Thread-1114] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:19:20,115 [Thread-1114] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:19:20,116 [pool-111-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1072259984_0041_r_000000_0
2018-07-14 20:19:20,120 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:20,120 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@53cbc375
2018-07-14 20:19:20,121 [pool-111-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:19:20,123 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1072259984_0041_r_000000_0 Thread started: EventFetcher for fetching Map
```

```
Completion Events
2018-07-14 20:19:20,123 [localfetcher#29] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#29 about to shuffle
output of map attempt_local1072259984_0041_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:19:20,123 [localfetcher#29] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1072259984_0041_m_000000_0
2018-07-14 20:19:20,123 [localfetcher#29] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:19:20,125 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:19:20,125 [pool-111-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:20,125 [pool-111-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:19:20,126 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:20,127 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:19:20,130 [pool-111-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:19:20,130 [pool-111-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:19:20,130 [pool-111-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:19:20,130 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:20,130 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:19:20,130 [pool-111-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:20,136 [pool-111-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:20,136 [pool-111-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:19:20,138 [pool-111-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R: w[30,4]
2018-07-14 20:19:20,138 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1072259984_0041_r_000000_0 is done. And is in the process of committing
2018-07-14 20:19:20,139 [pool-111-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:20,139 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1072259984_0041_r_000000_0 is allowed to commit now
2018-07-14 20:19:20,140 [pool-111-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1072259984_0041_r_000000_0' to
file:/tmp/temp-1902294536/tmp293696172/_temporary/0/task_local1072259984_0041_r_000000
2018-07-14 20:19:20,140 [pool-111-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:19:20,141 [pool-111-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1072259984_0041_r_000000_0' done.
2018-07-14 20:19:20,141 [pool-111-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1072259984_0041_r_000000_0
2018-07-14 20:19:20,141 [Thread-1114] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
reduce task executor complete.
2018-07-14 20:19:20,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1072259984_0041
2018-07-14 20:19:20,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 20:19:20,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[23,4] C:  R: w[30,4]
2018-07-14 20:19:20,274 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,275 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,276 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,283 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:19:20,285 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:19:19     2018-07-14 20:19:20
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1072259984_0041        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q,w     GROUP_BY
file:/tmp/temp-1902294536/tmp293696172,
job_local1564095694_0040        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp293696172"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1564095694_0040        ->      job_local1072259984_0041,
job_local1072259984_0041


2018-07-14 20:19:20,286 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,286 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,287 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,291 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,292 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,293 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:20,296 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:19:20,296 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:20,297 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:19:20,297 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:19:20,308 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:19:20,308 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
({(119),(114),(110),(105),(104),(102),(101)},500)
grunt> dump q;
2018-07-14 20:19:33,127 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:19:33,146 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:33,147 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:19:33,147 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:19:33,147 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:19:33,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:19:33,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:19:33,150 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:19:33,150 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:19:33,161 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:33,161 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:19:33,161 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:33,163 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:19:33,163 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:19:33,163 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:19:33,163 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:19:33,164 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:19:33,164 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:19:33,164 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:19:33,165 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:19:33,165 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:19:33,165 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531579773165-0
2018-07-14 20:19:33,172 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:19:33,174 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:33,181 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:19:33,183 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:19:33,183 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:19:33,183 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:19:33,184 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:19:33,206 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:19:33,221 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2061315230_0042
2018-07-14 20:19:33,293 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:19:33,293 [Thread-1143] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:19:33,297 [Thread-1143] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:19:33,297 [Thread-1143] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:19:33,297 [Thread-1143] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:33,297 [Thread-1143] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:19:33,297 [Thread-1143] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:19:33,297 [Thread-1143] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:19:33,299 [Thread-1143] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:19:33,300 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2061315230_0042_m_000000_0
2018-07-14 20:19:33,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:33,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:19:33,306 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:19:33,306 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:19:33,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:19:33,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:19:33,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:19:33,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:19:33,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:19:33,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:19:33,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:33,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:19:33,329 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:19:33,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:19:33,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:19:33,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:19:33,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:19:33,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:19:33,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:19:33,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2061315230_0042_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:19:33,343 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:19:33,343 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2061315230_0042_m_000000_0' done.
2018-07-14 20:19:33,343 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2061315230_0042_m_000000_0
2018-07-14 20:19:33,343 [Thread-1143] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:19:33,343 [Thread-1143] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:19:33,344 [pool-114-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2061315230_0042_r_000000_0
2018-07-14 20:19:33,349 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:33,349 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@78959f12
2018-07-14 20:19:33,350 [pool-114-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:19:33,350 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2061315230_0042_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:19:33,351 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#30 about to shuffle
output of map attempt_local2061315230_0042_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:19:33,352 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local2061315230_0042_m_000000_0
2018-07-14 20:19:33,352 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:19:33,352 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:19:33,353 [pool-114-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:33,353 [pool-114-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:19:33,353 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:33,354 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:19:33,477 [pool-114-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:19:33,477 [pool-114-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:19:33,477 [pool-114-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:19:33,477 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:33,477 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:19:33,478 [pool-114-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:33,485 [pool-114-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
```

```
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:33,485 [pool-114-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:19:33,487 [pool-114-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:19:33,488 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2061315230_0042_r_000000_0 is done. And is in the process of committing
2018-07-14 20:19:33,489 [pool-114-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:33,489 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local2061315230_0042_r_000000_0 is allowed to commit now
2018-07-14 20:19:33,490 [pool-114-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2061315230_0042_r_000000_0' to
file:/tmp/temp-1902294536/tmp-902546838/_temporary/0/task_local2061315230_0042_r_00000
0
2018-07-14 20:19:33,491 [pool-114-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:19:33,491 [pool-114-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2061315230_0042_r_000000_0' done.
2018-07-14 20:19:33,491 [pool-114-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2061315230_0042_r_000000_0
2018-07-14 20:19:33,491 [Thread-1143] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:19:33,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2061315230_0042
2018-07-14 20:19:33,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:19:33,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:19:33,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:19:33,674 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:33,675 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:33,675 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:33,677 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:19:33,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:19:33,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:19:33,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:19:33,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:19:33,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
```

```
map-reduce job(s) waiting for submission.
2018-07-14 20:19:33,684 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:33,685 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:33,685 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:19:33,691 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:19:33,698 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:19:33,698 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:19:33,698 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:19:33,722 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:19:33,739 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1240771768_0043
2018-07-14 20:19:33,822 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:19:33,823 [Thread-1172] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:19:33,828 [Thread-1172] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:19:33,828 [Thread-1172] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:19:33,828 [Thread-1172] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:33,828 [Thread-1172] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:19:33,828 [Thread-1172] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:19:33,828 [Thread-1172] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:19:33,830 [Thread-1172] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:19:33,830 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1240771768_0043_m_000000_0
2018-07-14 20:19:33,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:33,835 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------
```

```
2018-07-14 20:19:33,835 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-902546838/part-r-00000:0+62
2018-07-14 20:19:33,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:19:33,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:19:33,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:19:33,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:19:33,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:19:33,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:19:33,852 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:33,852 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:19:33,853 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R:
2018-07-14 20:19:33,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:19:33,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:19:33,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:19:33,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:19:33,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:19:33,856 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:19:33,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1240771768_0043_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:19:33,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:19:33,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1240771768_0043_m_000000_0' done.
2018-07-14 20:19:33,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1240771768_0043_m_000000_0
2018-07-14 20:19:33,857 [Thread-1172] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:19:33,857 [Thread-1172] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:19:33,860 [pool-117-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1240771768_0043_r_000000_0
2018-07-14 20:19:33,866 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:19:33,866 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@214da618
2018-07-14 20:19:33,869 [pool-117-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
```

```
2018-07-14 20:19:33,870 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1240771768_0043_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:19:33,871 [localfetcher#31] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#31 about to shuffle
output of map attempt_local1240771768_0043_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:19:33,871 [localfetcher#31] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1240771768_0043_m_000000_0
2018-07-14 20:19:33,871 [localfetcher#31] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:19:33,874 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:19:33,875 [pool-117-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:33,875 [pool-117-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:19:33,876 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:33,876 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:19:33,877 [pool-117-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:19:33,878 [pool-117-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:19:33,878 [pool-117-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:19:33,878 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:19:33,878 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:19:33,878 [pool-117-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:33,885 [pool-117-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:19:33,885 [pool-117-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:19:33,885 [pool-117-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R:
2018-07-14 20:19:33,886 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1240771768_0043_r_000000_0 is done. And is in the process of committing
2018-07-14 20:19:33,887 [pool-117-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:19:33,887 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1240771768_0043_r_000000_0 is allowed to commit now
2018-07-14 20:19:33,889 [pool-117-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1240771768_0043_r_000000_0' to
file:/tmp/temp-1902294536/tmp617044652/_temporary/0/task_local1240771768_0043_r_000000
2018-07-14 20:19:33,889 [pool-117-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:19:33,889 [pool-117-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1240771768_0043_r_000000_0' done.
2018-07-14 20:19:33,889 [pool-117-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1240771768_0043_r_000000_0
2018-07-14 20:19:33,889 [Thread-1172] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:19:34,023 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1240771768_0043
2018-07-14 20:19:34,023 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 20:19:34,023 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[23,4] C:  R:
2018-07-14 20:19:34,024 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,025 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,025 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,026 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:19:34,029 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:19:33     2018-07-14 20:19:34
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1240771768_0043        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q       GROUP_BY
file:/tmp/temp-1902294536/tmp617044652,
job_local2061315230_0042        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp617044652"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local2061315230_0042        ->      job_local1240771768_0043,
job_local1240771768_0043


2018-07-14 20:19:34,029 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,030 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:19:34,030 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,031 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,032 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,033 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:19:34,034 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:19:34,034 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:19:34,034 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:19:34,035 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:19:34,043 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:19:34,043 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt> w = foreach q generate MAX(p.sum);
grunt> describe q;
q: {group: chararray,p: {(EmpId: int,sum: long)}}
grunt> w = foreach q generate details.EmpId,MAX(p.sum);
grunt> dump w;
2018-07-14 20:20:57,956 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:20:57,974 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:57,975 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:20:57,975 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:20:57,975 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:20:57,978 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:20:57,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:20:57,981 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:20:57,981 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 20:20:57,981 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 20:20:57,992 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:57,993 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:20:57,993 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:20:57,994 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:20:57,994 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:20:57,994 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:20:57,994 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:20:57,995 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:20:57,995 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:20:57,995 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:20:57,996 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:20:57,996 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:20:57,996 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531579857996-0
2018-07-14 20:20:58,005 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:20:58,005 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:20:58,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:20:58,007 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:20:58,007 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:20:58,007 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531579858007-0
2018-07-14 20:20:58,013 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 2
map-reduce job(s) waiting for submission.
2018-07-14 20:20:58,015 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,023 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:20:58,025 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:20:58,026 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:20:58,026 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:20:58,026 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
```

```
2018-07-14 20:20:58,047 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:20:58,062 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local625857616_0044
2018-07-14 20:20:58,130 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:20:58,131 [Thread-1201] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:20:58,132 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,134 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:20:58,134 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:20:58,134 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:58,134 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:20:58,135 [Thread-1201] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:20:58,135 [Thread-1201] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:20:58,137 [Thread-1201] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:20:58,137 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local625857616_0044_m_000000_0
2018-07-14 20:20:58,140 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:20:58,140 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:20:58,141 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:20:58,141 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:20:58,141 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:20:58,143 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:20:58,144 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:20:58,146 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:20:58,146 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
```

```
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:20:58,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:20:58,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:20:58,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:20:58,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:20:58,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:20:58,165 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:20:58,167 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:20:58,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:20:58,173 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:20:58,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:20:58,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:20:58,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:20:58,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:20:58,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:20:58,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:20:58,183 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local625857616_0044_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:20:58,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:20:58,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local625857616_0044_m_000000_0' done.
2018-07-14 20:20:58,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local625857616_0044_m_000000_0
2018-07-14 20:20:58,186 [Thread-1201] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:20:58,186 [Thread-1201] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:20:58,186 [pool-120-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local625857616_0044_r_000000_0
2018-07-14 20:20:58,187 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:20:58,194 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:20:58,194 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@18cd45c6
2018-07-14 20:20:58,195 [pool-120-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
```

```
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:20:58,197 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local625857616_0044_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:20:58,198 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#32 about to shuffle
output of map attempt_local625857616_0044_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:20:58,198 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local625857616_0044_m_000000_0
2018-07-14 20:20:58,198 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:20:58,199 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:20:58,202 [pool-120-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:20:58,202 [pool-120-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:20:58,204 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:20:58,204 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:20:58,205 [pool-120-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:20:58,205 [pool-120-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:20:58,205 [pool-120-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:20:58,205 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:20:58,206 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:20:58,206 [pool-120-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:20:58,216 [pool-120-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:20:58,216 [pool-120-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:20:58,218 [pool-120-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:20:58,219 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local625857616_0044_r_000000_0 is done. And is in the process of committing
2018-07-14 20:20:58,220 [pool-120-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:20:58,220 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local625857616_0044_r_000000_0 is allowed to commit now
2018-07-14 20:20:58,221 [pool-120-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local625857616_0044_r_000000_0' to
file:/tmp/temp-1902294536/tmp388072264/_temporary/0/task_local625857616_0044_r_000000
2018-07-14 20:20:58,222 [pool-120-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:20:58,222 [pool-120-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
'attempt_local625857616_0044_r_000000_0' done.
2018-07-14 20:20:58,222 [pool-120-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local625857616_0044_r_000000_0
2018-07-14 20:20:58,222 [Thread-1201] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:20:58,238 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:58,238 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:20:58,242 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1673570258_0045
2018-07-14 20:20:58,372 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:20:58,373 [Thread-1228] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:20:58,378 [Thread-1228] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:20:58,378 [Thread-1228] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:20:58,378 [Thread-1228] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:58,378 [Thread-1228] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:20:58,378 [Thread-1228] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:20:58,380 [Thread-1228] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:20:58,380 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1673570258_0045_m_000000_0
2018-07-14 20:20:58,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:20:58,387 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:20:58,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:20:58,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:20:58,395 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:20:58,395 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:20:58,396 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
```

```
being processed per job phase (AliasName[line,offset]): M: details[8,10],details[-1,-1]
C:  R:
2018-07-14 20:20:58,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:20:58,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1673570258_0045_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:20:58,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:20:58,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1673570258_0045_m_000000_0 is allowed
to commit now
2018-07-14 20:20:58,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1673570258_0045_m_000000_0' to
file:/tmp/temp-1902294536/tmp1455759624/_temporary/0/task_local1673570258_0045_m_00000
0
2018-07-14 20:20:58,400 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:20:58,400 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1673570258_0045_m_000000_0' done.
2018-07-14 20:20:58,400 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1673570258_0045_m_000000_0
2018-07-14 20:20:58,400 [Thread-1228] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:20:58,515 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local625857616_0044
2018-07-14 20:20:58,515 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:20:58,515 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:20:58,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1673570258_0045
2018-07-14 20:20:58,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 20:20:58,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1] C:  R:
2018-07-14 20:20:58,518 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:20:58,519 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,520 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,521 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,527 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,529 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,530 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,531 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
```

```
2018-07-14 20:20:58,532 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:20:58,532 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:20:58,532 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:20:58,533 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:20:58,544 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:20:58,546 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,547 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:58,547 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:20:58,552 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:20:58,560 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:20:58,561 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:20:58,561 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:20:58,581 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:20:58,596 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local395787851_0046
2018-07-14 20:20:58,669 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:20:58,669 [Thread-1251] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:20:58,673 [Thread-1251] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:20:58,673 [Thread-1251] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:20:58,673 [Thread-1251] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:20:58,673 [Thread-1251] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:20:58,673 [Thread-1251] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:20:58,673 [Thread-1251] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:20:58,679 [Thread-1251] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
```

```
2018-07-14 20:20:58,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local395787851_0046_m_000000_0
2018-07-14 20:20:58,683 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:20:58,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:20:58,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp388072264/part-r-00000:0+62
2018-07-14 20:20:58,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:20:58,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:20:58,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:20:58,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:20:58,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:20:58,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:20:58,700 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:20:58,700 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:20:58,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: w[32,4],q[23,4] C:
w[32,4],q[23,4] R: w[32,4]
2018-07-14 20:20:58,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:20:58,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:20:58,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:20:58,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:20:58,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:20:58,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:20:58,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local395787851_0046_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:20:58,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:20:58,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local395787851_0046_m_000000_0' done.
2018-07-14 20:20:58,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local395787851_0046_m_000000_0
```

```
2018-07-14 20:20:58,711 [Thread-1251] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:20:58,711 [Thread-1251] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:20:58,712 [pool-125-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local395787851_0046_r_000000_0
2018-07-14 20:20:58,717 [pool-125-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:20:58,717 [pool-125-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3585cb1c
2018-07-14 20:20:58,718 [pool-125-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:20:58,723 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local395787851_0046_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:20:58,724 [localfetcher#33] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#33 about to shuffle
output of map attempt_local395787851_0046_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:20:58,724 [localfetcher#33] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local395787851_0046_m_000000_0
2018-07-14 20:20:58,724 [localfetcher#33] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:20:58,726 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:20:58,727 [pool-125-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:20:58,727 [pool-125-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:20:58,728 [pool-125-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:20:58,734 [pool-125-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:20:58,734 [pool-125-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:20:58,737 [pool-125-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: w[32,4],q[23,4] C:
```

```
w[32,4],q[23,4] R: w[32,4]
2018-07-14 20:20:58,738 [pool-125-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:20:58,747 [pool-125-thread-1] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:20:58,747 [pool-125-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:20:58,750 [Thread-1251] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:20:58,751 [Thread-1251] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local395787851_0046
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,Amitabh,20000,1), 2nd
:(102,Shahrukh,10000,2) (common cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should
be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:529)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,Amitabh,20000,1), 2nd :(102,Shahrukh,10000,2) (common
cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextInteger(POUserFunc.java:375)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:347)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.processPlan(POForEach.java:404)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:321)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.runPipeline(PigGenericMapReduce.java:465)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.processOnePackageOutput(PigGenericMapReduce.java:433)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:413)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:262)
        at org.apache.hadoop.mapreduce.Reducer.run(Reducer.java:171)
        at org.apache.hadoop.mapred.ReduceTask.runNewReducer(ReduceTask.java:627)
        at org.apache.hadoop.mapred.ReduceTask.run(ReduceTask.java:389)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$ReduceTaskRunnable.run(LocalJobRunner.java
:319)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
```

```
2018-07-14 20:20:58,870 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local395787851_0046
2018-07-14 20:20:58,871 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 20:20:58,871 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: w[32,4],q[23,4] C: w[32,4],q[23,4] R: w[32,4]
2018-07-14 20:20:58,872 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 20:20:58,872 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local395787851_0046 has failed! Stop running all dependent jobs
2018-07-14 20:20:58,872 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:20:58,873 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,874 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,874 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 20:20:58,875 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:20:57     2018-07-14 20:20:58
GROUP_BY

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1673570258_0045        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details MAP_ONLY
job_local625857616_0044 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local395787851_0046 q,w     GROUP_BY,COMBINER       Message: Job failed!
file:/tmp/temp-1902294536/tmp-583298746,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp-583298746"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local625857616_0044 ->      job_local395787851_0046,
```

```
job_local1673570258_0045        ->      job_local395787851_0046,
job_local395787851_0046


2018-07-14 20:20:58,876 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,877 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,877 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,879 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,879 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,880 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:20:58,881 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 20:20:58,882 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias w
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> 2018-07-14 20:21:04,717 [communication thread] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce

grunt> w = foreach q generate all.EmpId,MAX(p.sum);
2018-07-14 20:21:20,857 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
33, column 32>  mismatched input ',' expecting LEFT_PAREN
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe q
2018-07-14 20:21:26,478 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:21:26,479 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
q: {group: chararray,p: {(EmpId: int,sum: long)}}
grunt> w = foreach q generate MAX(p.sum) as max;
grunt> dump w;
2018-07-14 20:24:41,676 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:24:41,693 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:41,694 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:24:41,694 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:24:41,694 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:24:41,697 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:24:41,697 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:24:41,698 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:24:41,700 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
```

```
2018-07-14 20:24:41,700 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:24:41,710 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:41,710 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:24:41,711 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:41,712 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:24:41,713 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:24:41,714 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:24:41,714 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:24:41,714 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:24:41,714 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:24:41,716 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:24:41,717 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:24:41,717 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:24:41,717 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580081717-0
2018-07-14 20:24:41,727 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:24:41,728 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:41,737 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:24:41,738 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:24:41,739 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:24:41,739 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:24:41,739 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:24:41,760 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:24:41,774 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local32778805_0047
2018-07-14 20:24:41,909 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:24:41,910 [Thread-1278] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter set in config null
2018-07-14 20:24:41,916 [Thread-1278] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:24:41,916 [Thread-1278] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:24:41,916 [Thread-1278] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:41,917 [Thread-1278] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:24:41,917 [Thread-1278] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:24:41,917 [Thread-1278] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:24:41,919 [Thread-1278] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:24:41,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local32778805_0047_m_000000_0
2018-07-14 20:24:41,924 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:41,925 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:24:41,927 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:24:41,927 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:24:41,943 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:24:41,943 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:24:41,943 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:24:41,943 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:24:41,943 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:24:41,944 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:24:41,945 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:41,945 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:24:41,947 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
```

```
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:24:41,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:24:41,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:24:41,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:24:41,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:24:41,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:24:41,958 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:24:41,959 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local32778805_0047_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:24:41,960 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:24:41,960 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local32778805_0047_m_000000_0' done.
2018-07-14 20:24:41,960 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local32778805_0047_m_000000_0
2018-07-14 20:24:41,960 [Thread-1278] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:24:41,961 [Thread-1278] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:24:41,961 [pool-128-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local32778805_0047_r_000000_0
2018-07-14 20:24:41,968 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:41,969 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4f26d130
2018-07-14 20:24:41,969 [pool-128-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:24:41,971 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local32778805_0047_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:24:41,971 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#34 about to shuffle
output of map attempt_local32778805_0047_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:24:41,972 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local32778805_0047_m_000000_0
2018-07-14 20:24:41,972 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:24:41,974 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:24:41,975 [pool-128-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:41,975 [pool-128-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:24:41,976 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
```

```
2018-07-14 20:24:41,976 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:24:41,977 [pool-128-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:24:41,977 [pool-128-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:24:41,977 [pool-128-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:24:41,977 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:41,977 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:24:41,977 [pool-128-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:41,985 [pool-128-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:41,985 [pool-128-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:24:41,988 [pool-128-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:24:41,989 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local32778805_0047_r_000000_0 is done. And is in the process of committing
2018-07-14 20:24:41,990 [pool-128-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:41,990 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local32778805_0047_r_000000_0 is allowed to commit now
2018-07-14 20:24:41,991 [pool-128-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local32778805_0047_r_000000_0' to
file:/tmp/temp-1902294536/tmp1696216383/_temporary/0/task_local32778805_0047_r_000000
2018-07-14 20:24:41,992 [pool-128-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:24:41,992 [pool-128-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local32778805_0047_r_000000_0' done.
2018-07-14 20:24:41,992 [pool-128-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local32778805_0047_r_000000_0
2018-07-14 20:24:41,992 [Thread-1278] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:24:42,110 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local32778805_0047
2018-07-14 20:24:42,110 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:24:42,110 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:24:42,111 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:24:42,112 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,112 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,113 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,116 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:24:42,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:24:42,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:24:42,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:24:42,117 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:24:42,125 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:24:42,127 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,129 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:42,129 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:24:42,135 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:24:42,143 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:24:42,143 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:24:42,143 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:24:42,165 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:24:42,180 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1694791260_0048
2018-07-14 20:24:42,262 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:24:42,263 [Thread-1307] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:24:42,266 [Thread-1307] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:24:42,267 [Thread-1307] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:24:42,267 [Thread-1307] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:42,267 [Thread-1307] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:24:42,267 [Thread-1307] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:24:42,267 [Thread-1307] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:24:42,269 [Thread-1307] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:24:42,269 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1694791260_0048_m_000000_0
2018-07-14 20:24:42,274 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:42,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:24:42,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1696216383/part-r-00000:0+62
2018-07-14 20:24:42,290 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:24:42,290 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:24:42,290 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:24:42,290 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:24:42,290 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:24:42,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:24:42,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:42,292 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:24:42,293 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: w[33,4],q[23,4] C:
w[33,4],q[23,4] R: w[33,4]
2018-07-14 20:24:42,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:24:42,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:24:42,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:24:42,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:24:42,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:24:42,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:24:42,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1694791260_0048_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:24:42,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:24:42,305 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task 'attempt_local1694791260_0048_m_000000_0' done.
2018-07-14 20:24:42,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1694791260_0048_m_000000_0
2018-07-14 20:24:42,305 [Thread-1307] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:24:42,306 [Thread-1307] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:24:42,306 [pool-131-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1694791260_0048_r_000000_0
2018-07-14 20:24:42,318 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:42,318 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1af23e5c
2018-07-14 20:24:42,319 [pool-131-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:24:42,319 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1694791260_0048_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:24:42,321 [localfetcher#35] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#35 about to shuffle
output of map attempt_local1694791260_0048_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:24:42,321 [localfetcher#35] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1694791260_0048_m_000000_0
2018-07-14 20:24:42,321 [localfetcher#35] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:24:42,322 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:24:42,322 [pool-131-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:42,322 [pool-131-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:24:42,326 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:42,326 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:24:42,326 [pool-131-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:24:42,327 [pool-131-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:24:42,327 [pool-131-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:24:42,327 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:42,327 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:24:42,327 [pool-131-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:42,334 [pool-131-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:42,334 [pool-131-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
```

```
- SchemaTupleBackend has already been initialized
2018-07-14 20:24:42,337 [pool-131-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: w[33,4],q[23,4] C:
w[33,4],q[23,4] R: w[33,4]
2018-07-14 20:24:42,337 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1694791260_0048_r_000000_0 is done. And is in the process of committing
2018-07-14 20:24:42,339 [pool-131-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:42,339 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1694791260_0048_r_000000_0 is allowed to commit now
2018-07-14 20:24:42,340 [pool-131-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1694791260_0048_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1282530421/_temporary/0/task_local1694791260_0048_r_0000
00
2018-07-14 20:24:42,341 [pool-131-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:24:42,341 [pool-131-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1694791260_0048_r_000000_0' done.
2018-07-14 20:24:42,341 [pool-131-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1694791260_0048_r_000000_0
2018-07-14 20:24:42,341 [Thread-1307] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:24:42,463 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1694791260_0048
2018-07-14 20:24:42,463 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q,w
2018-07-14 20:24:42,463 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: w[33,4],q[23,4] C: w[33,4],q[23,4] R: w[33,4]
2018-07-14 20:24:42,465 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,465 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,466 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,467 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:24:42,469 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:24:41     2018-07-14 20:24:42
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1694791260_0048        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q,w     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-1282530421,
job_local32778805_0047  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER

Input(s):
```

```
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1282530421"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local32778805_0047  ->      job_local1694791260_0048,
job_local1694791260_0048


2018-07-14 20:24:42,471 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,472 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,473 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,475 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,476 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,477 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:42,478 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:24:42,478 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:42,479 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:24:42,479 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:24:42,487 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:24:42,487 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(500)
grunt> dump q;
2018-07-14 20:24:51,926 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:24:51,945 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:51,945 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:24:51,946 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:24:51,946 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:24:51,949 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:24:51,951 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:24:51,952 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:24:51,952 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:24:51,974 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:51,974 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:24:51,975 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:51,977 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:24:51,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:24:51,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:24:51,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:24:51,978 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:24:51,978 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:24:51,979 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:24:51,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:24:51,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:24:51,979 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580091979-0
2018-07-14 20:24:51,994 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:24:51,995 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,006 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:24:52,009 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:24:52,009 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:24:52,009 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:24:52,011 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:24:52,037 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
```

```
2018-07-14 20:24:52,058 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local828777705_0049
2018-07-14 20:24:52,143 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:24:52,143 [Thread-1336] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:24:52,147 [Thread-1336] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:24:52,147 [Thread-1336] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:24:52,147 [Thread-1336] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:52,147 [Thread-1336] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:24:52,147 [Thread-1336] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:24:52,147 [Thread-1336] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:24:52,150 [Thread-1336] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:24:52,150 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local828777705_0049_m_000000_0
2018-07-14 20:24:52,155 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:52,157 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:24:52,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:24:52,160 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:24:52,173 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:24:52,173 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:24:52,173 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:24:52,173 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:24:52,173 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:24:52,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:24:52,176 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 20:24:52,176 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:24:52,179 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:24:52,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:24:52,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:24:52,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:24:52,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:24:52,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:24:52,196 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:24:52,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local828777705_0049_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:24:52,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:24:52,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local828777705_0049_m_000000_0' done.
2018-07-14 20:24:52,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local828777705_0049_m_000000_0
2018-07-14 20:24:52,307 [Thread-1336] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:24:52,308 [Thread-1336] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:24:52,309 [pool-134-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local828777705_0049_r_000000_0
2018-07-14 20:24:52,314 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:52,314 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@38bda658
2018-07-14 20:24:52,315 [pool-134-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:24:52,319 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local828777705_0049_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:24:52,322 [localfetcher#36] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#36 about to shuffle
output of map attempt_local828777705_0049_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:24:52,322 [localfetcher#36] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local828777705_0049_m_000000_0
2018-07-14 20:24:52,322 [localfetcher#36] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:24:52,323 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:24:52,323 [pool-134-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
```

```
2018-07-14 20:24:52,323 [pool-134-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:24:52,325 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:52,325 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:24:52,327 [pool-134-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:24:52,327 [pool-134-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:24:52,327 [pool-134-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:24:52,327 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:52,327 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:24:52,327 [pool-134-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:52,334 [pool-134-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:52,334 [pool-134-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:24:52,337 [pool-134-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:24:52,338 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local828777705_0049_r_000000_0 is done. And is in the process of committing
2018-07-14 20:24:52,339 [pool-134-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:52,339 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local828777705_0049_r_000000_0 is allowed to commit now
2018-07-14 20:24:52,341 [pool-134-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local828777705_0049_r_000000_0' to
file:/tmp/temp-1902294536/tmp-612188667/_temporary/0/task_local828777705_0049_r_000000
2018-07-14 20:24:52,341 [pool-134-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:24:52,341 [pool-134-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local828777705_0049_r_000000_0' done.
2018-07-14 20:24:52,341 [pool-134-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local828777705_0049_r_000000_0
2018-07-14 20:24:52,342 [Thread-1336] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:24:52,496 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local828777705_0049
2018-07-14 20:24:52,496 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:24:52,496 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:24:52,497 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
```

```
2018-07-14 20:24:52,498 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,498 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,498 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,502 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:24:52,502 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:24:52,502 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:24:52,502 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:24:52,503 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:24:52,508 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:24:52,509 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,510 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:52,511 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:24:52,516 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:24:52,525 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:24:52,525 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:24:52,525 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:24:52,549 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:24:52,563 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1399939460_0050
2018-07-14 20:24:52,641 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:24:52,642 [Thread-1365] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:24:52,648 [Thread-1365] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:24:52,648 [Thread-1365] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:24:52,648 [Thread-1365] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:52,648 [Thread-1365] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
```

```
Instead, use mapreduce.job.reduces
2018-07-14 20:24:52,648 [Thread-1365] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:24:52,648 [Thread-1365] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:24:52,650 [Thread-1365] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:24:52,650 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1399939460_0050_m_000000_0
2018-07-14 20:24:52,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:52,663 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:24:52,664 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-612188667/part-r-00000:0+62
2018-07-14 20:24:52,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:24:52,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:24:52,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:24:52,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:24:52,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:24:52,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:24:52,680 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:52,680 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:24:52,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R:
2018-07-14 20:24:52,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:24:52,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:24:52,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:24:52,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:24:52,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:24:52,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:24:52,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1399939460_0050_m_000000_0 is done. And
```

```
is in the process of committing
2018-07-14 20:24:52,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:24:52,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1399939460_0050_m_000000_0' done.
2018-07-14 20:24:52,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1399939460_0050_m_000000_0
2018-07-14 20:24:52,686 [Thread-1365] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:24:52,687 [Thread-1365] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:24:52,689 [pool-137-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1399939460_0050_r_000000_0
2018-07-14 20:24:52,696 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:24:52,697 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5370058f
2018-07-14 20:24:52,698 [pool-137-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:24:52,698 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1399939460_0050_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:24:52,699 [localfetcher#37] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#37 about to shuffle
output of map attempt_local1399939460_0050_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:24:52,700 [localfetcher#37] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1399939460_0050_m_000000_0
2018-07-14 20:24:52,700 [localfetcher#37] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:24:52,700 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:24:52,701 [pool-137-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:52,701 [pool-137-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:24:52,702 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:52,702 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:24:52,703 [pool-137-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:24:52,703 [pool-137-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:24:52,703 [pool-137-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:24:52,703 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:24:52,704 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:24:52,704 [pool-137-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
```

```
2018-07-14 20:24:52,718 [pool-137-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:24:52,718 [pool-137-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:24:52,718 [pool-137-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R:
2018-07-14 20:24:52,719 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1399939460_0050_r_000000_0 is done. And is in the process of committing
2018-07-14 20:24:52,720 [pool-137-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:24:52,720 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1399939460_0050_r_000000_0 is allowed to commit now
2018-07-14 20:24:52,722 [pool-137-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1399939460_0050_r_000000_0' to
file:/tmp/temp-1902294536/tmp327287845/_temporary/0/task_local1399939460_0050_r_000000
2018-07-14 20:24:52,722 [pool-137-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:24:52,722 [pool-137-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1399939460_0050_r_000000_0' done.
2018-07-14 20:24:52,722 [pool-137-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1399939460_0050_r_000000_0
2018-07-14 20:24:52,722 [Thread-1365] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:24:52,842 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1399939460_0050
2018-07-14 20:24:52,842 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 20:24:52,842 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[23,4] C:  R:
2018-07-14 20:24:52,844 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,844 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,844 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,846 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:24:52,846 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:24:51     2018-07-14 20:24:52
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1399939460_0050        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     q       GROUP_BY
file:/tmp/temp-1902294536/tmp327287845,
job_local828777705_0049 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp327287845"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local828777705_0049 ->      job_local1399939460_0050,
job_local1399939460_0050


2018-07-14 20:24:52,847 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,848 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,848 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,850 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,851 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,851 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:24:52,852 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:24:52,852 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:24:52,852 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:24:52,852 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:24:52,860 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:24:52,860 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt> w = foreach q generate MAX(p.sum) as max;
grunt> d = filter q BY sum==max;
2018-07-14 20:25:28,249 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 35, column 16> Invalid field projection. Projected field [sum] does not exist in
schema: group:chararray,p:bag{:tuple(EmpId:int,sum:double)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> d = filter q BY p.sum==max;
2018-07-14 20:25:34,003 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 35, column 23> Invalid field projection. Projected field [max] does not exist in
schema: group:chararray,p:bag{:tuple(EmpId:int,sum:double)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt> d = filter q BY p.sum==q.max;
2018-07-14 20:25:45,211 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 35, column 23> Invalid scalar projection: q
```

```
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt> dump q;
2018-07-14 20:25:50,355 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:25:50,372 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:25:50,373 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:25:50,373 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:25:50,373 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:25:50,374 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:25:50,375 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:25:50,377 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:25:50,377 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:25:50,387 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:25:50,387 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:25:50,387 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:50,388 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:25:50,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:25:50,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:25:50,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:25:50,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:25:50,388 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:25:50,390 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:25:50,390 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:25:50,390 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:25:50,390 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
```

```
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580150390-0
2018-07-14 20:25:50,398 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:25:50,399 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:50,407 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:25:50,408 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:25:50,408 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:25:50,408 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:25:50,408 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:25:50,428 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:25:50,442 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1742577904_0051
2018-07-14 20:25:50,553 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:25:50,553 [Thread-1394] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:25:50,558 [Thread-1394] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:25:50,558 [Thread-1394] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:25:50,558 [Thread-1394] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:25:50,558 [Thread-1394] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:25:50,558 [Thread-1394] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:25:50,558 [Thread-1394] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:25:50,561 [Thread-1394] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:25:50,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1742577904_0051_m_000000_0
2018-07-14 20:25:50,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:25:50,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------
```

```
2018-07-14 20:25:50,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:25:50,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:25:50,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:25:50,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:25:50,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:25:50,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:25:50,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:25:50,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:25:50,590 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:25:50,590 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:25:50,593 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:25:50,595 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:25:50,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:25:50,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:25:50,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:25:50,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:25:50,604 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:25:50,605 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1742577904_0051_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:25:50,606 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:25:50,606 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1742577904_0051_m_000000_0' done.
2018-07-14 20:25:50,606 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1742577904_0051_m_000000_0
2018-07-14 20:25:50,606 [Thread-1394] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:25:50,607 [Thread-1394] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:25:50,607 [pool-140-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1742577904_0051_r_000000_0
2018-07-14 20:25:50,612 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:25:50,612 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@495a4aa1
```

```
2018-07-14 20:25:50,614 [pool-140-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:25:50,614 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1742577904_0051_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:25:50,616 [localfetcher#38] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#38 about to shuffle
output of map attempt_local1742577904_0051_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:25:50,616 [localfetcher#38] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1742577904_0051_m_000000_0
2018-07-14 20:25:50,616 [localfetcher#38] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:25:50,620 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:25:50,620 [pool-140-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:25:50,620 [pool-140-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:25:50,621 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:25:50,621 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:25:50,622 [pool-140-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:25:50,622 [pool-140-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:25:50,622 [pool-140-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:25:50,622 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:25:50,622 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:25:50,623 [pool-140-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:25:50,630 [pool-140-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:25:50,630 [pool-140-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:25:50,634 [pool-140-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3] R: p[17,3]
2018-07-14 20:25:50,637 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1742577904_0051_r_000000_0 is done. And is in the process of committing
2018-07-14 20:25:50,638 [pool-140-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:25:50,638 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1742577904_0051_r_000000_0 is allowed to commit now
2018-07-14 20:25:50,640 [pool-140-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1742577904_0051_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1669290128/_temporary/0/task_local1742577904_0051_r_0000
```

```
00
2018-07-14 20:25:50,641 [pool-140-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:25:50,641 [pool-140-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1742577904_0051_r_000000_0' done.
2018-07-14 20:25:50,641 [pool-140-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1742577904_0051_r_000000_0
2018-07-14 20:25:50,641 [Thread-1394] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:25:50,755 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1742577904_0051
2018-07-14 20:25:50,755 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses,p,s
2018-07-14 20:25:50,755 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1],p[17,3],s[12,3] C: p[17,3],s[12,3]
R: p[17,3]
2018-07-14 20:25:50,758 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:25:50,762 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:50,765 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:50,768 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:50,776 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:25:50,776 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:25:50,777 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:25:50,777 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:25:50,781 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:25:50,823 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:25:50,827 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:50,829 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:25:50,830 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:25:50,837 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:25:50,845 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:25:50,845 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
```

```
2018-07-14 20:25:50,845 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:25:50,870 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:25:50,882 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local626675504_0052
2018-07-14 20:25:50,947 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:25:50,947 [Thread-1423] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:25:50,950 [Thread-1423] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:25:50,950 [Thread-1423] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:25:50,950 [Thread-1423] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:25:50,950 [Thread-1423] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:25:50,951 [Thread-1423] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:25:50,951 [Thread-1423] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:25:50,955 [Thread-1423] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:25:50,955 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local626675504_0052_m_000000_0
2018-07-14 20:25:50,959 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:25:50,959 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:25:50,960 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1669290128/part-r-00000:0+62
2018-07-14 20:25:50,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:25:50,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:25:50,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:25:50,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:25:50,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:25:50,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
```

```
2018-07-14 20:25:50,976 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:25:50,976 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:25:50,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R:
2018-07-14 20:25:50,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:25:50,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:25:50,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:25:50,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:25:50,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:25:50,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:25:50,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local626675504_0052_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:25:50,980 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:25:50,980 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local626675504_0052_m_000000_0' done.
2018-07-14 20:25:50,980 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local626675504_0052_m_000000_0
2018-07-14 20:25:50,980 [Thread-1423] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:25:50,981 [Thread-1423] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:25:50,981 [pool-143-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local626675504_0052_r_000000_0
2018-07-14 20:25:50,988 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:25:50,988 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7eafb1f6
2018-07-14 20:25:50,988 [pool-143-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:25:50,989 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local626675504_0052_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:25:50,990 [localfetcher#39] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#39 about to shuffle
output of map attempt_local626675504_0052_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:25:50,990 [localfetcher#39] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local626675504_0052_m_000000_0
2018-07-14 20:25:50,990 [localfetcher#39] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:25:50,991 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:25:50,992 [pool-143-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:25:50,992 [pool-143-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:25:50,992 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:25:50,992 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:25:50,993 [pool-143-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:25:50,993 [pool-143-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:25:50,993 [pool-143-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:25:50,993 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:25:50,993 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:25:50,993 [pool-143-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:25:50,999 [pool-143-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:25:50,999 [pool-143-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:25:50,999 [pool-143-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: q[23,4] C:  R:
2018-07-14 20:25:51,000 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local626675504_0052_r_000000_0 is done. And is in the process of committing
2018-07-14 20:25:51,001 [pool-143-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:25:51,001 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local626675504_0052_r_000000_0 is allowed to commit now
2018-07-14 20:25:51,003 [pool-143-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local626675504_0052_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1664617686/_temporary/0/task_local626675504_0052_r_00000
0
2018-07-14 20:25:51,003 [pool-143-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:25:51,003 [pool-143-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local626675504_0052_r_000000_0' done.
2018-07-14 20:25:51,003 [pool-143-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local626675504_0052_r_000000_0
2018-07-14 20:25:51,004 [Thread-1423] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:25:51,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local626675504_0052
2018-07-14 20:25:51,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases q
2018-07-14 20:25:51,149 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: q[23,4] C:  R:
2018-07-14 20:25:51,157 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,159 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,161 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,164 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:25:51,165 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:25:50     2018-07-14 20:25:51
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1742577904_0051        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     expenses,p,s    GROUP_BY,COMBINER
job_local626675504_0052 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     q       GROUP_BY        file:/tmp/temp-1902294536/tmp-1664617686,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1664617686"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1742577904_0051        ->      job_local626675504_0052,
job_local626675504_0052


2018-07-14 20:25:51,166 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,168 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,169 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,171 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,172 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,173 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:25:51,174 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:25:51,174 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:25:51,175 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:25:51,175 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
```

```
2018-07-14 20:25:51,182 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:25:51,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt> dump expenses;;
2018-07-14 20:26:07,479 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 20:26:07,493 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:07,493 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:07,493 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:26:07,493 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:26:07,494 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:26:07,496 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:26:07,496 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:26:07,504 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:07,504 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:07,505 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,505 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:26:07,506 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:26:07,506 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:26:07,506 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:26:07,506 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:26:07,506 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580167506-0
2018-07-14 20:26:07,511 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:26:07,512 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,522 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:26:07,523 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:26:07,524 [JobControl] INFO
```

```
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:26:07,524 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:26:07,524 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:26:07,545 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:26:07,557 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local726781730_0053
2018-07-14 20:26:07,622 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:26:07,622 [Thread-1452] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:26:07,625 [Thread-1452] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:26:07,625 [Thread-1452] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:26:07,625 [Thread-1452] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:07,625 [Thread-1452] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:26:07,625 [Thread-1452] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:26:07,627 [Thread-1452] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:26:07,627 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local726781730_0053_m_000000_0
2018-07-14 20:26:07,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:26:07,632 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:26:07,634 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:26:07,634 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:26:07,641 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:26:07,641 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:26:07,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: expenses[2,11],expenses[-1,-1]
C:  R:
```

```
2018-07-14 20:26:07,643 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:26:07,643 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local726781730_0053_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:26:07,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:26:07,647 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local726781730_0053_m_000000_0 is allowed to
commit now
2018-07-14 20:26:07,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local726781730_0053_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1671742367/_temporary/0/task_local726781730_0053_m_00000
0
2018-07-14 20:26:07,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:26:07,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local726781730_0053_m_000000_0' done.
2018-07-14 20:26:07,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local726781730_0053_m_000000_0
2018-07-14 20:26:07,648 [Thread-1452] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:26:07,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local726781730_0053
2018-07-14 20:26:07,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses
2018-07-14 20:26:07,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[2,11],expenses[-1,-1] C:  R:
2018-07-14 20:26:07,830 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,831 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,833 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,836 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:26:07,837 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:26:07     2018-07-14 20:26:07     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local726781730_0053 1       0       n/a     n/a     n/a     n/a     0       0
0       0       expenses        MAP_ONLY
file:/tmp/temp-1902294536/tmp-1671742367,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 9 records in: "file:/tmp/temp-1902294536/tmp-1671742367"
```

```
Counters:
Total records written : 9
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local726781730_0053


2018-07-14 20:26:07,839 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,841 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,842 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:07,843 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:26:07,844 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:07,844 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:07,845 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:26:07,857 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:26:07,857 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,200)
(102,100)
(110,400)
(114,200)
(119,200)
(105,100)
(101,100)
(104,300)
(102,400)
grunt> ex_sum= foreach ex_g GENERATE group as EmpId,SUM(expenses.Expense) as sum;
2018-07-14 20:26:33,443 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 35, column 16> Undefined alias: ex_g
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
>> (EmpId:Int,Expense:Int);
2018-07-14 20:26:38,780 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:38,780 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt> dump expenses;
2018-07-14 20:26:42,834 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: UNKNOWN
2018-07-14 20:26:42,848 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:42,849 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:42,849 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:26:42,850 [main] INFO
```

```
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:26:42,850 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:26:42,851 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:26:42,851 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:26:42,862 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:42,862 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:42,863 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:42,864 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:26:42,864 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:26:42,865 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:26:42,866 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:26:42,866 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:26:42,866 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580202866-0
2018-07-14 20:26:42,871 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:26:42,872 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:42,879 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:26:42,880 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:26:42,880 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:26:42,880 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:26:42,880 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:26:42,900 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:26:42,916 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local124474653_0054
2018-07-14 20:26:42,980 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:26:42,980 [Thread-1475] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:26:42,983 [Thread-1475] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:26:42,983 [Thread-1475] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:26:42,983 [Thread-1475] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:42,983 [Thread-1475] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:26:42,984 [Thread-1475] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:26:42,985 [Thread-1475] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:26:42,985 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local124474653_0054_m_000000_0
2018-07-14 20:26:42,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:26:42,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:26:42,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:26:42,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:26:43,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:26:43,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:26:43,010 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: expenses[35,11],expenses[-1,-1]
C:  R:
2018-07-14 20:26:43,010 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:26:43,010 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local124474653_0054_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:26:43,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:26:43,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local124474653_0054_m_000000_0 is allowed to
commit now
2018-07-14 20:26:43,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local124474653_0054_m_000000_0' to
file:/tmp/temp-1902294536/tmp-576776523/_temporary/0/task_local124474653_0054_m_000000
2018-07-14 20:26:43,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
```

```
2018-07-14 20:26:43,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local124474653_0054_m_000000_0' done.
2018-07-14 20:26:43,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local124474653_0054_m_000000_0
2018-07-14 20:26:43,013 [Thread-1475] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:26:43,181 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local124474653_0054
2018-07-14 20:26:43,181 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases expenses
2018-07-14 20:26:43,181 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1] C:  R:
2018-07-14 20:26:43,184 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:43,185 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:43,185 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:43,187 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:26:43,187 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:26:42     2018-07-14 20:26:43     UNKNOWN

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local124474653_0054 1       0       n/a     n/a     n/a     n/a     0       0
0       0       expenses        MAP_ONLY
file:/tmp/temp-1902294536/tmp-576776523,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 9 records in: "file:/tmp/temp-1902294536/tmp-576776523"

Counters:
Total records written : 9
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local124474653_0054


2018-07-14 20:26:43,188 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:43,189 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:43,189 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:43,192 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:26:43,192 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:43,193 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:43,193 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:26:43,205 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:26:43,205 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,200)
(102,100)
(110,400)
(114,200)
(119,200)
(105,100)
(101,100)
(104,300)
(102,400)
grunt> ex_g= group expenses by EmpId;
grunt> dump ex_g;
2018-07-14 20:26:55,371 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:26:55,388 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:55,388 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:55,388 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:26:55,388 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:26:55,390 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:26:55,391 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:26:55,391 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:26:55,401 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:55,402 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:55,402 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,403 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:26:55,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:26:55,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:26:55,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:26:55,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:26:55,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:26:55,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:26:55,404 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:26:55,404 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:26:55,404 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580215404-0
2018-07-14 20:26:55,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:26:55,410 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,419 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:26:55,421 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:26:55,421 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:26:55,421 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:26:55,421 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:26:55,444 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:26:55,461 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1024379841_0055
2018-07-14 20:26:55,525 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:26:55,526 [Thread-1498] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:26:55,529 [Thread-1498] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:26:55,529 [Thread-1498] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:26:55,529 [Thread-1498] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:55,529 [Thread-1498] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:26:55,529 [Thread-1498] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 20:26:55,530 [Thread-1498] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:26:55,532 [Thread-1498] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:26:55,533 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1024379841_0055_m_000000_0
2018-07-14 20:26:55,537 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:26:55,538 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:26:55,539 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:26:55,539 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:26:55,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:26:55,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:26:55,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:26:55,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:26:55,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:26:55,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:26:55,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:26:55,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:26:55,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:26:55,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:26:55,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:26:55,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:26:55,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 20:26:55,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:26:55,558 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:26:55,559 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1024379841_0055_m_000000_0 is done. And
```

```
is in the process of committing
2018-07-14 20:26:55,560 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:26:55,560 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1024379841_0055_m_000000_0' done.
2018-07-14 20:26:55,560 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1024379841_0055_m_000000_0
2018-07-14 20:26:55,560 [Thread-1498] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:26:55,562 [Thread-1498] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:26:55,562 [pool-150-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1024379841_0055_r_000000_0
2018-07-14 20:26:55,568 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:26:55,568 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@34c62521
2018-07-14 20:26:55,569 [pool-150-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:26:55,569 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1024379841_0055_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:26:55,570 [localfetcher#40] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#40 about to shuffle
output of map attempt_local1024379841_0055_m_000000_0 decomp: 125 len: 129 to MEMORY
2018-07-14 20:26:55,570 [localfetcher#40] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local1024379841_0055_m_000000_0
2018-07-14 20:26:55,570 [localfetcher#40] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 20:26:55,572 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:26:55,573 [pool-150-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:26:55,573 [pool-150-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:26:55,573 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:26:55,573 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 20:26:55,576 [pool-150-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:26:55,576 [pool-150-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 20:26:55,576 [pool-150-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:26:55,576 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:26:55,576 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 20:26:55,577 [pool-150-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
```

```
2018-07-14 20:26:55,582 [pool-150-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:26:55,582 [pool-150-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:26:55,583 [pool-150-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:26:55,584 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1024379841_0055_r_000000_0 is done. And is in the process of committing
2018-07-14 20:26:55,585 [pool-150-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:26:55,585 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1024379841_0055_r_000000_0 is allowed to commit now
2018-07-14 20:26:55,586 [pool-150-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1024379841_0055_r_000000_0' to
file:/tmp/temp-1902294536/tmp1150971921/_temporary/0/task_local1024379841_0055_r_00000
0
2018-07-14 20:26:55,586 [pool-150-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:26:55,586 [pool-150-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1024379841_0055_r_000000_0' done.
2018-07-14 20:26:55,586 [pool-150-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1024379841_0055_r_000000_0
2018-07-14 20:26:55,586 [Thread-1498] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:26:55,729 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1024379841_0055
2018-07-14 20:26:55,729 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,expenses
2018-07-14 20:26:55,729 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:26:55,735 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,736 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,738 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,740 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:26:55,741 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:26:55     2018-07-14 20:26:55
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1024379841_0055        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,expenses   GROUP_BY
file:/tmp/temp-1902294536/tmp1150971921,
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp1150971921"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1024379841_0055


2018-07-14 20:26:55,742 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,744 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,745 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:26:55,749 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:26:55,750 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:26:55,750 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:26:55,750 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:26:55,766 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:26:55,766 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> ex_sum= foreach ex_g GENERATE group as EmpId,SUM(expenses.Expense) as sum;
grunt> dump ex_sum;
2018-07-14 20:27:07,975 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:27:07,991 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:27:07,991 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:27:07,991 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:27:07,992 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:27:07,993 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
```

```
concatenation threshold: 100 optimistic? false
2018-07-14 20:27:07,993 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:27:07,995 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:27:07,995 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:27:08,003 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:27:08,004 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:27:08,004 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,005 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:27:08,006 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:27:08,006 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:27:08,006 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:27:08,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:27:08,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:27:08,008 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:27:08,009 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:27:08,009 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:27:08,009 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580228008-0
2018-07-14 20:27:08,017 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:27:08,018 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,026 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:27:08,027 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:27:08,028 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:27:08,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:27:08,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
```

```
2018-07-14 20:27:08,052 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:27:08,065 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local966116924_0056
2018-07-14 20:27:08,184 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:27:08,184 [Thread-1527] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:27:08,188 [Thread-1527] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:27:08,188 [Thread-1527] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:27:08,188 [Thread-1527] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:27:08,188 [Thread-1527] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:27:08,188 [Thread-1527] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:27:08,188 [Thread-1527] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:27:08,190 [Thread-1527] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:27:08,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local966116924_0056_m_000000_0
2018-07-14 20:27:08,195 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:27:08,196 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:27:08,198 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:27:08,198 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:27:08,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:27:08,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:27:08,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:27:08,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:27:08,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:27:08,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:27:08,220 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:27:08,220 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:27:08,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:27:08,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:27:08,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:27:08,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:27:08,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:27:08,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:27:08,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:27:08,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local966116924_0056_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:27:08,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:27:08,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local966116924_0056_m_000000_0' done.
2018-07-14 20:27:08,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local966116924_0056_m_000000_0
2018-07-14 20:27:08,234 [Thread-1527] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:27:08,234 [Thread-1527] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:27:08,235 [pool-153-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local966116924_0056_r_000000_0
2018-07-14 20:27:08,243 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:27:08,243 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6247830e
2018-07-14 20:27:08,244 [pool-153-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:27:08,245 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local966116924_0056_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:27:08,251 [localfetcher#41] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#41 about to shuffle
output of map attempt_local966116924_0056_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:27:08,252 [localfetcher#41] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local966116924_0056_m_000000_0
2018-07-14 20:27:08,252 [localfetcher#41] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:27:08,253 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 20:27:08,254 [pool-153-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:27:08,255 [pool-153-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:27:08,255 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:27:08,255 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:27:08,257 [pool-153-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:27:08,257 [pool-153-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:27:08,257 [pool-153-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:27:08,257 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:27:08,257 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:27:08,257 [pool-153-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:27:08,265 [pool-153-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:27:08,265 [pool-153-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:27:08,267 [pool-153-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:27:08,268 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local966116924_0056_r_000000_0 is done. And is in the process of committing
2018-07-14 20:27:08,270 [pool-153-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:27:08,270 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local966116924_0056_r_000000_0 is allowed to commit now
2018-07-14 20:27:08,272 [pool-153-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local966116924_0056_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1859908239/_temporary/0/task_local966116924_0056_r_00000
0
2018-07-14 20:27:08,272 [pool-153-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:27:08,272 [pool-153-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local966116924_0056_r_000000_0' done.
2018-07-14 20:27:08,272 [pool-153-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local966116924_0056_r_000000_0
2018-07-14 20:27:08,272 [Thread-1527] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:27:08,386 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local966116924_0056
2018-07-14 20:27:08,386 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:27:08,386 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:27:08,394 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,398 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,400 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,407 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:27:08,409 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:27:08     2018-07-14 20:27:08
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local966116924_0056 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-1859908239,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1859908239"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local966116924_0056


2018-07-14 20:27:08,411 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,416 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,418 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:27:08,437 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:27:08,439 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:27:08,441 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:27:08,442 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:27:08,459 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:27:08,459 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> describe ex_sum;
ex_sum: {EmpId: int,sum: long}
grunt> ex_m = foreach ex_sum generate MAX(sum);
2018-07-14 20:28:02,873 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1045:
<line 39, column 31> Could not infer the matching function for org.apache.pig.builtin.MAX
as multiple or none of them fit. Please use an explicit cast.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ex_m = foreach ex_sum generate MAX(ex_sum.sum);
2018-07-14 20:28:09,136 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1045:
<line 39, column 31> Could not infer the matching function for org.apache.pig.builtin.MAX
as multiple or none of them fit. Please use an explicit cast.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt> describe ex_sum;
ex_sum: {EmpId: int,sum: long}
grunt> ex_m = foreach ex_sum generate EmpId,MAX(ex_sum.sum);
2018-07-14 20:28:30,408 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1045:
<line 39, column 37> Could not infer the matching function for org.apache.pig.builtin.MAX
as multiple or none of them fit. Please use an explicit cast.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt> dump ex_sum;
2018-07-14 20:28:48,070 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:28:48,088 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:28:48,088 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:28:48,088 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:28:48,088 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:28:48,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:28:48,090 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:28:48,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:28:48,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:28:48,100 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:28:48,101 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:28:48,101 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,101 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:28:48,102 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:28:48,102 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:28:48,102 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:28:48,102 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:28:48,102 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:28:48,103 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:28:48,103 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:28:48,103 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:28:48,103 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580328103-0
2018-07-14 20:28:48,110 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:28:48,111 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,119 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:28:48,121 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:28:48,121 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:28:48,121 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:28:48,121 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:28:48,143 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:28:48,157 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1162989812_0057
2018-07-14 20:28:48,218 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:28:48,218 [Thread-1556] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:28:48,221 [Thread-1556] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:28:48,221 [Thread-1556] INFO
org.apache.hadoop.conf.Configuration.deprecation -
```

```
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:28:48,222 [Thread-1556] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:28:48,222 [Thread-1556] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:28:48,222 [Thread-1556] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:28:48,222 [Thread-1556] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:28:48,223 [Thread-1556] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:28:48,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1162989812_0057_m_000000_0
2018-07-14 20:28:48,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:28:48,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:28:48,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:28:48,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:28:48,245 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:28:48,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:28:48,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:28:48,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:28:48,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:28:48,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:28:48,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:28:48,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:28:48,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:28:48,249 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:28:48,250 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:28:48,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:28:48,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:28:48,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:28:48,257 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:28:48,258 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1162989812_0057_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:28:48,259 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:28:48,259 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1162989812_0057_m_000000_0' done.
2018-07-14 20:28:48,259 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1162989812_0057_m_000000_0
2018-07-14 20:28:48,259 [Thread-1556] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:28:48,259 [Thread-1556] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:28:48,260 [pool-156-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1162989812_0057_r_000000_0
2018-07-14 20:28:48,266 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:28:48,266 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@76e4975f
2018-07-14 20:28:48,267 [pool-156-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:28:48,268 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1162989812_0057_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:28:48,270 [localfetcher#42] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#42 about to shuffle
output of map attempt_local1162989812_0057_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:28:48,270 [localfetcher#42] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1162989812_0057_m_000000_0
2018-07-14 20:28:48,270 [localfetcher#42] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:28:48,270 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:28:48,271 [pool-156-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:28:48,271 [pool-156-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:28:48,272 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:28:48,272 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:28:48,272 [pool-156-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
```

```
2018-07-14 20:28:48,273 [pool-156-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:28:48,273 [pool-156-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:28:48,273 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:28:48,273 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:28:48,273 [pool-156-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:28:48,280 [pool-156-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:28:48,280 [pool-156-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:28:48,283 [pool-156-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:28:48,283 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1162989812_0057_r_000000_0 is done. And is in the process of committing
2018-07-14 20:28:48,284 [pool-156-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:28:48,284 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1162989812_0057_r_000000_0 is allowed to commit now
2018-07-14 20:28:48,285 [pool-156-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1162989812_0057_r_000000_0' to
file:/tmp/temp-1902294536/tmp965341689/_temporary/0/task_local1162989812_0057_r_000000
2018-07-14 20:28:48,286 [pool-156-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:28:48,286 [pool-156-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1162989812_0057_r_000000_0' done.
2018-07-14 20:28:48,286 [pool-156-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1162989812_0057_r_000000_0
2018-07-14 20:28:48,286 [Thread-1556] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:28:48,421 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1162989812_0057
2018-07-14 20:28:48,421 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:28:48,421 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:28:48,432 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,436 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,438 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,444 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:28:48,445 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:
```

```
HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:28:48     2018-07-14 20:28:48
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1162989812_0057        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp965341689,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp965341689"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1162989812_0057


2018-07-14 20:28:48,448 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,449 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,450 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:28:48,452 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:28:48,452 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:28:48,452 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:28:48,453 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:28:48,471 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:28:48,471 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> ex_m = group all ex_suml
>> ex_m = group all ex_sum;
2018-07-14 20:29:08,575 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
39, column 13>  Syntax error, unexpected symbol at or near 'all'
Details at logfile: /home/acadgild/pig_1531574068461.log
```

```
grunt> ex_m = group all ex_sum;
2018-07-14 20:29:11,776 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
39, column 13>  Syntax error, unexpected symbol at or near 'all'
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ex_m = group  ex_sum all;
grunt> dump ex_m;
2018-07-14 20:29:27,719 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:29:27,732 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:29:27,733 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:29:27,733 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:29:27,733 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:29:27,735 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:29:27,736 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:29:27,737 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:29:27,737 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:29:27,748 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:29:27,748 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:29:27,749 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:27,749 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:29:27,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:29:27,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:29:27,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:29:27,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:29:27,750 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:29:27,751 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:29:27,751 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:29:27,751 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
```

```
process to move generated code to distributed cacche
2018-07-14 20:29:27,751 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580367751-0
2018-07-14 20:29:27,764 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:29:27,765 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:27,777 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:29:27,778 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:29:27,778 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:29:27,778 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:29:27,778 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:29:27,802 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:29:27,816 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1155474936_0058
2018-07-14 20:29:27,877 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:29:27,878 [Thread-1585] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:29:27,881 [Thread-1585] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:29:27,881 [Thread-1585] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:29:27,881 [Thread-1585] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:29:27,881 [Thread-1585] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:29:27,881 [Thread-1585] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:29:27,881 [Thread-1585] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:29:27,883 [Thread-1585] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:29:27,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1155474936_0058_m_000000_0
2018-07-14 20:29:27,892 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:29:27,894 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:
```

```
------------------------

2018-07-14 20:29:27,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:29:27,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:29:27,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:29:27,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:29:27,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:29:27,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:29:27,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:29:27,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:29:27,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:29:27,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:29:27,917 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:29:27,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:29:27,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:29:27,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:29:27,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:29:27,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:29:27,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:29:27,927 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1155474936_0058_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:29:27,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:29:27,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1155474936_0058_m_000000_0' done.
2018-07-14 20:29:27,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1155474936_0058_m_000000_0
2018-07-14 20:29:27,928 [Thread-1585] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:29:27,928 [Thread-1585] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:29:27,929 [pool-159-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1155474936_0058_r_000000_0
2018-07-14 20:29:27,933 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
```

```
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:29:27,933 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@26832738
2018-07-14 20:29:27,934 [pool-159-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:29:27,938 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1155474936_0058_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:29:27,941 [localfetcher#43] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#43 about to shuffle
output of map attempt_local1155474936_0058_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:29:27,941 [localfetcher#43] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1155474936_0058_m_000000_0
2018-07-14 20:29:27,941 [localfetcher#43] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:29:27,942 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:29:27,943 [pool-159-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:29:27,943 [pool-159-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:29:27,944 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:29:27,944 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:29:27,945 [pool-159-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:29:27,945 [pool-159-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:29:27,945 [pool-159-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:29:27,945 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:29:27,946 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:29:27,946 [pool-159-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:29:27,959 [pool-159-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:29:27,959 [pool-159-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:29:27,961 [pool-159-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:29:27,961 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1155474936_0058_r_000000_0 is done. And is in the process of committing
2018-07-14 20:29:27,963 [pool-159-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:29:27,963 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1155474936_0058_r_000000_0 is allowed to commit now
```

```
2018-07-14 20:29:27,964 [pool-159-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1155474936_0058_r_000000_0' to
file:/tmp/temp-1902294536/tmp1043280299/_temporary/0/task_local1155474936_0058_r_00000
0
2018-07-14 20:29:27,965 [pool-159-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:29:27,965 [pool-159-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1155474936_0058_r_000000_0' done.
2018-07-14 20:29:27,965 [pool-159-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1155474936_0058_r_000000_0
2018-07-14 20:29:27,965 [Thread-1585] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:29:28,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1155474936_0058
2018-07-14 20:29:28,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:29:28,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:29:28,183 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:29:28,184 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,186 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,186 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,191 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:29:28,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:29:28,192 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:29:28,192 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:29:28,193 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:29:28,216 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:29:28,219 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,222 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:29:28,222 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:29:28,230 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:29:28,240 [JobControl] INFO
```

```
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:29:28,240 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:29:28,241 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:29:28,260 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:29:28,274 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1579110820_0059
2018-07-14 20:29:28,338 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:29:28,339 [Thread-1614] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:29:28,342 [Thread-1614] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:29:28,342 [Thread-1614] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:29:28,342 [Thread-1614] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:29:28,342 [Thread-1614] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:29:28,342 [Thread-1614] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:29:28,342 [Thread-1614] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:29:28,343 [Thread-1614] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:29:28,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1579110820_0059_m_000000_0
2018-07-14 20:29:28,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:29:28,348 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:29:28,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1043280299/part-r-00000:0+62
2018-07-14 20:29:28,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:29:28,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:29:28,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:29:28,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:29:28,366 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:29:28,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:29:28,367 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:29:28,367 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:29:28,367 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
2018-07-14 20:29:28,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:29:28,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:29:28,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:29:28,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:29:28,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:29:28,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:29:28,375 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1579110820_0059_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:29:28,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:29:28,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1579110820_0059_m_000000_0' done.
2018-07-14 20:29:28,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1579110820_0059_m_000000_0
2018-07-14 20:29:28,376 [Thread-1614] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:29:28,376 [Thread-1614] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:29:28,377 [pool-162-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1579110820_0059_r_000000_0
2018-07-14 20:29:28,382 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:29:28,383 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@177be2ec
2018-07-14 20:29:28,383 [pool-162-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:29:28,386 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1579110820_0059_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:29:28,388 [localfetcher#44] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#44 about to shuffle
output of map attempt_local1579110820_0059_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:29:28,388 [localfetcher#44] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1579110820_0059_m_000000_0
2018-07-14 20:29:28,388 [localfetcher#44] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
```

```
2018-07-14 20:29:28,388 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:29:28,389 [pool-162-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:29:28,389 [pool-162-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:29:28,389 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:29:28,389 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:29:28,390 [pool-162-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:29:28,390 [pool-162-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:29:28,390 [pool-162-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:29:28,390 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:29:28,390 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:29:28,390 [pool-162-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:29:28,398 [pool-162-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:29:28,398 [pool-162-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:29:28,398 [pool-162-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
2018-07-14 20:29:28,399 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1579110820_0059_r_000000_0 is done. And is in the process of committing
2018-07-14 20:29:28,399 [pool-162-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:29:28,399 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1579110820_0059_r_000000_0 is allowed to commit now
2018-07-14 20:29:28,400 [pool-162-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1579110820_0059_r_000000_0' to
file:/tmp/temp-1902294536/tmp2051572829/_temporary/0/task_local1579110820_0059_r_00000
0
2018-07-14 20:29:28,401 [pool-162-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:29:28,401 [pool-162-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1579110820_0059_r_000000_0' done.
2018-07-14 20:29:28,401 [pool-162-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1579110820_0059_r_000000_0
2018-07-14 20:29:28,401 [Thread-1614] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:29:28,540 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1579110820_0059
2018-07-14 20:29:28,540 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m
2018-07-14 20:29:28,540 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
detailed locations: M: ex_m[39,7] C:  R:
2018-07-14 20:29:28,549 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,551 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,552 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,557 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:29:28,558 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:29:27     2018-07-14 20:29:28
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1155474936_0058        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1579110820_0059        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m    GROUP_BY
file:/tmp/temp-1902294536/tmp2051572829,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp2051572829"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1155474936_0058        ->      job_local1579110820_0059,
job_local1579110820_0059


2018-07-14 20:29:28,558 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,560 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,560 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,565 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,566 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,566 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:29:28,567 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:29:28,568 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:29:28,568 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:29:28,568 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:29:28,576 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:29:28,576 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt>
grunt>
grunt> ex_m = foreach ex_m generate group, MAX(sum);
2018-07-14 20:30:12,853 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 40, column 40> Invalid field projection. Projected field [sum] does not exist in
schema: group:chararray,ex_sum:bag{:tuple(EmpId:int,sum:double)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe ex_m;
ex_m: {group: chararray,ex_sum: {(EmpId: int,sum: long)}}
grunt> ex_max = foreach ex_m generate group, MAX(ex_m.sum);
2018-07-14 20:30:33,898 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 40, column 42> Invalid scalar projection: ex_m
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ex_max = foreach ex_m generate group, MAX(sum);
2018-07-14 20:30:47,078 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 40, column 42> Invalid field projection. Projected field [sum] does not exist in
schema: group:chararray,ex_sum:bag{:tuple(EmpId:int,sum:double)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt> describe ex_m;
ex_m: {group: chararray,ex_sum: {(EmpId: int,sum: long)}}
grunt>
grunt> ex_max = foreach ex_m generate group, MAX(ex_sum.sum);
grunt> dump ex_max;
2018-07-14 20:31:25,798 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:31:25,813 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:25,814 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:25,814 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:31:25,814 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:31:25,816 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:31:25,817 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:31:25,817 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:31:25,818 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
```

```
plan size before optimization: 2
2018-07-14 20:31:25,818 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:31:25,827 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:25,827 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:25,828 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:25,829 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:31:25,829 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:31:25,829 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:31:25,829 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:31:25,829 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:31:25,829 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:31:25,830 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:31:25,831 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:31:25,831 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:31:25,831 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580485830-0
2018-07-14 20:31:25,840 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:31:25,841 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:25,848 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:31:25,849 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:31:25,849 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:25,849 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:31:25,850 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:31:25,872 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:31:25,886 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local467039160_0060
2018-07-14 20:31:25,958 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
```

```
2018-07-14 20:31:25,962 [Thread-1643] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:31:25,965 [Thread-1643] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:31:25,965 [Thread-1643] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:31:25,965 [Thread-1643] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:25,965 [Thread-1643] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:31:25,965 [Thread-1643] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:25,966 [Thread-1643] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:31:25,967 [Thread-1643] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:31:25,967 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local467039160_0060_m_000000_0
2018-07-14 20:31:25,971 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:25,972 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:31:25,974 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:31:25,974 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:31:25,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:31:25,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:31:25,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:31:25,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:31:25,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:31:25,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:31:25,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:25,992 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:31:25,995 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:31:25,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:31:25,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:31:25,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:31:25,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:31:25,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:31:26,005 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:31:26,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local467039160_0060_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:31:26,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:31:26,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local467039160_0060_m_000000_0' done.
2018-07-14 20:31:26,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local467039160_0060_m_000000_0
2018-07-14 20:31:26,008 [Thread-1643] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:31:26,008 [Thread-1643] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:31:26,009 [pool-165-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local467039160_0060_r_000000_0
2018-07-14 20:31:26,016 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:26,016 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@438a8d7f
2018-07-14 20:31:26,017 [pool-165-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:31:26,023 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local467039160_0060_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:31:26,027 [localfetcher#45] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#45 about to shuffle
output of map attempt_local467039160_0060_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:31:26,027 [localfetcher#45] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local467039160_0060_m_000000_0
2018-07-14 20:31:26,027 [localfetcher#45] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:31:26,028 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:31:26,029 [pool-165-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:26,029 [pool-165-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
```

```
2018-07-14 20:31:26,029 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:26,029 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:31:26,030 [pool-165-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:31:26,030 [pool-165-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:31:26,030 [pool-165-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:31:26,030 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:26,030 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:31:26,030 [pool-165-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:26,038 [pool-165-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:26,038 [pool-165-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:31:26,040 [pool-165-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:31:26,041 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local467039160_0060_r_000000_0 is done. And is in the process of committing
2018-07-14 20:31:26,042 [pool-165-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:26,042 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local467039160_0060_r_000000_0 is allowed to commit now
2018-07-14 20:31:26,044 [pool-165-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local467039160_0060_r_000000_0' to
file:/tmp/temp-1902294536/tmp-879023785/_temporary/0/task_local467039160_0060_r_000000
2018-07-14 20:31:26,044 [pool-165-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:31:26,044 [pool-165-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local467039160_0060_r_000000_0' done.
2018-07-14 20:31:26,044 [pool-165-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local467039160_0060_r_000000_0
2018-07-14 20:31:26,045 [Thread-1643] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:31:26,165 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local467039160_0060
2018-07-14 20:31:26,165 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:31:26,165 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:31:26,168 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:31:26,173 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:31:26,177 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,179 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,184 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:31:26,184 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:31:26,184 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:31:26,184 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:31:26,186 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:31:26,198 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:31:26,199 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,201 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:26,201 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:26,207 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:31:26,216 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:26,216 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:31:26,216 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:31:26,236 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:31:26,252 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local277576839_0061
2018-07-14 20:31:26,318 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:31:26,319 [Thread-1672] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:31:26,323 [Thread-1672] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:31:26,323 [Thread-1672] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:31:26,323 [Thread-1672] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:26,323 [Thread-1672] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:31:26,323 [Thread-1672] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:26,323 [Thread-1672] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:31:26,326 [Thread-1672] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:31:26,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local277576839_0061_m_000000_0
2018-07-14 20:31:26,330 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:26,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:31:26,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-879023785/part-r-00000:0+62
2018-07-14 20:31:26,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:31:26,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:31:26,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:31:26,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:31:26,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:31:26,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:31:26,348 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:26,348 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:31:26,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[40,9],ex_m[39,7]
C: ex_max[40,9],ex_m[39,7] R: ex_max[40,9]
2018-07-14 20:31:26,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:31:26,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:31:26,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:31:26,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:31:26,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:31:26,357 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:31:26,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local277576839_0061_m_000000_0 is done. And
is in the process of committing
```

```
2018-07-14 20:31:26,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:31:26,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local277576839_0061_m_000000_0' done.
2018-07-14 20:31:26,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local277576839_0061_m_000000_0
2018-07-14 20:31:26,360 [Thread-1672] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:31:26,360 [Thread-1672] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:31:26,361 [pool-168-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local277576839_0061_r_000000_0
2018-07-14 20:31:26,366 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:26,366 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4bb4e587
2018-07-14 20:31:26,367 [pool-168-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:31:26,368 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local277576839_0061_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:31:26,369 [localfetcher#46] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#46 about to shuffle
output of map attempt_local277576839_0061_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:31:26,369 [localfetcher#46] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local277576839_0061_m_000000_0
2018-07-14 20:31:26,369 [localfetcher#46] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:31:26,371 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:31:26,372 [pool-168-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:26,372 [pool-168-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:31:26,372 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:26,372 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:31:26,374 [pool-168-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:31:26,374 [pool-168-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:31:26,374 [pool-168-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:31:26,374 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:26,374 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:31:26,374 [pool-168-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:26,380 [pool-168-thread-1] INFO
```

```
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:26,380 [pool-168-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:31:26,382 [pool-168-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[40,9],ex_m[39,7]
C: ex_max[40,9],ex_m[39,7] R: ex_max[40,9]
2018-07-14 20:31:26,382 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local277576839_0061_r_000000_0 is done. And is in the process of committing
2018-07-14 20:31:26,383 [pool-168-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:26,383 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local277576839_0061_r_000000_0 is allowed to commit now
2018-07-14 20:31:26,384 [pool-168-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local277576839_0061_r_000000_0' to
file:/tmp/temp-1902294536/tmp1980691737/_temporary/0/task_local277576839_0061_r_000000
2018-07-14 20:31:26,386 [pool-168-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:31:26,386 [pool-168-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local277576839_0061_r_000000_0' done.
2018-07-14 20:31:26,386 [pool-168-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local277576839_0061_r_000000_0
2018-07-14 20:31:26,386 [Thread-1672] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:31:26,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local277576839_0061
2018-07-14 20:31:26,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:31:26,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[40,9],ex_m[39,7] C: ex_max[40,9],ex_m[39,7] R: ex_max[40,9]
2018-07-14 20:31:26,529 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,532 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,533 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,545 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:31:26,549 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:31:25     2018-07-14 20:31:26
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local277576839_0061 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp1980691737,
job_local467039160_0060 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp1980691737"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local467039160_0060 ->      job_local277576839_0061,
job_local277576839_0061


2018-07-14 20:31:26,552 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,553 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,555 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,557 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,558 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,559 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:26,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:31:26,561 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:26,562 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:26,562 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:31:26,572 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:26,572 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,500)
grunt> ex_max = foreach ex_m generate group, MAX(ex_sum.sum) as max;
grunt> dump ex_max;
2018-07-14 20:31:40,378 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:31:40,395 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:40,395 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:40,395 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:31:40,395 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:31:40,397 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:31:40,399 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:31:40,399 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:31:40,400 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:31:40,400 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:31:40,410 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:40,410 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:40,411 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:40,412 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:31:40,413 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:31:40,413 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:31:40,413 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:31:40,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:31:40,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:31:40,416 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:31:40,416 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:31:40,416 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:31:40,416 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580500416-0
2018-07-14 20:31:40,426 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:31:40,427 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:40,435 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:31:40,439 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:31:40,440 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:40,440 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
2018-07-14 20:31:40,440 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:31:40,461 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:31:40,476 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1004043501_0062
2018-07-14 20:31:40,553 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:31:40,554 [Thread-1701] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:31:40,558 [Thread-1701] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:31:40,558 [Thread-1701] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:31:40,558 [Thread-1701] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:40,558 [Thread-1701] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:31:40,558 [Thread-1701] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:40,558 [Thread-1701] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:31:40,560 [Thread-1701] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:31:40,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1004043501_0062_m_000000_0
2018-07-14 20:31:40,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:40,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:31:40,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:31:40,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:31:40,585 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:31:40,585 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:31:40,585 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:31:40,586 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:31:40,586 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
```

```
2018-07-14 20:31:40,586 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:31:40,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:40,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:31:40,590 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:31:40,592 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:31:40,592 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:31:40,592 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:31:40,592 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:31:40,592 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:31:40,601 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:31:40,603 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1004043501_0062_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:31:40,604 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:31:40,604 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1004043501_0062_m_000000_0' done.
2018-07-14 20:31:40,604 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1004043501_0062_m_000000_0
2018-07-14 20:31:40,605 [Thread-1701] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:31:40,605 [Thread-1701] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:31:40,605 [pool-171-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1004043501_0062_r_000000_0
2018-07-14 20:31:40,613 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:40,613 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@10099108
2018-07-14 20:31:40,614 [pool-171-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:31:40,615 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1004043501_0062_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:31:40,616 [localfetcher#47] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#47 about to shuffle
output of map attempt_local1004043501_0062_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:31:40,616 [localfetcher#47] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1004043501_0062_m_000000_0
2018-07-14 20:31:40,616 [localfetcher#47] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:31:40,617 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:31:40,618 [pool-171-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:40,618 [pool-171-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:31:40,618 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:40,619 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:31:40,620 [pool-171-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:31:40,620 [pool-171-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:31:40,620 [pool-171-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:31:40,621 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:40,621 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:31:40,621 [pool-171-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:40,627 [pool-171-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:40,628 [pool-171-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:31:40,629 [pool-171-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:31:40,630 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1004043501_0062_r_000000_0 is done. And is in the process of committing
2018-07-14 20:31:40,631 [pool-171-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:40,631 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1004043501_0062_r_000000_0 is allowed to commit now
2018-07-14 20:31:40,632 [pool-171-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1004043501_0062_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1639748848/_temporary/0/task_local1004043501_0062_r_0000
00
2018-07-14 20:31:40,632 [pool-171-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:31:40,632 [pool-171-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1004043501_0062_r_000000_0' done.
2018-07-14 20:31:40,632 [pool-171-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1004043501_0062_r_000000_0
2018-07-14 20:31:40,633 [Thread-1701] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:31:40,761 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1004043501_0062
2018-07-14 20:31:40,761 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:31:40,761 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:31:40,764 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:31:40,772 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:40,777 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:40,780 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:40,788 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:31:40,790 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:31:40,791 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:31:40,791 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:31:40,796 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:31:40,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:31:40,826 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:40,830 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:40,830 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:40,838 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:31:40,846 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:40,846 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:31:40,846 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:31:40,865 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:31:40,878 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1613517257_0063
2018-07-14 20:31:40,944 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:31:40,945 [Thread-1730] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:31:40,950 [Thread-1730] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
```

```
2018-07-14 20:31:40,950 [Thread-1730] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:31:40,950 [Thread-1730] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:40,950 [Thread-1730] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:31:40,950 [Thread-1730] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:40,950 [Thread-1730] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:31:40,953 [Thread-1730] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:31:40,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1613517257_0063_m_000000_0
2018-07-14 20:31:40,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:40,958 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:31:40,958 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1639748848/part-r-00000:0+62
2018-07-14 20:31:40,973 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:31:40,973 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:31:40,973 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:31:40,973 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:31:40,973 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:31:40,974 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:31:40,975 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:40,975 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:31:40,976 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[41,9],ex_m[39,7]
C: ex_max[41,9],ex_m[39,7] R: ex_max[41,9]
2018-07-14 20:31:40,977 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:31:40,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:31:40,978 [LocalJobRunner Map Task Executor #0] INFO
```

org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:31:40,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:31:40,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:31:40,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:31:40,988 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1613517257_0063_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:31:40,989 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:31:40,989 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1613517257_0063_m_000000_0' done.
2018-07-14 20:31:40,989 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1613517257_0063_m_000000_0
2018-07-14 20:31:40,989 [Thread-1730] INFO org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:31:40,989 [Thread-1730] INFO org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:31:40,990 [pool-174-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1613517257_0063_r_000000_0
2018-07-14 20:31:40,994 [pool-174-thread-1] INFO org.apache.hadoop.mapred.Task - Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:40,994 [pool-174-thread-1] INFO org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@18f0d7cb
2018-07-14 20:31:40,995 [pool-174-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:31:40,996 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1613517257_0063_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:31:40,997 [localfetcher#48] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#48 about to shuffle
output of map attempt_local1613517257_0063_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:31:40,997 [localfetcher#48] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1613517257_0063_m_000000_0
2018-07-14 20:31:40,997 [localfetcher#48] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:31:40,998 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:31:40,998 [pool-174-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:40,998 [pool-174-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:31:40,999 [pool-174-thread-1] INFO org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:41,000 [pool-174-thread-1] INFO org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:31:41,003 [pool-174-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:31:41,003 [pool-174-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from

```
disk
2018-07-14 20:31:41,003 [pool-174-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:31:41,003 [pool-174-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:41,003 [pool-174-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:31:41,004 [pool-174-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:41,014 [pool-174-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:41,014 [pool-174-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:31:41,016 [pool-174-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[41,9],ex_m[39,7]
C: ex_max[41,9],ex_m[39,7] R: ex_max[41,9]
2018-07-14 20:31:41,017 [pool-174-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1613517257_0063_r_000000_0 is done. And is in the process of committing
2018-07-14 20:31:41,017 [pool-174-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:41,017 [pool-174-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1613517257_0063_r_000000_0 is allowed to commit now
2018-07-14 20:31:41,018 [pool-174-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1613517257_0063_r_000000_0' to
file:/tmp/temp-1902294536/tmp890997638/_temporary/0/task_local1613517257_0063_r_000000
2018-07-14 20:31:41,019 [pool-174-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:31:41,019 [pool-174-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1613517257_0063_r_000000_0' done.
2018-07-14 20:31:41,019 [pool-174-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1613517257_0063_r_000000_0
2018-07-14 20:31:41,020 [Thread-1730] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:31:41,146 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1613517257_0063
2018-07-14 20:31:41,146 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:31:41,146 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[41,9],ex_m[39,7] C: ex_max[41,9],ex_m[39,7] R: ex_max[41,9]
2018-07-14 20:31:41,156 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,160 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,162 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,167 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:31:41,170 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:31:40     2018-07-14 20:31:41
GROUP_BY
```

```
Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1004043501_0062        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1613517257_0063        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp890997638,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp890997638"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1004043501_0062        ->      job_local1613517257_0063,
job_local1613517257_0063


2018-07-14 20:31:41,172 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,173 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,174 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,176 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,177 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,178 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:41,180 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:31:41,180 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:41,180 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:41,181 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:31:41,191 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:41,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,500)
grunt> dump ex_max;
2018-07-14 20:31:46,495 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:31:46,513 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:46,513 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:46,513 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:31:46,513 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:31:46,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:31:46,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:31:46,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:31:46,518 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:31:46,518 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:31:46,528 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:46,528 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:46,529 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:46,530 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:31:46,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:31:46,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:31:46,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:31:46,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:31:46,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:31:46,531 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:31:46,531 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:31:46,531 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:31:46,531 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580506531-0
2018-07-14 20:31:46,541 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:31:46,547 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
```

```
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:46,553 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:31:46,556 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:31:46,558 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:46,558 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:31:46,559 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:31:46,578 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:31:46,592 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1711556306_0064
2018-07-14 20:31:46,652 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:31:46,653 [Thread-1759] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:31:46,656 [Thread-1759] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:31:46,656 [Thread-1759] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:31:46,656 [Thread-1759] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:46,656 [Thread-1759] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:31:46,656 [Thread-1759] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:46,656 [Thread-1759] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:31:46,659 [Thread-1759] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:31:46,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1711556306_0064_m_000000_0
2018-07-14 20:31:46,664 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:46,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:31:46,667 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:31:46,667 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
```

```
2018-07-14 20:31:46,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:31:46,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:31:46,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:31:46,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:31:46,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:31:46,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:31:46,683 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:46,683 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:31:46,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:31:46,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:31:46,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:31:46,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:31:46,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:31:46,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:31:46,694 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:31:46,695 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1711556306_0064_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:31:46,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:31:46,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1711556306_0064_m_000000_0' done.
2018-07-14 20:31:46,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1711556306_0064_m_000000_0
2018-07-14 20:31:46,696 [Thread-1759] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:31:46,696 [Thread-1759] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:31:46,696 [pool-177-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1711556306_0064_r_000000_0
2018-07-14 20:31:46,701 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:46,701 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6c5f219
2018-07-14 20:31:46,702 [pool-177-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:31:46,702 [EventFetcher for fetching Map Completion Events] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1711556306_0064_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:31:46,703 [localfetcher#49] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#49 about to shuffle
output of map attempt_local1711556306_0064_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:31:46,704 [localfetcher#49] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1711556306_0064_m_000000_0
2018-07-14 20:31:46,704 [localfetcher#49] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:31:46,706 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:31:46,706 [pool-177-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:46,706 [pool-177-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:31:46,707 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:46,707 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:31:46,709 [pool-177-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:31:46,709 [pool-177-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:31:46,709 [pool-177-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:31:46,709 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:46,710 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:31:46,710 [pool-177-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:46,819 [pool-177-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:46,819 [pool-177-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:31:46,820 [pool-177-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:31:46,821 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1711556306_0064_r_000000_0 is done. And is in the process of committing
2018-07-14 20:31:46,822 [pool-177-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:46,822 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1711556306_0064_r_000000_0 is allowed to commit now
2018-07-14 20:31:46,823 [pool-177-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1711556306_0064_r_000000_0' to
file:/tmp/temp-1902294536/tmp32739841/_temporary/0/task_local1711556306_0064_r_000000
2018-07-14 20:31:46,823 [pool-177-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:31:46,823 [pool-177-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1711556306_0064_r_000000_0' done.
```

```
2018-07-14 20:31:46,823 [pool-177-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1711556306_0064_r_000000_0
2018-07-14 20:31:46,823 [Thread-1759] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:31:47,020 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1711556306_0064
2018-07-14 20:31:47,020 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:31:47,020 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:31:47,022 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:31:47,024 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,025 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,026 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,028 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:31:47,028 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:31:47,028 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:31:47,028 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:31:47,030 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:31:47,042 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:31:47,043 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,045 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:47,045 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:47,051 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:31:47,061 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:31:47,061 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:31:47,061 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:31:47,081 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
```

```
2018-07-14 20:31:47,096 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2000800424_0065
2018-07-14 20:31:47,162 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:31:47,163 [Thread-1788] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:31:47,166 [Thread-1788] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:31:47,166 [Thread-1788] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:31:47,167 [Thread-1788] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:47,167 [Thread-1788] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:31:47,167 [Thread-1788] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:31:47,167 [Thread-1788] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:31:47,169 [Thread-1788] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:31:47,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2000800424_0065_m_000000_0
2018-07-14 20:31:47,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:47,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:31:47,175 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp32739841/part-r-00000:0+62
2018-07-14 20:31:47,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:31:47,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:31:47,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:31:47,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:31:47,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:31:47,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:31:47,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:47,192 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
```

```
2018-07-14 20:31:47,193 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[41,9],ex_m[39,7]
C: ex_max[41,9],ex_m[39,7] R: ex_max[41,9]
2018-07-14 20:31:47,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:31:47,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:31:47,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:31:47,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:31:47,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:31:47,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:31:47,202 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2000800424_0065_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:31:47,203 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:31:47,203 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2000800424_0065_m_000000_0' done.
2018-07-14 20:31:47,203 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2000800424_0065_m_000000_0
2018-07-14 20:31:47,203 [Thread-1788] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:31:47,204 [Thread-1788] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:31:47,204 [pool-180-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2000800424_0065_r_000000_0
2018-07-14 20:31:47,210 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:31:47,210 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3c74eb1
2018-07-14 20:31:47,210 [pool-180-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:31:47,211 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2000800424_0065_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:31:47,212 [localfetcher#50] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#50 about to shuffle
output of map attempt_local2000800424_0065_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:31:47,212 [localfetcher#50] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local2000800424_0065_m_000000_0
2018-07-14 20:31:47,212 [localfetcher#50] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:31:47,213 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:31:47,214 [pool-180-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:47,214 [pool-180-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
```

```
2018-07-14 20:31:47,215 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:47,215 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:31:47,216 [pool-180-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:31:47,216 [pool-180-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:31:47,216 [pool-180-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:31:47,216 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:31:47,216 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:31:47,216 [pool-180-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:47,222 [pool-180-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:31:47,222 [pool-180-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:31:47,224 [pool-180-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[41,9],ex_m[39,7]
C: ex_max[41,9],ex_m[39,7] R: ex_max[41,9]
2018-07-14 20:31:47,225 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2000800424_0065_r_000000_0 is done. And is in the process of committing
2018-07-14 20:31:47,226 [pool-180-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:31:47,226 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local2000800424_0065_r_000000_0 is allowed to commit now
2018-07-14 20:31:47,227 [pool-180-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2000800424_0065_r_000000_0' to
file:/tmp/temp-1902294536/tmp1369944136/_temporary/0/task_local2000800424_0065_r_00000
0
2018-07-14 20:31:47,228 [pool-180-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:31:47,228 [pool-180-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2000800424_0065_r_000000_0' done.
2018-07-14 20:31:47,228 [pool-180-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2000800424_0065_r_000000_0
2018-07-14 20:31:47,228 [Thread-1788] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:31:47,364 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2000800424_0065
2018-07-14 20:31:47,364 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:31:47,364 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[41,9],ex_m[39,7] C: ex_max[41,9],ex_m[39,7] R: ex_max[41,9]
2018-07-14 20:31:47,371 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,373 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,373 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:31:47,376 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:31:47,381 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:31:46     2018-07-14 20:31:47
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1711556306_0064        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local2000800424_0065        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp1369944136,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp1369944136"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1711556306_0064        ->      job_local2000800424_0065,
job_local2000800424_0065


2018-07-14 20:31:47,382 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,383 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,383 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,387 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,387 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,388 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:31:47,389 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:31:47,389 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:31:47,389 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:31:47,389 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:31:47,398 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
```

```
2018-07-14 20:31:47,398 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,500)
grunt> describe ex_max;
ex_max: {group: chararray,max: long}
grunt>
grunt> ex_max = foreach ex_m generate group, ex_sum.EmpId as EmpId, MAX(ex_sum.sum) as max;
grunt> describe ex_max;
ex_max: {group: chararray,EmpId: {(EmpId: int)},max: long}
grunt> dump ex_max;
2018-07-14 20:33:22,762 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:33:22,779 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:22,780 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:33:22,780 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:33:22,780 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:33:22,793 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:33:22,794 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:33:22,794 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:33:22,794 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:33:22,804 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:22,804 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:33:22,805 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:22,806 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:33:22,806 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:33:22,806 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:33:22,806 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:33:22,807 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:33:22,807 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:33:22,808 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:33:22,808 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:33:22,808 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:33:22,808 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580602808-0
2018-07-14 20:33:22,818 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:33:22,819 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:22,828 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:33:22,831 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:33:22,831 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:33:22,831 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:33:22,831 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:33:22,851 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:33:22,866 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1472761447_0066
2018-07-14 20:33:22,950 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:33:22,951 [Thread-1817] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:33:22,955 [Thread-1817] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:33:22,955 [Thread-1817] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:33:22,956 [Thread-1817] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:22,956 [Thread-1817] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:33:22,956 [Thread-1817] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:33:22,956 [Thread-1817] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:33:22,958 [Thread-1817] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:33:22,958 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1472761447_0066_m_000000_0
2018-07-14 20:33:22,964 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:22,964 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
```

```
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:33:22,966 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:33:22,966 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:33:22,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:33:22,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:33:22,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:33:22,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:33:22,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:33:22,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:33:22,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:22,985 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:33:22,986 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:33:22,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:33:22,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:33:22,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:33:22,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:33:22,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:33:22,995 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:33:22,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1472761447_0066_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:33:22,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:33:22,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1472761447_0066_m_000000_0' done.
2018-07-14 20:33:22,997 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1472761447_0066_m_000000_0
2018-07-14 20:33:22,997 [Thread-1817] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:33:22,998 [Thread-1817] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for reduce tasks
2018-07-14 20:33:22,998 [pool-183-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1472761447_0066_r_000000_0
2018-07-14 20:33:23,003 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:23,003 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@76365cec
2018-07-14 20:33:23,005 [pool-183-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:33:23,006 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1472761447_0066_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:33:23,011 [localfetcher#51] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#51 about to shuffle
output of map attempt_local1472761447_0066_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:33:23,012 [localfetcher#51] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1472761447_0066_m_000000_0
2018-07-14 20:33:23,012 [localfetcher#51] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:33:23,012 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:33:23,013 [pool-183-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:23,013 [pool-183-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:33:23,013 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:23,013 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:33:23,015 [pool-183-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:33:23,015 [pool-183-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:33:23,015 [pool-183-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:33:23,015 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:23,015 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:33:23,015 [pool-183-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:23,022 [pool-183-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:23,022 [pool-183-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:33:23,024 [pool-183-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:33:23,025 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Task -
```

```
Task:attempt_local1472761447_0066_r_000000_0 is done. And is in the process of committing
2018-07-14 20:33:23,026 [pool-183-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:23,026 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1472761447_0066_r_000000_0 is allowed to commit now
2018-07-14 20:33:23,027 [pool-183-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1472761447_0066_r_000000_0' to
file:/tmp/temp-1902294536/tmp1643165066/_temporary/0/task_local1472761447_0066_r_00000
0
2018-07-14 20:33:23,028 [pool-183-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:33:23,028 [pool-183-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1472761447_0066_r_000000_0' done.
2018-07-14 20:33:23,028 [pool-183-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1472761447_0066_r_000000_0
2018-07-14 20:33:23,028 [Thread-1817] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:33:23,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1472761447_0066
2018-07-14 20:33:23,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:33:23,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:33:23,154 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:33:23,158 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,160 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,162 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,168 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:33:23,171 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:33:23,172 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:33:23,172 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:33:23,176 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:33:23,209 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:33:23,213 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,218 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:23,218 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
```

```
Instead, use fs.defaultFS
2018-07-14 20:33:23,228 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:33:23,237 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:33:23,237 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:33:23,237 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:33:23,258 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:33:23,273 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1674606399_0067
2018-07-14 20:33:23,332 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:33:23,332 [Thread-1846] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:33:23,336 [Thread-1846] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:33:23,336 [Thread-1846] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:33:23,336 [Thread-1846] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:23,336 [Thread-1846] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:33:23,336 [Thread-1846] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:33:23,336 [Thread-1846] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:33:23,338 [Thread-1846] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:33:23,338 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1674606399_0067_m_000000_0
2018-07-14 20:33:23,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:23,343 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:33:23,343 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1643165066/part-r-00000:0+62
2018-07-14 20:33:23,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:33:23,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 20:33:23,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:33:23,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:33:23,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:33:23,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:33:23,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:23,442 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:33:23,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[42,9]
2018-07-14 20:33:23,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:33:23,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:33:23,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:33:23,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:33:23,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:33:23,444 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:33:23,447 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1674606399_0067_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:33:23,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:33:23,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1674606399_0067_m_000000_0' done.
2018-07-14 20:33:23,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1674606399_0067_m_000000_0
2018-07-14 20:33:23,448 [Thread-1846] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:33:23,448 [Thread-1846] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:33:23,448 [pool-186-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1674606399_0067_r_000000_0
2018-07-14 20:33:23,453 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:23,454 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@556e3d41
2018-07-14 20:33:23,455 [pool-186-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:33:23,456 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1674606399_0067_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:33:23,457 [localfetcher#52] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#52 about to shuffle
output of map attempt_local1674606399_0067_m_000000_0 decomp: 113 len: 117 to MEMORY
```

```
2018-07-14 20:33:23,457 [localfetcher#52] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1674606399_0067_m_000000_0
2018-07-14 20:33:23,457 [localfetcher#52] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:33:23,458 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:33:23,458 [pool-186-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:23,458 [pool-186-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:33:23,459 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:23,459 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:33:23,460 [pool-186-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:33:23,460 [pool-186-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:33:23,460 [pool-186-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:33:23,460 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:23,460 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:33:23,461 [pool-186-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:23,469 [pool-186-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:23,469 [pool-186-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:33:23,472 [pool-186-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[42,9]
2018-07-14 20:33:23,473 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1674606399_0067_r_000000_0 is done. And is in the process of committing
2018-07-14 20:33:23,474 [pool-186-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:23,474 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1674606399_0067_r_000000_0 is allowed to commit now
2018-07-14 20:33:23,475 [pool-186-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1674606399_0067_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1615356931/_temporary/0/task_local1674606399_0067_r_0000
00
2018-07-14 20:33:23,475 [pool-186-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:33:23,475 [pool-186-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1674606399_0067_r_000000_0' done.
2018-07-14 20:33:23,475 [pool-186-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1674606399_0067_r_000000_0
2018-07-14 20:33:23,475 [Thread-1846] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:33:23,643 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1674606399_0067
2018-07-14 20:33:23,643 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:33:23,643 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_m[39,7] C:  R: ex_max[42,9]
2018-07-14 20:33:23,652 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,658 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,662 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:33:23,676 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:33:22     2018-07-14 20:33:23
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1472761447_0066        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1674606399_0067        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY
file:/tmp/temp-1902294536/tmp-1615356931,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1615356931"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1472761447_0066        ->      job_local1674606399_0067,
job_local1674606399_0067


2018-07-14 20:33:23,678 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,680 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,681 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,684 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,685 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,687 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:23,688 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:33:23,689 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:23,689 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:33:23,689 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:33:23,700 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:33:23,700 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119),(114),(110),(105),(104),(102),(101)},500)
grunt> dump ex_max.EmpId;
2018-07-14 20:33:29,302 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:33:29,318 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:29,318 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:33:29,318 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:33:29,318 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:33:29,321 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:33:29,321 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:33:29,322 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:33:29,322 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:33:29,331 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:29,331 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:33:29,332 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,332 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:33:29,333 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:33:29,333 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:33:29,333 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:33:29,333 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:33:29,333 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:33:29,334 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:33:29,334 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:33:29,334 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:33:29,334 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580609334-0
2018-07-14 20:33:29,341 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:33:29,342 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,351 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:33:29,353 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:33:29,353 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:33:29,353 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:33:29,353 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:33:29,375 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:33:29,389 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1240010084_0068
2018-07-14 20:33:29,450 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:33:29,450 [Thread-1875] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:33:29,454 [Thread-1875] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:33:29,454 [Thread-1875] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:33:29,454 [Thread-1875] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:29,454 [Thread-1875] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:33:29,454 [Thread-1875] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:33:29,454 [Thread-1875] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:33:29,457 [Thread-1875] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for map tasks
2018-07-14 20:33:29,458 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1240010084_0068_m_000000_0
2018-07-14 20:33:29,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:29,463 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:33:29,465 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:33:29,465 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:33:29,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:33:29,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:33:29,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:33:29,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:33:29,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:33:29,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:33:29,482 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:29,482 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:33:29,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:33:29,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:33:29,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:33:29,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:33:29,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:33:29,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:33:29,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:33:29,498 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1240010084_0068_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:33:29,499 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
```

```
2018-07-14 20:33:29,499 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1240010084_0068_m_000000_0' done.
2018-07-14 20:33:29,499 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1240010084_0068_m_000000_0
2018-07-14 20:33:29,499 [Thread-1875] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:33:29,500 [Thread-1875] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:33:29,500 [pool-189-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1240010084_0068_r_000000_0
2018-07-14 20:33:29,504 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:29,505 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@417ec61c
2018-07-14 20:33:29,505 [pool-189-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:33:29,506 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1240010084_0068_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:33:29,507 [localfetcher#53] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#53 about to shuffle
output of map attempt_local1240010084_0068_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:33:29,507 [localfetcher#53] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1240010084_0068_m_000000_0
2018-07-14 20:33:29,507 [localfetcher#53] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:33:29,508 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:33:29,508 [pool-189-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:29,508 [pool-189-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:33:29,509 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:29,509 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:33:29,509 [pool-189-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:33:29,509 [pool-189-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:33:29,509 [pool-189-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:33:29,509 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:29,510 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:33:29,510 [pool-189-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:29,516 [pool-189-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 20:33:29,516 [pool-189-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:33:29,518 [pool-189-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:33:29,519 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1240010084_0068_r_000000_0 is done. And is in the process of committing
2018-07-14 20:33:29,522 [pool-189-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:29,522 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1240010084_0068_r_000000_0 is allowed to commit now
2018-07-14 20:33:29,523 [pool-189-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1240010084_0068_r_000000_0' to
file:/tmp/temp-1902294536/tmp-123163963/_temporary/0/task_local1240010084_0068_r_00000
0
2018-07-14 20:33:29,525 [pool-189-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:33:29,525 [pool-189-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1240010084_0068_r_000000_0' done.
2018-07-14 20:33:29,525 [pool-189-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1240010084_0068_r_000000_0
2018-07-14 20:33:29,526 [Thread-1875] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:33:29,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1240010084_0068
2018-07-14 20:33:29,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:33:29,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:33:29,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:33:29,652 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,652 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,653 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,655 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:33:29,655 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:33:29,655 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:33:29,655 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:33:29,656 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:33:29,663 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
```

```
map-reduce job(s) waiting for submission.
2018-07-14 20:33:29,664 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,666 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:29,666 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:33:29,673 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:33:29,680 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:33:29,680 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:33:29,681 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:33:29,701 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:33:29,715 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local45524956_0069
2018-07-14 20:33:29,776 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:33:29,777 [Thread-1904] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:33:29,780 [Thread-1904] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:33:29,780 [Thread-1904] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:33:29,780 [Thread-1904] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:29,780 [Thread-1904] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:33:29,780 [Thread-1904] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:33:29,780 [Thread-1904] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:33:29,783 [Thread-1904] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:33:29,783 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local45524956_0069_m_000000_0
2018-07-14 20:33:29,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:29,788 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------
```

```
2018-07-14 20:33:29,788 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-123163963/part-r-00000:0+62
2018-07-14 20:33:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:33:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:33:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:33:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:33:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:33:29,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:33:29,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:29,805 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:33:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[42,9]
2018-07-14 20:33:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:33:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:33:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:33:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:33:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:33:29,807 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:33:29,808 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local45524956_0069_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:33:29,809 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:33:29,809 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local45524956_0069_m_000000_0' done.
2018-07-14 20:33:29,809 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local45524956_0069_m_000000_0
2018-07-14 20:33:29,809 [Thread-1904] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:33:29,810 [Thread-1904] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:33:29,810 [pool-192-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local45524956_0069_r_000000_0
2018-07-14 20:33:29,818 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:33:29,818 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@59f3bcc6
2018-07-14 20:33:29,818 [pool-192-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
```

```
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:33:29,821 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local45524956_0069_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:33:29,821 [localfetcher#54] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#54 about to shuffle
output of map attempt_local45524956_0069_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:33:29,821 [localfetcher#54] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local45524956_0069_m_000000_0
2018-07-14 20:33:29,821 [localfetcher#54] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:33:29,822 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:33:29,822 [pool-192-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:29,822 [pool-192-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:33:29,823 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:29,823 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:33:29,824 [pool-192-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:33:29,824 [pool-192-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:33:29,824 [pool-192-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:33:29,824 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:33:29,824 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:33:29,824 [pool-192-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:29,832 [pool-192-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:33:29,832 [pool-192-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:33:29,834 [pool-192-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[42,9]
2018-07-14 20:33:29,835 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local45524956_0069_r_000000_0 is done. And is in the process of committing
2018-07-14 20:33:29,836 [pool-192-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:33:29,836 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local45524956_0069_r_000000_0 is allowed to commit now
2018-07-14 20:33:29,837 [pool-192-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local45524956_0069_r_000000_0' to
file:/tmp/temp-1902294536/tmp1527580405/_temporary/0/task_local45524956_0069_r_000000
2018-07-14 20:33:29,837 [pool-192-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:33:29,837 [pool-192-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
'attempt_local45524956_0069_r_000000_0' done.
2018-07-14 20:33:29,837 [pool-192-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local45524956_0069_r_000000_0
2018-07-14 20:33:29,837 [Thread-1904] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:33:29,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local45524956_0069
2018-07-14 20:33:29,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:33:29,980 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_m[39,7] C:  R: ex_max[42,9]
2018-07-14 20:33:29,982 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,982 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,982 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,984 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:33:29,985 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:33:29     2018-07-14 20:33:29
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1240010084_0068        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local45524956_0069  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY
file:/tmp/temp-1902294536/tmp1527580405,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp1527580405"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1240010084_0068        ->      job_local45524956_0069,
job_local45524956_0069


2018-07-14 20:33:29,986 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:33:29,987 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,988 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,989 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,990 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,991 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:33:29,992 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:33:29,992 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:33:29,992 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:33:29,992 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:33:30,053 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:33:30,053 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119),(114),(110),(105),(104),(102),(101)},500)
2018-07-14 20:33:30,054 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1000: Error
during parsing. Encountered " <PATH> "ex_max.EmpId "" at line 16, column 6.
Was expecting one of:
    <EOF>
    "cat" ...
    "clear" ...
    "fs" ...
    "sh" ...
    "cd" ...
    "cp" ...
    "copyFromLocal" ...
    "copyToLocal" ...
    "dump" ...
    "\\d" ...
    "describe" ...
    "\\de" ...
    "aliases" ...
    "explain" ...
    "\\e" ...
    "help" ...
    "history" ...
    "kill" ...
    "ls" ...
    "mv" ...
    "mkdir" ...
    "pwd" ...
    "quit" ...
    "\\q" ...
    "register" ...
    "rm" ...
    "rmf" ...
    "set" ...
    "illustrate" ...
    "\\i" ...
    "run" ...
    "exec" ...
    "%default" ...
    "%declare" ...
```

```
    "scriptDone" ...
    "" ...
    "" ...
    <EOL> ...
    ";" ...
    "@" ...
    <IDENTIFIER> ...

Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ex_max = foreach ex_m generate group, ex_sum.EmpId as EmpId, MAX(ex_sum.sum) as max;
grunt> xx = filter expense BY ex_max.EmpId;
2018-07-14 20:33:58,405 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 44, column 12> Undefined alias: expense
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> xx = filter expenses BY ex_max.EmpId;
2018-07-14 20:34:03,999 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1058:
<line 44, column 5> Filter's condition must evaluate to boolean. Found: bag
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> xx = filter ex_sum BY ex_max.EmpId;
2018-07-14 20:34:41,807 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1058:
<line 44, column 5> Filter's condition must evaluate to boolean. Found: bag
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> dump ex_max;;
2018-07-14 20:35:31,539 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:35:31,556 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:35:31,557 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:35:31,557 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:35:31,557 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:35:31,558 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:35:31,559 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:35:31,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:35:31,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:35:31,572 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:35:31,572 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:35:31,573 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:31,573 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:35:31,573 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:35:31,574 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:35:31,574 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:35:31,574 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:35:31,574 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:35:31,575 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:35:31,576 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:35:31,576 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:35:31,576 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580731576-0
2018-07-14 20:35:31,586 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:35:31,587 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:31,595 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:35:31,596 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:35:31,596 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:35:31,596 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:35:31,596 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:35:31,617 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:35:31,631 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local924103656_0070
2018-07-14 20:35:31,693 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:35:31,693 [Thread-1933] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:35:31,697 [Thread-1933] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:35:31,697 [Thread-1933] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:35:31,697 [Thread-1933] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:35:31,697 [Thread-1933] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:35:31,697 [Thread-1933] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
```

```
Instead, use fs.defaultFS
2018-07-14 20:35:31,697 [Thread-1933] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:35:31,699 [Thread-1933] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:35:31,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local924103656_0070_m_000000_0
2018-07-14 20:35:31,704 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:35:31,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:35:31,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:35:31,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:35:31,723 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:35:31,723 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:35:31,723 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:35:31,723 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:35:31,723 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:35:31,723 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:35:31,724 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:35:31,724 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:35:31,727 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:35:31,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:35:31,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:35:31,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:35:31,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:35:31,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:35:31,737 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
```

```
2018-07-14 20:35:31,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local924103656_0070_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:35:31,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:35:31,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local924103656_0070_m_000000_0' done.
2018-07-14 20:35:31,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local924103656_0070_m_000000_0
2018-07-14 20:35:31,739 [Thread-1933] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:35:31,740 [Thread-1933] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:35:31,740 [pool-195-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local924103656_0070_r_000000_0
2018-07-14 20:35:31,744 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:35:31,744 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@d87f6bf
2018-07-14 20:35:31,745 [pool-195-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:35:31,746 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local924103656_0070_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:35:31,746 [localfetcher#55] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#55 about to shuffle
output of map attempt_local924103656_0070_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:35:31,746 [localfetcher#55] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local924103656_0070_m_000000_0
2018-07-14 20:35:31,746 [localfetcher#55] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:35:31,747 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:35:31,747 [pool-195-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:35:31,747 [pool-195-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:35:31,751 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:35:31,751 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:35:31,752 [pool-195-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:35:31,752 [pool-195-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:35:31,752 [pool-195-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:35:31,752 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:35:31,752 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
```

```
2018-07-14 20:35:31,752 [pool-195-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:35:31,758 [pool-195-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:35:31,758 [pool-195-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:35:31,762 [pool-195-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:35:31,763 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local924103656_0070_r_000000_0 is done. And is in the process of committing
2018-07-14 20:35:31,764 [pool-195-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:35:31,764 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local924103656_0070_r_000000_0 is allowed to commit now
2018-07-14 20:35:31,765 [pool-195-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local924103656_0070_r_000000_0' to
file:/tmp/temp-1902294536/tmp-857361120/_temporary/0/task_local924103656_0070_r_000000
2018-07-14 20:35:31,765 [pool-195-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:35:31,765 [pool-195-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local924103656_0070_r_000000_0' done.
2018-07-14 20:35:31,765 [pool-195-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local924103656_0070_r_000000_0
2018-07-14 20:35:31,766 [Thread-1933] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:35:31,896 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local924103656_0070
2018-07-14 20:35:31,896 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:35:31,896 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:35:31,897 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:35:31,898 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:31,899 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:31,900 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:31,904 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:35:31,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:35:31,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:35:31,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:35:31,906 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:35:31,924 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:35:31,925 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:31,927 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:35:31,927 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:35:31,933 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:35:31,943 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:35:31,943 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:35:31,943 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:35:31,963 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:35:31,976 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1078640829_0071
2018-07-14 20:35:32,093 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:35:32,094 [Thread-1962] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:35:32,097 [Thread-1962] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:35:32,097 [Thread-1962] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:35:32,097 [Thread-1962] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:35:32,097 [Thread-1962] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:35:32,097 [Thread-1962] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:35:32,097 [Thread-1962] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:35:32,099 [Thread-1962] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:35:32,100 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1078640829_0071_m_000000_0
2018-07-14 20:35:32,104 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:35:32,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
```

```
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:35:32,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-857361120/part-r-00000:0+62
2018-07-14 20:35:32,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:35:32,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:35:32,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:35:32,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:35:32,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:35:32,124 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:35:32,124 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:35:32,124 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:35:32,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[43,9]
2018-07-14 20:35:32,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:35:32,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:35:32,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:35:32,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:35:32,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:35:32,126 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:35:32,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1078640829_0071_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:35:32,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:35:32,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1078640829_0071_m_000000_0' done.
2018-07-14 20:35:32,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1078640829_0071_m_000000_0
2018-07-14 20:35:32,131 [Thread-1962] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:35:32,131 [Thread-1962] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:35:32,132 [pool-198-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1078640829_0071_r_000000_0
2018-07-14 20:35:32,138 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:35:32,138 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
```

```
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@56f7be3b
2018-07-14 20:35:32,140 [pool-198-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:35:32,141 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1078640829_0071_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:35:32,143 [localfetcher#56] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#56 about to shuffle
output of map attempt_local1078640829_0071_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:35:32,143 [localfetcher#56] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1078640829_0071_m_000000_0
2018-07-14 20:35:32,143 [localfetcher#56] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:35:32,146 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:35:32,147 [pool-198-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:35:32,147 [pool-198-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:35:32,148 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:35:32,148 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:35:32,148 [pool-198-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:35:32,148 [pool-198-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:35:32,149 [pool-198-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:35:32,149 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:35:32,149 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:35:32,149 [pool-198-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:35:32,157 [pool-198-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:35:32,159 [pool-198-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:35:32,161 [pool-198-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[43,9]
2018-07-14 20:35:32,162 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1078640829_0071_r_000000_0 is done. And is in the process of committing
2018-07-14 20:35:32,163 [pool-198-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:35:32,163 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1078640829_0071_r_000000_0 is allowed to commit now
2018-07-14 20:35:32,164 [pool-198-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1078640829_0071_r_000000_0' to
```

```
file:/tmp/temp-1902294536/tmp-1514121917/_temporary/0/task_local1078640829_0071_r_0000
00
2018-07-14 20:35:32,164 [pool-198-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:35:32,164 [pool-198-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1078640829_0071_r_000000_0' done.
2018-07-14 20:35:32,164 [pool-198-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1078640829_0071_r_000000_0
2018-07-14 20:35:32,164 [Thread-1962] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:35:32,295 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1078640829_0071
2018-07-14 20:35:32,295 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:35:32,295 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_m[39,7] C:  R: ex_max[43,9]
2018-07-14 20:35:32,304 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,307 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,310 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,313 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:35:32,314 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:35:31     2018-07-14 20:35:32
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1078640829_0071        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY
file:/tmp/temp-1902294536/tmp-1514121917,
job_local924103656_0070 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1514121917"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local924103656_0070 ->      job_local1078640829_0071,
```

```
job_local1078640829_0071


2018-07-14 20:35:32,316 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,318 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,319 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,323 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,324 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,324 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:35:32,326 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:35:32,326 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:35:32,326 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:35:32,327 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:35:32,336 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:35:32,336 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119),(114),(110),(105),(104),(102),(101)},500)
grunt> describe ex_max;
ex_max: {group: chararray,EmpId: {(EmpId: int)},max: long}
grunt> describe ex_sum;
ex_sum: {EmpId: int,sum: long}
grunt> xx= filter ex_sum BY sum==ex_max.max;
grunt> dump xx;
2018-07-14 20:36:45,653 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY,FILTER
2018-07-14 20:36:45,669 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:45,670 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:36:45,670 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:36:45,670 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:36:45,675 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:36:45,676 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:36:45,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 20:36:45,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 20:36:45,678 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 20:36:45,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 20:36:45,685 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:45,685 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:36:45,685 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:45,690 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:36:45,691 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:36:45,691 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:36:45,691 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:36:45,692 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:36:45,692 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:36:45,693 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:36:45,693 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:36:45,693 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:36:45,693 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580805693-0
2018-07-14 20:36:45,700 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:36:45,702 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:45,710 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:36:45,711 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:36:45,711 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:36:45,712 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:36:45,712 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:36:45,732 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:36:45,748 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1233372169_0072
2018-07-14 20:36:45,807 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
```

```
track the job: http://localhost:8080/
2018-07-14 20:36:45,808 [Thread-1991] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:36:45,812 [Thread-1991] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:36:45,812 [Thread-1991] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:36:45,812 [Thread-1991] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:45,812 [Thread-1991] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:36:45,812 [Thread-1991] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:36:45,812 [Thread-1991] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:36:45,815 [Thread-1991] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:36:45,815 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1233372169_0072_m_000000_0
2018-07-14 20:36:45,821 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:36:45,821 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:36:45,824 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:36:45,824 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:36:45,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:36:45,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:36:45,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:36:45,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:36:45,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:36:45,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:36:45,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:36:45,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
```

```
2018-07-14 20:36:45,845 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:36:45,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:36:45,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:36:45,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:36:45,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:36:45,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:36:45,852 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:36:45,853 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1233372169_0072_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:36:45,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:36:45,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1233372169_0072_m_000000_0' done.
2018-07-14 20:36:45,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1233372169_0072_m_000000_0
2018-07-14 20:36:45,854 [Thread-1991] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:36:45,855 [Thread-1991] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:36:45,855 [pool-201-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1233372169_0072_r_000000_0
2018-07-14 20:36:45,861 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:36:45,862 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4c31a808
2018-07-14 20:36:45,862 [pool-201-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:36:45,863 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1233372169_0072_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:36:45,865 [localfetcher#57] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#57 about to shuffle
output of map attempt_local1233372169_0072_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:36:45,865 [localfetcher#57] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1233372169_0072_m_000000_0
2018-07-14 20:36:45,865 [localfetcher#57] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:36:45,866 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:36:45,867 [pool-201-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:36:45,867 [pool-201-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
```

```
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:36:45,867 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:36:45,867 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:36:45,868 [pool-201-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:36:45,868 [pool-201-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:36:45,868 [pool-201-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:36:45,868 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:36:45,868 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:36:45,868 [pool-201-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:36:45,875 [pool-201-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:36:45,875 [pool-201-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:36:45,876 [pool-201-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:36:45,877 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1233372169_0072_r_000000_0 is done. And is in the process of committing
2018-07-14 20:36:45,878 [pool-201-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:36:45,878 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1233372169_0072_r_000000_0 is allowed to commit now
2018-07-14 20:36:45,879 [pool-201-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1233372169_0072_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1507488921/_temporary/0/task_local1233372169_0072_r_0000
00
2018-07-14 20:36:45,879 [pool-201-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:36:45,879 [pool-201-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1233372169_0072_r_000000_0' done.
2018-07-14 20:36:45,879 [pool-201-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1233372169_0072_r_000000_0
2018-07-14 20:36:45,879 [Thread-1991] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:36:46,009 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1233372169_0072
2018-07-14 20:36:46,009 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:36:46,009 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:36:46,010 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
```

```
2018-07-14 20:36:46,011 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,011 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,011 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,013 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:36:46,013 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:36:46,013 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:36:46,013 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:36:46,014 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:36:46,023 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:36:46,024 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,026 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:46,026 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:36:46,033 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:36:46,044 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:36:46,044 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:36:46,045 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:36:46,066 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:36:46,080 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local573964620_0073
2018-07-14 20:36:46,153 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:36:46,154 [Thread-2020] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:36:46,157 [Thread-2020] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:36:46,158 [Thread-2020] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:36:46,158 [Thread-2020] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:46,158 [Thread-2020] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
```

```
Instead, use mapreduce.job.reduces
2018-07-14 20:36:46,158 [Thread-2020] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:36:46,159 [Thread-2020] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:36:46,161 [Thread-2020] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:36:46,161 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local573964620_0073_m_000000_0
2018-07-14 20:36:46,165 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:36:46,165 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:36:46,166 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1507488921/part-r-00000:0+62
2018-07-14 20:36:46,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:36:46,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:36:46,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:36:46,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:36:46,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:36:46,184 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:36:46,185 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:36:46,185 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:36:46,186 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[43,9]
2018-07-14 20:36:46,187 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:36:46,187 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:36:46,187 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:36:46,187 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:36:46,187 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:36:46,189 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:36:46,189 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task:attempt_local573964620_0073_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:36:46,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:36:46,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local573964620_0073_m_000000_0' done.
2018-07-14 20:36:46,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local573964620_0073_m_000000_0
2018-07-14 20:36:46,191 [Thread-2020] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:36:46,192 [Thread-2020] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:36:46,193 [pool-204-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local573964620_0073_r_000000_0
2018-07-14 20:36:46,198 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:36:46,198 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@24e674b7
2018-07-14 20:36:46,198 [pool-204-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:36:46,199 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local573964620_0073_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:36:46,200 [localfetcher#58] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#58 about to shuffle
output of map attempt_local573964620_0073_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:36:46,200 [localfetcher#58] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local573964620_0073_m_000000_0
2018-07-14 20:36:46,200 [localfetcher#58] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:36:46,204 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:36:46,205 [pool-204-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:36:46,206 [pool-204-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:36:46,210 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:36:46,211 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:36:46,212 [pool-204-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:36:46,212 [pool-204-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:36:46,212 [pool-204-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:36:46,212 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:36:46,212 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:36:46,212 [pool-204-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:36:46,219 [pool-204-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:36:46,219 [pool-204-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:36:46,222 [pool-204-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
ex_max[43,9]
2018-07-14 20:36:46,225 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local573964620_0073_r_000000_0 is done. And is in the process of committing
2018-07-14 20:36:46,226 [pool-204-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:36:46,226 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local573964620_0073_r_000000_0 is allowed to commit now
2018-07-14 20:36:46,228 [pool-204-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local573964620_0073_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1427457561/_temporary/0/task_local573964620_0073_r_00000
0
2018-07-14 20:36:46,228 [pool-204-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:36:46,228 [pool-204-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local573964620_0073_r_000000_0' done.
2018-07-14 20:36:46,228 [pool-204-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local573964620_0073_r_000000_0
2018-07-14 20:36:46,228 [Thread-2020] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:36:46,355 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local573964620_0073
2018-07-14 20:36:46,355 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:36:46,355 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_m[39,7] C:  R: ex_max[43,9]
2018-07-14 20:36:46,361 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:36:46,367 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,370 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,372 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,378 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:36:46,379 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:36:46,380 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:36:46,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:36:46,407 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,410 [JobControl] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:46,410 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:36:46,416 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:36:46,423 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:36:46,423 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:36:46,423 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:36:46,443 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:36:46,456 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1314861035_0074
2018-07-14 20:36:46,518 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:36:46,519 [Thread-2049] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:36:46,522 [Thread-2049] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:36:46,522 [Thread-2049] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:36:46,522 [Thread-2049] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:46,522 [Thread-2049] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:36:46,522 [Thread-2049] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:36:46,524 [Thread-2049] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:36:46,524 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1314861035_0074_m_000000_0
2018-07-14 20:36:46,529 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:36:46,530 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:36:46,531 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1507488921/part-r-00000:0+62
2018-07-14 20:36:46,537 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 20:36:46,537 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:36:46,538 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: xx[44,4] C:  R:
2018-07-14 20:36:46,538 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:36:46,545 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:36:46,545 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:36:46,546 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:36:46,546 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1314861035_0074_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:36:46,547 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:36:46,547 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1314861035_0074_m_000000_0 is allowed
to commit now
2018-07-14 20:36:46,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1314861035_0074_m_000000_0' to
file:/tmp/temp-1902294536/tmp-2082211056/_temporary/0/task_local1314861035_0074_m_0000
00
2018-07-14 20:36:46,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:36:46,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1314861035_0074_m_000000_0' done.
2018-07-14 20:36:46,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1314861035_0074_m_000000_0
2018-07-14 20:36:46,548 [Thread-2049] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:36:46,720 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1314861035_0074
2018-07-14 20:36:46,720 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases xx
2018-07-14 20:36:46,720 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: xx[44,4] C:  R:
2018-07-14 20:36:46,724 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,725 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,726 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,728 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:36:46,731 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:36:45     2018-07-14 20:36:46
GROUP_BY,FILTER

Success!
```

```
Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1233372169_0072        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1314861035_0074        1       0       n/a     n/a     n/a     n/a     0
0       0       0       xx      MAP_ONLY
file:/tmp/temp-1902294536/tmp-2082211056,
job_local573964620_0073 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-2082211056"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1233372169_0072        ->
job_local573964620_0073,job_local1314861035_0074,
job_local573964620_0073 ->      job_local1314861035_0074,
job_local1314861035_0074


2018-07-14 20:36:46,732 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,734 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,735 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,737 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,737 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,738 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,741 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,742 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,744 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:36:46,745 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:36:46,745 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:36:46,745 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:36:46,745 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:36:46,753 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:36:46,753 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
(102,500)
grunt> describe ex_m;
ex_m: {group: chararray,ex_sum: {(EmpId: int,sum: long)}}
grunt> dump ex_m;
2018-07-14 20:37:23,564 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:37:23,584 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:23,585 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:37:23,585 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:37:23,585 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:37:23,587 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:37:23,588 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:37:23,588 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:37:23,588 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:37:23,599 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:23,600 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:37:23,600 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:23,601 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:37:23,601 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:37:23,602 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:37:23,602 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:37:23,602 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:37:23,602 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:37:23,603 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:37:23,604 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:37:23,604 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
```

```
2018-07-14 20:37:23,604 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580843603-0
2018-07-14 20:37:23,614 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:37:23,616 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:23,624 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:37:23,626 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:37:23,627 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:37:23,627 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:37:23,627 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:37:23,649 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:37:23,664 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local60761791_0075
2018-07-14 20:37:23,741 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:37:23,741 [Thread-2074] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:37:23,745 [Thread-2074] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:37:23,745 [Thread-2074] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:37:23,746 [Thread-2074] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:23,746 [Thread-2074] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:37:23,746 [Thread-2074] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:37:23,746 [Thread-2074] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:37:23,749 [Thread-2074] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:37:23,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local60761791_0075_m_000000_0
2018-07-14 20:37:23,754 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:23,755 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:
```

```
-----------------------

2018-07-14 20:37:23,757 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:37:23,757 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:37:23,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:37:23,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:37:23,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:37:23,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:37:23,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:37:23,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:37:23,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:23,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:37:23,780 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:37:23,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:37:23,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:37:23,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:37:23,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:37:23,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:37:23,792 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:37:23,793 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local60761791_0075_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:37:23,795 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:37:23,795 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local60761791_0075_m_000000_0' done.
2018-07-14 20:37:23,795 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local60761791_0075_m_000000_0
2018-07-14 20:37:23,795 [Thread-2074] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:37:23,796 [Thread-2074] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:37:23,797 [pool-209-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local60761791_0075_r_000000_0
2018-07-14 20:37:23,807 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
```

```
2018-07-14 20:37:23,807 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6e15b798
2018-07-14 20:37:23,808 [pool-209-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:37:23,809 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local60761791_0075_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:37:23,810 [localfetcher#59] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#59 about to shuffle
output of map attempt_local60761791_0075_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:37:23,810 [localfetcher#59] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local60761791_0075_m_000000_0
2018-07-14 20:37:23,810 [localfetcher#59] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:37:23,811 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:37:23,813 [pool-209-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:23,813 [pool-209-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:37:23,814 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:23,814 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:37:23,815 [pool-209-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:37:23,815 [pool-209-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:37:23,815 [pool-209-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:37:23,815 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:23,815 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:37:23,816 [pool-209-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:23,824 [pool-209-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:23,824 [pool-209-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:37:23,829 [pool-209-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:37:23,831 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local60761791_0075_r_000000_0 is done. And is in the process of committing
2018-07-14 20:37:23,833 [pool-209-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:23,833 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local60761791_0075_r_000000_0 is allowed to commit now
2018-07-14 20:37:23,834 [pool-209-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local60761791_0075_r_000000_0' to
file:/tmp/temp-1902294536/tmp389286920/_temporary/0/task_local60761791_0075_r_000000
2018-07-14 20:37:23,834 [pool-209-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:37:23,834 [pool-209-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local60761791_0075_r_000000_0' done.
2018-07-14 20:37:23,835 [pool-209-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local60761791_0075_r_000000_0
2018-07-14 20:37:23,835 [Thread-2074] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:37:24,046 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local60761791_0075
2018-07-14 20:37:24,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:37:24,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:37:24,050 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:37:24,055 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,057 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,060 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,067 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:37:24,067 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:37:24,067 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:37:24,067 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:37:24,070 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:37:24,085 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:37:24,087 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,089 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:24,089 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:37:24,096 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:37:24,105 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:37:24,105 [JobControl] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:37:24,105 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:37:24,125 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:37:24,139 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1442042877_0076
2018-07-14 20:37:24,208 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:37:24,209 [Thread-2103] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:37:24,213 [Thread-2103] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:37:24,213 [Thread-2103] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:37:24,213 [Thread-2103] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:24,213 [Thread-2103] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:37:24,213 [Thread-2103] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:37:24,214 [Thread-2103] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:37:24,215 [Thread-2103] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:37:24,216 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1442042877_0076_m_000000_0
2018-07-14 20:37:24,220 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:24,220 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:37:24,221 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp389286920/part-r-00000:0+62
2018-07-14 20:37:24,236 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:37:24,236 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:37:24,236 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:37:24,236 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:37:24,236 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:37:24,237 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:37:24,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:24,238 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:37:24,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
2018-07-14 20:37:24,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:37:24,240 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:37:24,240 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:37:24,240 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 20:37:24,240 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:37:24,240 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:37:24,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1442042877_0076_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:37:24,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:37:24,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1442042877_0076_m_000000_0' done.
2018-07-14 20:37:24,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1442042877_0076_m_000000_0
2018-07-14 20:37:24,242 [Thread-2103] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:37:24,242 [Thread-2103] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:37:24,242 [pool-212-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1442042877_0076_r_000000_0
2018-07-14 20:37:24,250 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:24,250 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2deb3ff
2018-07-14 20:37:24,251 [pool-212-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:37:24,252 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1442042877_0076_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:37:24,253 [localfetcher#60] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#60 about to shuffle
output of map attempt_local1442042877_0076_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 20:37:24,253 [localfetcher#60] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local1442042877_0076_m_000000_0
2018-07-14 20:37:24,253 [localfetcher#60] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 20:37:24,254 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 20:37:24,256 [pool-212-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:24,256 [pool-212-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:37:24,257 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:24,257 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:37:24,257 [pool-212-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:37:24,258 [pool-212-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 20:37:24,258 [pool-212-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:37:24,258 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:24,258 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 20:37:24,258 [pool-212-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:24,265 [pool-212-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:24,265 [pool-212-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:37:24,265 [pool-212-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[39,7] C:  R:
2018-07-14 20:37:24,266 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1442042877_0076_r_000000_0 is done. And is in the process of committing
2018-07-14 20:37:24,266 [pool-212-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:24,266 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1442042877_0076_r_000000_0 is allowed to commit now
2018-07-14 20:37:24,267 [pool-212-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1442042877_0076_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1185416216/_temporary/0/task_local1442042877_0076_r_0000
00
2018-07-14 20:37:24,268 [pool-212-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:37:24,268 [pool-212-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1442042877_0076_r_000000_0' done.
2018-07-14 20:37:24,268 [pool-212-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1442042877_0076_r_000000_0
2018-07-14 20:37:24,268 [Thread-2103] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:37:24,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1442042877_0076
2018-07-14 20:37:24,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m
2018-07-14 20:37:24,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_m[39,7] C:  R:
2018-07-14 20:37:24,416 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,418 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,421 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,425 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:37:24,428 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:37:23     2018-07-14 20:37:24
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1442042877_0076        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m    GROUP_BY
file:/tmp/temp-1902294536/tmp-1185416216,
job_local60761791_0075  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1185416216"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local60761791_0075  ->      job_local1442042877_0076,
job_local1442042877_0076


2018-07-14 20:37:24,431 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,433 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,433 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,436 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,437 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,437 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:24,440 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:37:24,440 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:24,441 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:37:24,441 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:37:24,451 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:37:24,451 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt> ex_max = foreach ex_m generate group,MAX(ex_sum.sum) as max;
grunt> dump ex_max;
2018-07-14 20:37:55,574 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:37:55,590 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:55,590 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:37:55,590 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:37:55,590 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:37:55,595 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:37:55,598 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:37:55,598 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:37:55,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:37:55,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:37:55,607 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:55,608 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:37:55,608 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:55,609 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:37:55,609 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:37:55,609 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:37:55,609 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:37:55,609 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:37:55,610 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:37:55,611 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:37:55,612 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:37:55,612 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:37:55,612 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580875612-0
2018-07-14 20:37:55,620 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:37:55,622 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:55,630 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:37:55,631 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:37:55,631 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:37:55,631 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:37:55,631 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:37:55,652 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:37:55,666 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1895697216_0077
2018-07-14 20:37:55,729 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:37:55,729 [Thread-2132] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:37:55,732 [Thread-2132] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:37:55,732 [Thread-2132] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:37:55,733 [Thread-2132] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:55,733 [Thread-2132] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:37:55,733 [Thread-2132] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:37:55,733 [Thread-2132] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:37:55,735 [Thread-2132] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:37:55,735 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1895697216_0077_m_000000_0
2018-07-14 20:37:55,739 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:55,740 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:37:55,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:37:55,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:37:55,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:37:55,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:37:55,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:37:55,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:37:55,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:37:55,772 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:37:55,772 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:55,772 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:37:55,774 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:37:55,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:37:55,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:37:55,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:37:55,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:37:55,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:37:55,784 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:37:55,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1895697216_0077_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:37:55,786 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:37:55,786 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1895697216_0077_m_000000_0' done.
2018-07-14 20:37:55,786 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1895697216_0077_m_000000_0
```

```
2018-07-14 20:37:55,786 [Thread-2132] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:37:55,786 [Thread-2132] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:37:55,787 [pool-215-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1895697216_0077_r_000000_0
2018-07-14 20:37:55,791 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:55,791 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@b885fb3
2018-07-14 20:37:55,792 [pool-215-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:37:55,793 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1895697216_0077_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:37:55,796 [localfetcher#61] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#61 about to shuffle
output of map attempt_local1895697216_0077_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:37:55,797 [localfetcher#61] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1895697216_0077_m_000000_0
2018-07-14 20:37:55,797 [localfetcher#61] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:37:55,797 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:37:55,798 [pool-215-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:55,798 [pool-215-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:37:55,799 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:37:55,800 [pool-215-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:55,806 [pool-215-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:55,807 [pool-215-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:37:55,811 [pool-215-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
```

```
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:37:55,813 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1895697216_0077_r_000000_0 is done. And is in the process of committing
2018-07-14 20:37:55,814 [pool-215-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:55,814 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1895697216_0077_r_000000_0 is allowed to commit now
2018-07-14 20:37:55,815 [pool-215-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1895697216_0077_r_000000_0' to
file:/tmp/temp-1902294536/tmp2078851358/_temporary/0/task_local1895697216_0077_r_00000
0
2018-07-14 20:37:55,815 [pool-215-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:37:55,815 [pool-215-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1895697216_0077_r_000000_0' done.
2018-07-14 20:37:55,815 [pool-215-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1895697216_0077_r_000000_0
2018-07-14 20:37:55,815 [Thread-2132] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:37:55,930 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1895697216_0077
2018-07-14 20:37:55,930 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:37:55,930 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:37:55,931 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:37:55,932 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:55,933 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:55,934 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:55,936 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:37:55,936 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:37:55,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:37:55,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:37:55,938 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:37:55,950 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:37:55,952 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:55,954 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
```

```
Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:55,954 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:37:55,961 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:37:55,970 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:37:55,970 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:37:55,970 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:37:55,998 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:37:56,018 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local398286676_0078
2018-07-14 20:37:56,126 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:37:56,126 [Thread-2161] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:37:56,132 [Thread-2161] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:37:56,132 [Thread-2161] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:37:56,132 [Thread-2161] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:56,132 [Thread-2161] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:37:56,132 [Thread-2161] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:37:56,132 [Thread-2161] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:37:56,134 [Thread-2161] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:37:56,135 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local398286676_0078_m_000000_0
2018-07-14 20:37:56,140 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:56,140 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:37:56,141 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp2078851358/part-r-00000:0+62
2018-07-14 20:37:56,168 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:37:56,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:37:56,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:37:56,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:37:56,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:37:56,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:37:56,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:56,169 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:37:56,171 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[45,9],ex_m[39,7]
C: ex_max[45,9],ex_m[39,7] R: ex_max[45,9]
2018-07-14 20:37:56,172 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:37:56,172 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:37:56,172 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:37:56,172 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:37:56,172 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:37:56,181 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:37:56,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local398286676_0078_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:37:56,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:37:56,183 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local398286676_0078_m_000000_0' done.
2018-07-14 20:37:56,183 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local398286676_0078_m_000000_0
2018-07-14 20:37:56,183 [Thread-2161] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:37:56,183 [Thread-2161] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:37:56,183 [pool-218-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local398286676_0078_r_000000_0
2018-07-14 20:37:56,188 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:37:56,188 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@19d4884e
2018-07-14 20:37:56,188 [pool-218-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:37:56,189 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local398286676_0078_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
```

```
2018-07-14 20:37:56,190 [localfetcher#62] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#62 about to shuffle
output of map attempt_local398286676_0078_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:37:56,190 [localfetcher#62] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local398286676_0078_m_000000_0
2018-07-14 20:37:56,190 [localfetcher#62] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:37:56,190 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:37:56,191 [pool-218-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:56,191 [pool-218-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:37:56,191 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:56,191 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:37:56,192 [pool-218-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:37:56,193 [pool-218-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:37:56,193 [pool-218-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:37:56,193 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:37:56,193 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:37:56,193 [pool-218-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:56,199 [pool-218-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:37:56,199 [pool-218-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:37:56,201 [pool-218-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[45,9],ex_m[39,7]
C: ex_max[45,9],ex_m[39,7] R: ex_max[45,9]
2018-07-14 20:37:56,201 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local398286676_0078_r_000000_0 is done. And is in the process of committing
2018-07-14 20:37:56,202 [pool-218-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:37:56,202 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local398286676_0078_r_000000_0 is allowed to commit now
2018-07-14 20:37:56,203 [pool-218-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local398286676_0078_r_000000_0' to
file:/tmp/temp-1902294536/tmp481602855/_temporary/0/task_local398286676_0078_r_000000
2018-07-14 20:37:56,203 [pool-218-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:37:56,203 [pool-218-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local398286676_0078_r_000000_0' done.
2018-07-14 20:37:56,203 [pool-218-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local398286676_0078_r_000000_0
2018-07-14 20:37:56,204 [Thread-2161] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
reduce task executor complete.
2018-07-14 20:37:56,327 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local398286676_0078
2018-07-14 20:37:56,327 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:37:56,327 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[45,9],ex_m[39,7] C: ex_max[45,9],ex_m[39,7] R: ex_max[45,9]
2018-07-14 20:37:56,336 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,340 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,342 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,348 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:37:56,352 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:37:55     2018-07-14 20:37:56
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1895697216_0077        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local398286676_0078 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp481602855,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp481602855"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1895697216_0077        ->      job_local398286676_0078,
job_local398286676_0078


2018-07-14 20:37:56,356 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,362 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,366 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,373 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,376 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,380 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:37:56,384 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:37:56,385 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:37:56,387 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:37:56,387 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:37:56,401 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:37:56,401 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,500)
grunt> ex_max = foreach ex_m generate group,MAX(ex_sum.sum) as max;
grunt> ans = filter ex_sum BY sum==ex_max.max;
grunt> dump ans;
2018-07-14 20:39:09,689 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY,FILTER
2018-07-14 20:39:09,708 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:09,708 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:39:09,709 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:39:09,709 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:39:09,711 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:39:09,712 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:39:09,712 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:39:09,713 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 20:39:09,713 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 20:39:09,713 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 20:39:09,713 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 20:39:09,722 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:09,722 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:39:09,723 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:09,724 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:39:09,724 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:39:09,724 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:39:09,724 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:39:09,724 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:39:09,724 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:39:09,725 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:39:09,726 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:39:09,726 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:39:09,726 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531580949726-0
2018-07-14 20:39:09,736 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:39:09,737 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:09,744 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:39:09,745 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:39:09,745 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:39:09,745 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:39:09,746 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:39:09,770 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:39:09,783 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local856273166_0079
2018-07-14 20:39:09,857 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:39:09,857 [Thread-2190] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:39:09,861 [Thread-2190] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:39:09,861 [Thread-2190] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
```

```
2018-07-14 20:39:09,861 [Thread-2190] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:09,861 [Thread-2190] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:39:09,861 [Thread-2190] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:39:09,861 [Thread-2190] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:39:09,863 [Thread-2190] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:39:09,863 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local856273166_0079_m_000000_0
2018-07-14 20:39:09,867 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:39:09,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:39:09,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:39:09,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:39:09,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:39:09,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:39:09,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:39:09,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:39:09,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:39:09,892 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:39:09,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:39:09,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:39:09,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:39:09,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:39:09,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:39:09,898 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:39:09,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:39:09,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:39:09,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:39:09,907 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local856273166_0079_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:39:09,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:39:09,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local856273166_0079_m_000000_0' done.
2018-07-14 20:39:09,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local856273166_0079_m_000000_0
2018-07-14 20:39:09,908 [Thread-2190] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:39:09,908 [Thread-2190] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:39:09,908 [pool-221-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local856273166_0079_r_000000_0
2018-07-14 20:39:09,913 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:39:09,913 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3079a2d5
2018-07-14 20:39:09,914 [pool-221-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:39:09,914 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local856273166_0079_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:39:09,915 [localfetcher#63] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#63 about to shuffle
output of map attempt_local856273166_0079_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:39:09,915 [localfetcher#63] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local856273166_0079_m_000000_0
2018-07-14 20:39:09,916 [localfetcher#63] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:39:09,916 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:39:09,917 [pool-221-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:39:09,917 [pool-221-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:39:09,921 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:39:09,921 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:39:09,922 [pool-221-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:39:09,922 [pool-221-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
```

```
disk
2018-07-14 20:39:09,922 [pool-221-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:39:09,922 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:39:09,922 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:39:09,923 [pool-221-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:39:09,928 [pool-221-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:39:09,928 [pool-221-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:39:09,930 [pool-221-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:39:09,931 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local856273166_0079_r_000000_0 is done. And is in the process of committing
2018-07-14 20:39:09,932 [pool-221-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:39:09,932 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local856273166_0079_r_000000_0 is allowed to commit now
2018-07-14 20:39:09,933 [pool-221-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local856273166_0079_r_000000_0' to
file:/tmp/temp-1902294536/tmp-532868691/_temporary/0/task_local856273166_0079_r_000000
2018-07-14 20:39:09,934 [pool-221-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:39:09,934 [pool-221-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local856273166_0079_r_000000_0' done.
2018-07-14 20:39:09,934 [pool-221-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local856273166_0079_r_000000_0
2018-07-14 20:39:09,934 [Thread-2190] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:39:10,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local856273166_0079
2018-07-14 20:39:10,059 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:39:10,059 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:39:10,060 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 20:39:10,062 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,062 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,063 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,066 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:39:10,066 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:39:10,066 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:39:10,066 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:39:10,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:39:10,080 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:39:10,081 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,083 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:10,083 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:39:10,091 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:39:10,099 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:39:10,099 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:39:10,099 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:39:10,120 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:39:10,136 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local350050583_0080
2018-07-14 20:39:10,198 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:39:10,198 [Thread-2219] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:39:10,202 [Thread-2219] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:39:10,202 [Thread-2219] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:39:10,202 [Thread-2219] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:10,202 [Thread-2219] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:39:10,202 [Thread-2219] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:39:10,202 [Thread-2219] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:39:10,204 [Thread-2219] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:39:10,204 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
```

```
attempt_local350050583_0080_m_000000_0
2018-07-14 20:39:10,208 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:39:10,208 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:39:10,209 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-532868691/part-r-00000:0+62
2018-07-14 20:39:10,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:39:10,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:39:10,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:39:10,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:39:10,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:39:10,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:39:10,226 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:39:10,226 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:39:10,227 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:39:10,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:39:10,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:39:10,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:39:10,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:39:10,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:39:10,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:39:10,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local350050583_0080_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:39:10,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:39:10,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local350050583_0080_m_000000_0' done.
2018-07-14 20:39:10,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local350050583_0080_m_000000_0
2018-07-14 20:39:10,239 [Thread-2219] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
```

```
2018-07-14 20:39:10,240 [Thread-2219] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:39:10,240 [pool-224-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local350050583_0080_r_000000_0
2018-07-14 20:39:10,245 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:39:10,245 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@11c7a92
2018-07-14 20:39:10,245 [pool-224-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:39:10,349 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local350050583_0080_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:39:10,350 [localfetcher#64] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#64 about to shuffle
output of map attempt_local350050583_0080_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:39:10,350 [localfetcher#64] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local350050583_0080_m_000000_0
2018-07-14 20:39:10,350 [localfetcher#64] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:39:10,354 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:39:10,355 [pool-224-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:39:10,355 [pool-224-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:39:10,355 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:39:10,356 [pool-224-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:39:10,362 [pool-224-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:39:10,362 [pool-224-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:39:10,364 [pool-224-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:39:10,364 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Task -
```

```
Task:attempt_local350050583_0080_r_000000_0 is done. And is in the process of committing
2018-07-14 20:39:10,365 [pool-224-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:39:10,365 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local350050583_0080_r_000000_0 is allowed to commit now
2018-07-14 20:39:10,365 [pool-224-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local350050583_0080_r_000000_0' to
file:/tmp/temp-1902294536/tmp1642371036/_temporary/0/task_local350050583_0080_r_000000
2018-07-14 20:39:10,366 [pool-224-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:39:10,366 [pool-224-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local350050583_0080_r_000000_0' done.
2018-07-14 20:39:10,366 [pool-224-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local350050583_0080_r_000000_0
2018-07-14 20:39:10,366 [Thread-2219] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:39:10,551 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local350050583_0080
2018-07-14 20:39:10,552 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:39:10,552 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[46,9],ex_m[39,7] C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:39:10,556 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:39:10,561 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,563 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,564 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,568 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:39:10,568 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:39:10,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:39:10,580 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:39:10,581 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,584 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:10,584 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:39:10,591 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:39:10,599 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:39:10,599 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
2018-07-14 20:39:10,600 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:39:10,619 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:39:10,634 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2034815960_0081
2018-07-14 20:39:10,698 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:39:10,699 [Thread-2248] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:39:10,702 [Thread-2248] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:39:10,702 [Thread-2248] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:39:10,702 [Thread-2248] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:10,703 [Thread-2248] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:39:10,703 [Thread-2248] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:39:10,706 [Thread-2248] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:39:10,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2034815960_0081_m_000000_0
2018-07-14 20:39:10,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:39:10,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:39:10,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-532868691/part-r-00000:0+62
2018-07-14 20:39:10,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:39:10,718 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:39:10,719 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: ans[47,6] C:  R:
2018-07-14 20:39:10,720 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:39:10,727 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:39:10,727 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
```

```
2018-07-14 20:39:10,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:39:10,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2034815960_0081_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:39:10,731 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:39:10,731 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local2034815960_0081_m_000000_0 is allowed
to commit now
2018-07-14 20:39:10,732 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2034815960_0081_m_000000_0' to
file:/tmp/temp-1902294536/tmp1796831489/_temporary/0/task_local2034815960_0081_m_00000
0
2018-07-14 20:39:10,732 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:39:10,732 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2034815960_0081_m_000000_0' done.
2018-07-14 20:39:10,732 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2034815960_0081_m_000000_0
2018-07-14 20:39:10,733 [Thread-2248] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:39:10,901 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2034815960_0081
2018-07-14 20:39:10,901 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans
2018-07-14 20:39:10,901 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[47,6] C:  R:
2018-07-14 20:39:10,916 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,917 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,918 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,921 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:39:10,923 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:39:09     2018-07-14 20:39:10
GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local2034815960_0081        1       0       n/a     n/a     n/a     n/a     0
0       0       0       ans     MAP_ONLY
file:/tmp/temp-1902294536/tmp1796831489,
job_local350050583_0080 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local856273166_0079 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp1796831489"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local856273166_0079 ->      job_local350050583_0080,job_local2034815960_0081,
job_local350050583_0080 ->      job_local2034815960_0081,
job_local2034815960_0081


2018-07-14 20:39:10,924 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,925 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,926 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,929 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,930 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,930 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,932 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,933 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,934 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:39:10,936 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:39:10,936 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:39:10,937 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:39:10,937 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:39:10,947 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:39:10,947 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,500)
grunt> ans = JOIN details BY EmpId, ans BY empId
>> ;
2018-07-14 20:40:05,373 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 48, column 36> Invalid field projection. Projected field [empId] does not exist in
schema: EmpId:int,sum:double.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans = JOIN details BY EmpId, ans BY EmpId;
grunt> dump ans;
2018-07-14 20:40:17,672 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 20:40:17,687 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:17,687 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:40:17,688 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:40:17,688 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:40:17,690 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:40:17,691 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:40:17,692 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:40:17,693 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:40:17,693 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 20:40:17,693 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 20:40:17,693 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 20:40:17,693 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 20:40:17,701 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:17,701 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:40:17,702 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:17,703 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:40:17,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:40:17,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:40:17,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:40:17,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:40:17,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:40:17,706 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
```

```
2018-07-14 20:40:17,706 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:40:17,706 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:40:17,707 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581017706-0
2018-07-14 20:40:17,715 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:40:17,716 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:17,723 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:40:17,725 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:40:17,725 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:40:17,725 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:40:17,726 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:40:17,747 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:40:17,760 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1038115324_0082
2018-07-14 20:40:17,833 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:40:17,833 [Thread-2273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:40:17,838 [Thread-2273] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:40:17,838 [Thread-2273] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:40:17,838 [Thread-2273] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:17,838 [Thread-2273] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:40:17,838 [Thread-2273] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:40:17,838 [Thread-2273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:40:17,840 [Thread-2273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:40:17,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1038115324_0082_m_000000_0
2018-07-14 20:40:17,845 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:17,846 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
```

```
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:40:17,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:40:17,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:40:17,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:40:17,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:40:17,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:40:17,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:40:17,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:40:17,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:40:17,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:17,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:40:17,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:40:17,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:40:17,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:40:17,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:40:17,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:40:17,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:40:17,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:40:17,884 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1038115324_0082_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:40:17,884 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:40:17,884 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1038115324_0082_m_000000_0' done.
2018-07-14 20:40:17,884 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1038115324_0082_m_000000_0
2018-07-14 20:40:17,884 [Thread-2273] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:40:17,885 [Thread-2273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:40:17,885 [pool-229-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1038115324_0082_r_000000_0
2018-07-14 20:40:17,892 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:17,892 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6b9155e1
2018-07-14 20:40:17,892 [pool-229-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:40:17,893 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1038115324_0082_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:40:17,897 [localfetcher#65] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#65 about to shuffle
output of map attempt_local1038115324_0082_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:40:17,897 [localfetcher#65] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1038115324_0082_m_000000_0
2018-07-14 20:40:17,897 [localfetcher#65] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:40:17,898 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:40:17,898 [pool-229-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:40:17,898 [pool-229-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:40:17,899 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:40:17,899 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:40:17,900 [pool-229-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:40:17,901 [pool-229-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:40:17,901 [pool-229-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:40:17,901 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:40:17,901 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:40:17,902 [pool-229-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:40:17,909 [pool-229-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:17,910 [pool-229-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:40:17,911 [pool-229-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:40:17,912 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1038115324_0082_r_000000_0 is done. And is in the process of committing
2018-07-14 20:40:17,913 [pool-229-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:40:17,913 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1038115324_0082_r_000000_0 is allowed to commit now
2018-07-14 20:40:17,914 [pool-229-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1038115324_0082_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1044712746/_temporary/0/task_local1038115324_0082_r_0000
00
2018-07-14 20:40:17,914 [pool-229-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:40:17,914 [pool-229-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1038115324_0082_r_000000_0' done.
2018-07-14 20:40:17,914 [pool-229-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1038115324_0082_r_000000_0
2018-07-14 20:40:17,914 [Thread-2273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:40:18,035 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1038115324_0082
2018-07-14 20:40:18,035 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:40:18,035 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:40:18,038 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 20:40:18,042 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,044 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,046 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,051 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:40:18,052 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:40:18,053 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:40:18,053 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:40:18,056 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:40:18,071 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:40:18,072 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,073 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:18,073 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:40:18,080 [JobControl] WARN
```

```
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:40:18,090 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:40:18,090 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:40:18,090 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:40:18,109 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:40:18,123 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1455364279_0083
2018-07-14 20:40:18,198 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:40:18,199 [Thread-2302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:40:18,205 [Thread-2302] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:40:18,205 [Thread-2302] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:40:18,205 [Thread-2302] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:18,205 [Thread-2302] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:40:18,205 [Thread-2302] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:40:18,205 [Thread-2302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:40:18,208 [Thread-2302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:40:18,209 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1455364279_0083_m_000000_0
2018-07-14 20:40:18,214 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:18,215 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:40:18,217 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1044712746/part-r-00000:0+62
2018-07-14 20:40:18,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:40:18,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:40:18,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
```

```
2018-07-14 20:40:18,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:40:18,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:40:18,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:40:18,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:18,235 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:40:18,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:40:18,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:40:18,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:40:18,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:40:18,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:40:18,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:40:18,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:40:18,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1455364279_0083_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:40:18,253 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:40:18,253 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1455364279_0083_m_000000_0' done.
2018-07-14 20:40:18,253 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1455364279_0083_m_000000_0
2018-07-14 20:40:18,253 [Thread-2302] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:40:18,256 [Thread-2302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:40:18,256 [pool-232-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1455364279_0083_r_000000_0
2018-07-14 20:40:18,267 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:18,267 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3a6b6d6b
2018-07-14 20:40:18,268 [pool-232-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:40:18,270 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1455364279_0083_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:40:18,272 [localfetcher#66] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#66 about to shuffle
output of map attempt_local1455364279_0083_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:40:18,272 [localfetcher#66] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
```

```
for attempt_local1455364279_0083_m_000000_0
2018-07-14 20:40:18,272 [localfetcher#66] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:40:18,272 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:40:18,273 [pool-232-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:40:18,273 [pool-232-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:40:18,274 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:40:18,274 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:40:18,275 [pool-232-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:40:18,275 [pool-232-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:40:18,275 [pool-232-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:40:18,275 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:40:18,275 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:40:18,276 [pool-232-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:40:18,285 [pool-232-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:18,285 [pool-232-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:40:18,289 [pool-232-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:40:18,290 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1455364279_0083_r_000000_0 is done. And is in the process of committing
2018-07-14 20:40:18,292 [pool-232-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:40:18,292 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1455364279_0083_r_000000_0 is allowed to commit now
2018-07-14 20:40:18,292 [pool-232-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1455364279_0083_r_000000_0' to
file:/tmp/temp-1902294536/tmp-770156867/_temporary/0/task_local1455364279_0083_r_00000
0
2018-07-14 20:40:18,293 [pool-232-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:40:18,293 [pool-232-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1455364279_0083_r_000000_0' done.
2018-07-14 20:40:18,293 [pool-232-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1455364279_0083_r_000000_0
2018-07-14 20:40:18,293 [Thread-2302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:40:18,500 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1455364279_0083
```

```
2018-07-14 20:40:18,500 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:40:18,501 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[46,9],ex_m[39,7] C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:40:18,505 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:40:18,510 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,512 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,514 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,522 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:40:18,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:40:18,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:40:18,523 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:40:18,533 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 20:40:18,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:40:18,539 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:40:18,568 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:40:18,570 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:18,573 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:18,573 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:40:18,579 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:40:18,586 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:40:18,586 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:40:18,587 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:40:18,589 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:40:18,589 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
```

```
2018-07-14 20:40:18,589 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:40:18,589 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:40:18,610 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:40:18,623 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1997807376_0084
2018-07-14 20:40:18,686 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:40:18,687 [Thread-2331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:40:18,690 [Thread-2331] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:40:18,690 [Thread-2331] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:40:18,691 [Thread-2331] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:18,691 [Thread-2331] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:40:18,691 [Thread-2331] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:40:18,691 [Thread-2331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:40:18,693 [Thread-2331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:40:18,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1997807376_0084_m_000000_0
2018-07-14 20:40:18,697 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:18,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:40:18,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:40:18,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:40:18,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:40:18,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:40:18,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:40:18,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
```

```
2018-07-14 20:40:18,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:40:18,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:40:18,725 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:18,725 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:40:18,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:40:18,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:40:18,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:40:18,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:40:18,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:40:18,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:40:18,740 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:40:18,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1997807376_0084_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:40:18,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:40:18,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1997807376_0084_m_000000_0' done.
2018-07-14 20:40:18,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1997807376_0084_m_000000_0
2018-07-14 20:40:18,742 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1997807376_0084_m_000001_0
2018-07-14 20:40:18,743 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:18,744 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:40:18,745 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1044712746/part-r-00000:0+62
2018-07-14 20:40:18,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:40:18,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:40:18,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:40:18,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
```

```
2018-07-14 20:40:18,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:40:18,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:40:18,765 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:18,765 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:40:18,767 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:40:18,767 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:40:18,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:40:18,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:40:18,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:40:18,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:40:18,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:40:18,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 12; bufvoid = 104857600
2018-07-14 20:40:18,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 20:40:18,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:40:18,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1997807376_0084_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:40:18,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:40:18,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1997807376_0084_m_000001_0' done.
2018-07-14 20:40:18,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1997807376_0084_m_000001_0
2018-07-14 20:40:18,778 [Thread-2331] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:40:18,778 [Thread-2331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:40:18,779 [pool-235-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1997807376_0084_r_000000_0
2018-07-14 20:40:18,785 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:40:18,785 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7360abd0
2018-07-14 20:40:18,786 [pool-235-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:40:18,788 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1997807376_0084_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
```

```
2018-07-14 20:40:18,789 [localfetcher#67] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#67 about to shuffle
output of map attempt_local1997807376_0084_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:40:18,789 [localfetcher#67] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1997807376_0084_m_000000_0
2018-07-14 20:40:18,789 [localfetcher#67] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 20:40:18,790 [localfetcher#67] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#67 about to shuffle
output of map attempt_local1997807376_0084_m_000001_0 decomp: 16 len: 20 to MEMORY
2018-07-14 20:40:18,790 [localfetcher#67] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 16 bytes from map-output
for attempt_local1997807376_0084_m_000001_0
2018-07-14 20:40:18,790 [localfetcher#67] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 16, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->368
2018-07-14 20:40:18,791 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:40:18,792 [pool-235-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:40:18,792 [pool-235-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:40:18,793 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:40:18,794 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 352 bytes
2018-07-14 20:40:18,794 [pool-235-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 368 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:40:18,794 [pool-235-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 370 bytes from
disk
2018-07-14 20:40:18,794 [pool-235-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:40:18,794 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:40:18,795 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 358 bytes
2018-07-14 20:40:18,795 [pool-235-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:40:18,801 [pool-235-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:40:18,801 [pool-235-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:40:18,802 [pool-235-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:40:18,805 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1997807376_0084_r_000000_0 is done. And is in the process of committing
2018-07-14 20:40:18,806 [pool-235-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:40:18,806 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1997807376_0084_r_000000_0 is allowed to commit now
2018-07-14 20:40:18,807 [pool-235-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
```

```
'attempt_local1997807376_0084_r_000000_0' to
file:/tmp/temp-1902294536/tmp157831995/_temporary/0/task_local1997807376_0084_r_000000
2018-07-14 20:40:18,808 [pool-235-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:40:18,808 [pool-235-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1997807376_0084_r_000000_0' done.
2018-07-14 20:40:18,808 [pool-235-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1997807376_0084_r_000000_0
2018-07-14 20:40:18,808 [Thread-2331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:40:18,989 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1997807376_0084
2018-07-14 20:40:18,989 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,details
2018-07-14 20:40:18,990 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:40:19,003 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,007 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,010 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,028 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:40:19,030 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:40:17     2018-07-14 20:40:19
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1038115324_0082        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1455364279_0083        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local1997807376_0084        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans,details     HASH_JOIN
file:/tmp/temp-1902294536/tmp157831995,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp157831995"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
```

```
Job DAG:
job_local1038115324_0082        ->
job_local1455364279_0083,job_local1997807376_0084,
job_local1455364279_0083        ->      job_local1997807376_0084,
job_local1997807376_0084


2018-07-14 20:40:19,031 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,032 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,033 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,039 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,039 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,040 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,044 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,058 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,058 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:40:19,059 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:40:19,059 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:40:19,060 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:40:19,060 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:40:19,068 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:40:19,068 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,10000,2,102,500)
grunt>
grunt>
grunt>
grunt> ans2 = foreach ans generate EmpId,Name,Expense;
2018-07-14 20:42:54,220 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 49, column 28> Invalid field projection. Projected field [EmpId] does not exist in
schema:
details::EmpId:int,details::Name:chararray,details::Salary:int,details::Dept:int,ans::
EmpId:int,ans::sum:double.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe ans;
ans: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ans::EmpId: int,ans::sum: long}
grunt> ans2 = foreach ans generate details.EmpId,Name,Expense;
2018-07-14 20:43:14,400 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 49, column 47> Invalid field projection. Projected field [Expense] does not exist
in schema:
details::EmpId:int,details::Name:chararray,details::Salary:int,details::Dept:int,ans::
EmpId:int,ans::sum:double.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans2 = foreach ans generate details.EmpId,Name,sum;
grunt> dump ans2;
```

```
2018-07-14 20:43:29,602 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 20:43:29,616 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:29,616 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:43:29,617 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:43:29,617 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:43:29,620 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:43:29,622 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:43:29,622 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:43:29,624 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:43:29,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 5
2018-07-14 20:43:29,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 20:43:29,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 20:43:29,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 20:43:29,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 20:43:29,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 4
2018-07-14 20:43:29,633 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:29,633 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:43:29,633 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:29,635 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:29,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:29,636 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 20:43:29,636 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:43:29,636 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
```

```
2018-07-14 20:43:29,636 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581209636-0
2018-07-14 20:43:29,643 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:29,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:29,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:43:29,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:43:29,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:43:29,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:43:29,647 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:43:29,647 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:43:29,647 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:43:29,647 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581209647-0
2018-07-14 20:43:29,657 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 2
map-reduce job(s) waiting for submission.
2018-07-14 20:43:29,659 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:29,671 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:29,673 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:43:29,673 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:29,673 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:29,673 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:29,693 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:43:29,708 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1408952110_0085
2018-07-14 20:43:29,777 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:29,778 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:29,779 [Thread-2363] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:29,781 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
```

```
2018-07-14 20:43:29,781 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:29,781 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:29,781 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:29,785 [Thread-2363] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:29,787 [Thread-2363] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:29,788 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1408952110_0085_m_000000_0
2018-07-14 20:43:29,793 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:29,794 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:43:29,803 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:29,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:43:29,804 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:43:29,804 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:29,804 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:29,804 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:29,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:29,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:43:29,811 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],details[-1,-1]
C:  R:
2018-07-14 20:43:29,833 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:29,833 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1408952110_0085_m_000000_0 is done. And
is in the process of committing
```

```
2018-07-14 20:43:29,835 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:29,835 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1408952110_0085_m_000000_0 is allowed
to commit now
2018-07-14 20:43:29,836 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1408952110_0085_m_000000_0' to
file:/tmp/temp-1902294536/tmp-445499220/_temporary/0/task_local1408952110_0085_m_00000
0
2018-07-14 20:43:29,836 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1408952110_0085_m_000000_0' to
file:/tmp/temp-1902294536/tmp2025899979/_temporary/0/task_local1408952110_0085_m_00000
0
2018-07-14 20:43:29,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:29,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1408952110_0085_m_000000_0' done.
2018-07-14 20:43:29,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1408952110_0085_m_000000_0
2018-07-14 20:43:29,837 [Thread-2363] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:43:29,846 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:43:29,862 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:29,862 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:29,869 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local4885161_0086
2018-07-14 20:43:29,937 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:29,938 [Thread-2388] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:29,941 [Thread-2388] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:43:29,941 [Thread-2388] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:29,941 [Thread-2388] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:29,941 [Thread-2388] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:43:29,941 [Thread-2388] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:29,941 [Thread-2388] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:29,943 [Thread-2388] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:29,943 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local4885161_0086_m_000000_0
2018-07-14 20:43:29,948 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:29,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:29,950 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:43:29,950 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:43:29,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:29,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:43:29,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:29,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:29,978 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:29,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:29,981 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:29,981 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:29,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:43:29,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:29,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:29,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:29,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:43:29,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:43:29,989 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:29,990 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local4885161_0086_m_000000_0 is done. And is
in the process of committing
2018-07-14 20:43:29,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:29,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local4885161_0086_m_000000_0' done.
2018-07-14 20:43:29,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local4885161_0086_m_000000_0
2018-07-14 20:43:29,991 [Thread-2388] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
```

```
task executor complete.
2018-07-14 20:43:29,991 [Thread-2388] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:43:29,992 [pool-240-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local4885161_0086_r_000000_0
2018-07-14 20:43:29,997 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:29,997 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@68c9cd72
2018-07-14 20:43:29,998 [pool-240-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:43:29,999 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local4885161_0086_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:43:30,000 [localfetcher#68] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#68 about to shuffle
output of map attempt_local4885161_0086_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:43:30,000 [localfetcher#68] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local4885161_0086_m_000000_0
2018-07-14 20:43:30,000 [localfetcher#68] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:43:30,001 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:43:30,001 [pool-240-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:30,001 [pool-240-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:43:30,002 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:30,002 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:43:30,003 [pool-240-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:43:30,003 [pool-240-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:43:30,003 [pool-240-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:43:30,003 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:30,003 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:43:30,003 [pool-240-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:30,009 [pool-240-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:30,009 [pool-240-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,011 [pool-240-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
```

```
ex_sum[38,8]
2018-07-14 20:43:30,012 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local4885161_0086_r_000000_0 is done. And is in the process of committing
2018-07-14 20:43:30,013 [pool-240-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:30,013 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local4885161_0086_r_000000_0 is allowed to commit now
2018-07-14 20:43:30,014 [pool-240-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local4885161_0086_r_000000_0' to
file:/tmp/temp-1902294536/tmp972257833/_temporary/0/task_local4885161_0086_r_000000
2018-07-14 20:43:30,015 [pool-240-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:43:30,015 [pool-240-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local4885161_0086_r_000000_0' done.
2018-07-14 20:43:30,015 [pool-240-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local4885161_0086_r_000000_0
2018-07-14 20:43:30,015 [Thread-2388] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:43:30,141 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1408952110_0085
2018-07-14 20:43:30,141 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 20:43:30,141 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1] C:  R:
2018-07-14 20:43:30,143 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local4885161_0086
2018-07-14 20:43:30,143 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:43:30,143 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:43:30,148 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:43:30,150 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,152 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,153 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,157 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,158 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,159 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,166 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:30,167 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:30,167 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
```

```
2018-07-14 20:43:30,167 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:43:30,168 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:43:30,183 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:43:30,184 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,186 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:30,186 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:30,193 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:30,202 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:30,202 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:30,202 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:30,221 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:43:30,234 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1991890699_0087
2018-07-14 20:43:30,294 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:30,294 [Thread-2417] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:30,299 [Thread-2417] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:43:30,299 [Thread-2417] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:30,299 [Thread-2417] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:30,299 [Thread-2417] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:43:30,299 [Thread-2417] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:30,299 [Thread-2417] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:30,301 [Thread-2417] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:30,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1991890699_0087_m_000000_0
2018-07-14 20:43:30,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:30,306 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:30,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp972257833/part-r-00000:0+62
2018-07-14 20:43:30,323 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:30,323 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:43:30,323 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:30,323 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:30,323 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:30,324 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:30,325 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:30,325 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:43:30,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:30,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:30,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:30,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:43:30,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:43:30,335 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:30,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1991890699_0087_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:43:30,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:30,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1991890699_0087_m_000000_0' done.
2018-07-14 20:43:30,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1991890699_0087_m_000000_0
2018-07-14 20:43:30,336 [Thread-2417] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:43:30,337 [Thread-2417] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:43:30,337 [pool-243-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
```

```
attempt_local1991890699_0087_r_000000_0
2018-07-14 20:43:30,341 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:30,341 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3bff1186
2018-07-14 20:43:30,342 [pool-243-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:43:30,343 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1991890699_0087_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:43:30,344 [localfetcher#69] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#69 about to shuffle
output of map attempt_local1991890699_0087_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:43:30,344 [localfetcher#69] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1991890699_0087_m_000000_0
2018-07-14 20:43:30,344 [localfetcher#69] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:43:30,344 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:43:30,345 [pool-243-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:30,345 [pool-243-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:43:30,345 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:30,345 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:43:30,348 [pool-243-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:43:30,348 [pool-243-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:43:30,348 [pool-243-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:43:30,348 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:30,348 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:43:30,348 [pool-243-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:30,484 [pool-243-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:30,484 [pool-243-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,486 [pool-243-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:43:30,486 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1991890699_0087_r_000000_0 is done. And is in the process of committing
2018-07-14 20:43:30,487 [pool-243-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:30,487 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local1991890699_0087_r_000000_0 is allowed to commit now
2018-07-14 20:43:30,488 [pool-243-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1991890699_0087_r_000000_0' to
file:/tmp/temp-1902294536/tmp-748442288/_temporary/0/task_local1991890699_0087_r_00000
0
2018-07-14 20:43:30,488 [pool-243-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:43:30,488 [pool-243-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1991890699_0087_r_000000_0' done.
2018-07-14 20:43:30,488 [pool-243-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1991890699_0087_r_000000_0
2018-07-14 20:43:30,488 [Thread-2417] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:43:30,685 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1991890699_0087
2018-07-14 20:43:30,685 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:43:30,685 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[46,9],ex_m[39,7] C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:43:30,688 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 75%
complete
2018-07-14 20:43:30,692 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,694 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,695 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,699 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:30,700 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:30,701 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:43:30,701 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:43:30,703 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=342
2018-07-14 20:43:30,703 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:43:30,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:43:30,718 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:43:30,719 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:30,720 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 20:43:30,720 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:30,727 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:30,734 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:30,734 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:30,734 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:30,742 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:30,742 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:30,742 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:30,762 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:43:30,777 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1823028922_0088
2018-07-14 20:43:30,837 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:30,838 [Thread-2448] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:30,841 [Thread-2448] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:43:30,841 [Thread-2448] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:30,841 [Thread-2448] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:30,841 [Thread-2448] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:43:30,841 [Thread-2448] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:30,841 [Thread-2448] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:30,843 [Thread-2448] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:30,843 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1823028922_0088_m_000000_0
2018-07-14 20:43:30,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:30,849 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 280
Input split[0]:
   Length = 280
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:
```

```
-----------------------

2018-07-14 20:43:30,850 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp2025899979/part-m-00000:0+280
2018-07-14 20:43:30,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:30,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:43:30,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:30,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:30,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:30,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:30,867 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:30,867 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[48,6],ans[47,6],ans[48,6] C:  R: ans2[49,7]
2018-07-14 20:43:30,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:30,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:30,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:30,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:43:30,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:43:30,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:30,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1823028922_0088_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:43:30,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:30,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1823028922_0088_m_000000_0' done.
2018-07-14 20:43:30,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1823028922_0088_m_000000_0
2018-07-14 20:43:30,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1823028922_0088_m_000001_0
2018-07-14 20:43:30,874 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:30,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:
```

```
-----------------------

2018-07-14 20:43:30,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp972257833/part-r-00000:0+62
2018-07-14 20:43:30,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:30,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:43:30,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:30,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:30,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:30,894 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:30,895 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:30,895 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[48,6],ans[47,6],ans[48,6] C:  R: ans2[49,7]
2018-07-14 20:43:30,897 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,905 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:30,905 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:30,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:30,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:30,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:30,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 12; bufvoid = 104857600
2018-07-14 20:43:30,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 20:43:30,907 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:30,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1823028922_0088_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:43:30,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:30,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1823028922_0088_m_000001_0' done.
2018-07-14 20:43:30,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1823028922_0088_m_000001_0
2018-07-14 20:43:30,909 [Thread-2448] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:43:30,909 [Thread-2448] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:43:30,909 [pool-246-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1823028922_0088_r_000000_0
2018-07-14 20:43:30,918 [pool-246-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:30,918 [pool-246-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@72800a9
2018-07-14 20:43:30,919 [pool-246-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:43:30,920 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1823028922_0088_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:43:30,921 [localfetcher#70] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#70 about to shuffle
output of map attempt_local1823028922_0088_m_000001_0 decomp: 16 len: 20 to MEMORY
2018-07-14 20:43:30,921 [localfetcher#70] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 16 bytes from map-output
for attempt_local1823028922_0088_m_000001_0
2018-07-14 20:43:30,921 [localfetcher#70] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 16, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->16
2018-07-14 20:43:30,922 [localfetcher#70] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#70 about to shuffle
output of map attempt_local1823028922_0088_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:43:30,922 [localfetcher#70] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1823028922_0088_m_000000_0
2018-07-14 20:43:30,922 [localfetcher#70] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 16, usedMemory
->368
2018-07-14 20:43:30,925 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:43:30,932 [pool-246-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:43:30,932 [pool-246-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:43:30,933 [pool-246-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:43:30,933 [pool-246-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 352 bytes
2018-07-14 20:43:30,934 [pool-246-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 368 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:43:30,934 [pool-246-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 370 bytes from
disk
2018-07-14 20:43:30,934 [pool-246-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:43:30,934 [pool-246-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:30,934 [pool-246-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 358 bytes
2018-07-14 20:43:30,935 [pool-246-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:43:30,941 [pool-246-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 20:43:30,941 [pool-246-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,943 [pool-246-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[48,6],ans[47,6],ans[48,6] C:  R: ans2[49,7]
2018-07-14 20:43:30,943 [pool-246-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:30,952 [pool-246-thread-1] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:30,952 [pool-246-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:30,952 [Thread-2448] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:43:30,954 [Thread-2448] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local1823028922_0088
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,Amitabh,20000,1), 2nd
:(102,Shahrukh,10000,2) (common cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should
be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:529)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,Amitabh,20000,1), 2nd :(102,Shahrukh,10000,2) (common
cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextInteger(POUserFunc.java:375)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:347)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.processPlan(POForEach.java:404)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:321)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.runPipeline(PigGenericMapReduce.java:465)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.processOnePackageOutput(PigGenericMapReduce.java:433)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:405)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:262)
        at org.apache.hadoop.mapreduce.Reducer.run(Reducer.java:171)
        at org.apache.hadoop.mapred.ReduceTask.runNewReducer(ReduceTask.java:627)
        at org.apache.hadoop.mapred.ReduceTask.run(ReduceTask.java:389)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$ReduceTaskRunnable.run(LocalJobRunner.java
:319)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
```

```
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
2018-07-14 20:43:31,139 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1823028922_0088
2018-07-14 20:43:31,139 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,ans2
2018-07-14 20:43:31,139 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[48,6],ans[47,6],ans[48,6] C:  R: ans2[49,7]
2018-07-14 20:43:31,144 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 20:43:31,144 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local1823028922_0088 has failed! Stop running all dependent jobs
2018-07-14 20:43:31,144 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:43:31,150 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,152 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,152 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 20:43:31,153 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:43:29     2018-07-14 20:43:31
HASH_JOIN,GROUP_BY,FILTER

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1408952110_0085        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details MULTI_QUERY,MAP_ONLY
job_local1991890699_0087        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local4885161_0086   1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local1823028922_0088        ans,ans2        HASH_JOIN       Message: Job failed!
file:/tmp/temp-1902294536/tmp1465792008,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp1465792008"

Counters:
Total records written : 0
```

```
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1408952110_0085        ->      job_local1823028922_0088,
job_local4885161_0086   ->      job_local1991890699_0087,job_local1823028922_0088,
job_local1991890699_0087        ->      job_local1823028922_0088,
job_local1823028922_0088


2018-07-14 20:43:31,155 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,156 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,158 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,162 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,164 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,165 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,168 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,169 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,169 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:31,171 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 20:43:31,172 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias ans2
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt>
grunt> dump ans2;2018-07-14 20:43:36,913 [communication thread] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
ans = JOIN details BY EmpId, ans BY empId
>>
>>
>>
>>
>>
>> ;
2018-07-14 20:43:44,548 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:44,548 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:43:44,560 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 50, column 36> Invalid field projection. Projected field [empId] does not exist in
schema:
details::EmpId:int,details::Name:chararray,details::Salary:int,details::Dept:int,ans::
EmpId:int,ans::sum:double.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt>
```

```
grunt> dump ans;
2018-07-14 20:43:47,526 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 20:43:47,541 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:47,541 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:43:47,542 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:43:47,542 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:43:47,547 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:43:47,548 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:43:47,548 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:43:47,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:43:47,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 20:43:47,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 20:43:47,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 20:43:47,550 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 20:43:47,558 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:47,558 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:43:47,558 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:47,560 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:47,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:47,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:43:47,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:43:47,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:43:47,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Setting Parallelism to 1
2018-07-14 20:43:47,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:43:47,562 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:43:47,562 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:43:47,562 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581227561-0
2018-07-14 20:43:47,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:43:47,571 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:47,579 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:47,582 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:43:47,585 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:47,585 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:47,586 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:47,605 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:43:47,621 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1704337873_0089
2018-07-14 20:43:47,688 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:47,689 [Thread-2478] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:47,692 [Thread-2478] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:43:47,692 [Thread-2478] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:47,692 [Thread-2478] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:47,692 [Thread-2478] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:43:47,692 [Thread-2478] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:47,692 [Thread-2478] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:47,694 [Thread-2478] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:47,694 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1704337873_0089_m_000000_0
2018-07-14 20:43:47,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
```

```
2018-07-14 20:43:47,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:47,701 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:43:47,701 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:43:47,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:47,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:43:47,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:47,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:47,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:47,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:47,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:47,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:43:47,719 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:43:47,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:47,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:47,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:47,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:43:47,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:43:47,727 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:47,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1704337873_0089_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:43:47,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:47,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1704337873_0089_m_000000_0' done.
2018-07-14 20:43:47,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1704337873_0089_m_000000_0
2018-07-14 20:43:47,731 [Thread-2478] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
```

```
task executor complete.
2018-07-14 20:43:47,732 [Thread-2478] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:43:47,732 [pool-249-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1704337873_0089_r_000000_0
2018-07-14 20:43:47,739 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:47,740 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@66ce5b67
2018-07-14 20:43:47,741 [pool-249-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:43:47,742 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1704337873_0089_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:43:47,750 [localfetcher#71] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#71 about to shuffle
output of map attempt_local1704337873_0089_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:43:47,750 [localfetcher#71] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1704337873_0089_m_000000_0
2018-07-14 20:43:47,750 [localfetcher#71] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:43:47,751 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:43:47,752 [pool-249-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:47,752 [pool-249-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:43:47,753 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:47,753 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:43:47,754 [pool-249-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:43:47,754 [pool-249-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:43:47,754 [pool-249-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:43:47,754 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:47,754 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:43:47,754 [pool-249-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:47,762 [pool-249-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:47,762 [pool-249-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:47,766 [pool-249-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
```

```
ex_sum[38,8]
2018-07-14 20:43:47,766 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1704337873_0089_r_000000_0 is done. And is in the process of committing
2018-07-14 20:43:47,767 [pool-249-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:47,767 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1704337873_0089_r_000000_0 is allowed to commit now
2018-07-14 20:43:47,768 [pool-249-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1704337873_0089_r_000000_0' to
file:/tmp/temp-1902294536/tmp1218562418/_temporary/0/task_local1704337873_0089_r_00000
0
2018-07-14 20:43:47,768 [pool-249-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:43:47,768 [pool-249-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1704337873_0089_r_000000_0' done.
2018-07-14 20:43:47,768 [pool-249-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1704337873_0089_r_000000_0
2018-07-14 20:43:47,768 [Thread-2478] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:43:47,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1704337873_0089
2018-07-14 20:43:47,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:43:47,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:43:47,892 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 20:43:47,896 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:47,900 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:47,903 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:47,908 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:47,908 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:47,909 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:43:47,909 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:43:47,912 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:43:47,925 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:43:47,927 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:47,928 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 20:43:47,928 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:47,934 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:47,942 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:47,942 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:47,942 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:47,962 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:43:47,975 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1121388564_0090
2018-07-14 20:43:48,038 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:48,039 [Thread-2507] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:48,043 [Thread-2507] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:43:48,043 [Thread-2507] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:48,043 [Thread-2507] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:48,043 [Thread-2507] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:43:48,043 [Thread-2507] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:48,043 [Thread-2507] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:48,045 [Thread-2507] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:48,046 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1121388564_0090_m_000000_0
2018-07-14 20:43:48,052 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:48,053 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:48,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1218562418/part-r-00000:0+62
2018-07-14 20:43:48,090 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 20:43:48,090 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:43:48,090 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:48,090 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:48,090 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:48,091 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:48,092 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:48,093 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,094 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:43:48,095 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:48,095 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:48,095 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:48,095 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:43:48,095 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:43:48,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:48,104 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1121388564_0090_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:43:48,105 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:48,105 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1121388564_0090_m_000000_0' done.
2018-07-14 20:43:48,105 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1121388564_0090_m_000000_0
2018-07-14 20:43:48,105 [Thread-2507] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:43:48,106 [Thread-2507] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:43:48,106 [pool-252-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1121388564_0090_r_000000_0
2018-07-14 20:43:48,117 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:48,117 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@35a72f21
2018-07-14 20:43:48,118 [pool-252-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:43:48,120 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1121388564_0090_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:43:48,123 [localfetcher#72] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#72 about to shuffle
output of map attempt_local1121388564_0090_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:43:48,123 [localfetcher#72] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1121388564_0090_m_000000_0
2018-07-14 20:43:48,123 [localfetcher#72] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:43:48,124 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:43:48,125 [pool-252-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:48,125 [pool-252-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:43:48,126 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:48,126 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:43:48,127 [pool-252-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:43:48,128 [pool-252-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:43:48,128 [pool-252-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:43:48,128 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:48,128 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:43:48,128 [pool-252-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:48,139 [pool-252-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:48,139 [pool-252-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,142 [pool-252-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:43:48,145 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1121388564_0090_r_000000_0 is done. And is in the process of committing
2018-07-14 20:43:48,146 [pool-252-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:43:48,146 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1121388564_0090_r_000000_0 is allowed to commit now
2018-07-14 20:43:48,150 [pool-252-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1121388564_0090_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1605383280/_temporary/0/task_local1121388564_0090_r_0000
00
2018-07-14 20:43:48,151 [pool-252-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:43:48,151 [pool-252-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1121388564_0090_r_000000_0' done.
2018-07-14 20:43:48,151 [pool-252-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1121388564_0090_r_000000_0
2018-07-14 20:43:48,151 [Thread-2507] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
reduce task executor complete.
2018-07-14 20:43:48,342 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1121388564_0090
2018-07-14 20:43:48,342 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:43:48,342 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[46,9],ex_m[39,7] C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:43:48,345 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:43:48,357 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,361 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,364 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,372 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:43:48,373 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:43:48,374 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:43:48,374 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:43:48,377 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 20:43:48,378 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:43:48,384 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:43:48,410 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:43:48,412 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,415 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:48,415 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:48,423 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:43:48,430 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:48,430 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:48,430 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
```

```
2018-07-14 20:43:48,433 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:43:48,433 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:48,433 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:48,433 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:43:48,454 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:43:48,468 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local258449096_0091
2018-07-14 20:43:48,526 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:43:48,527 [Thread-2536] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:43:48,530 [Thread-2536] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:43:48,530 [Thread-2536] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:43:48,530 [Thread-2536] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:48,530 [Thread-2536] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:43:48,530 [Thread-2536] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:43:48,530 [Thread-2536] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:43:48,533 [Thread-2536] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:43:48,533 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local258449096_0091_m_000000_0
2018-07-14 20:43:48,538 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:48,538 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:48,540 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:43:48,540 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:43:48,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:48,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 20:43:48,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:48,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:48,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:48,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:48,558 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:48,558 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,560 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:43:48,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:48,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:48,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:48,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:43:48,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:43:48,562 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:48,563 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local258449096_0091_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:43:48,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:48,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local258449096_0091_m_000000_0' done.
2018-07-14 20:43:48,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local258449096_0091_m_000000_0
2018-07-14 20:43:48,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local258449096_0091_m_000001_0
2018-07-14 20:43:48,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:48,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:43:48,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1218562418/part-r-00000:0+62
2018-07-14 20:43:48,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:43:48,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 20:43:48,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:43:48,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:43:48,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:43:48,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:43:48,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:48,698 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,700 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:43:48,700 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:48,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:43:48,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:43:48,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:43:48,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:43:48,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 12; bufvoid = 104857600
2018-07-14 20:43:48,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 20:43:48,708 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:43:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local258449096_0091_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:43:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:43:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local258449096_0091_m_000001_0' done.
2018-07-14 20:43:48,709 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local258449096_0091_m_000001_0
2018-07-14 20:43:48,709 [Thread-2536] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:43:48,710 [Thread-2536] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:43:48,710 [pool-255-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local258449096_0091_r_000000_0
2018-07-14 20:43:48,715 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:43:48,715 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@41c49ddb
2018-07-14 20:43:48,716 [pool-255-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
```

```
2018-07-14 20:43:48,717 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local258449096_0091_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:43:48,719 [localfetcher#73] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#73 about to shuffle
output of map attempt_local258449096_0091_m_000001_0 decomp: 16 len: 20 to MEMORY
2018-07-14 20:43:48,719 [localfetcher#73] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 16 bytes from map-output
for attempt_local258449096_0091_m_000001_0
2018-07-14 20:43:48,719 [localfetcher#73] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 16, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->16
2018-07-14 20:43:48,720 [localfetcher#73] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#73 about to shuffle
output of map attempt_local258449096_0091_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:43:48,720 [localfetcher#73] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local258449096_0091_m_000000_0
2018-07-14 20:43:48,720 [localfetcher#73] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 16, usedMemory
->368
2018-07-14 20:43:48,721 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:43:48,721 [pool-255-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:43:48,721 [pool-255-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:43:48,722 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:43:48,722 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 352 bytes
2018-07-14 20:43:48,723 [pool-255-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 368 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:43:48,723 [pool-255-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 370 bytes from
disk
2018-07-14 20:43:48,723 [pool-255-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:43:48,723 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:43:48,723 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 358 bytes
2018-07-14 20:43:48,723 [pool-255-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:43:48,729 [pool-255-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:43:48,729 [pool-255-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,730 [pool-255-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:43:48,730 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local258449096_0091_r_000000_0 is done. And is in the process of committing
2018-07-14 20:43:48,731 [pool-255-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
```

```
2018-07-14 20:43:48,732 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local258449096_0091_r_000000_0 is allowed to commit now
2018-07-14 20:43:48,733 [pool-255-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local258449096_0091_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1821682154/_temporary/0/task_local258449096_0091_r_00000
0
2018-07-14 20:43:48,733 [pool-255-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:43:48,733 [pool-255-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local258449096_0091_r_000000_0' done.
2018-07-14 20:43:48,733 [pool-255-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local258449096_0091_r_000000_0
2018-07-14 20:43:48,734 [Thread-2536] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:43:48,884 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local258449096_0091
2018-07-14 20:43:48,885 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,details
2018-07-14 20:43:48,885 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1],ans[48,6],ans[47,6],ans[48,6] C:  R:
2018-07-14 20:43:48,886 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,887 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,887 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,890 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:43:48,892 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:43:47     2018-07-14 20:43:48
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1121388564_0090        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local1704337873_0089        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local258449096_0091 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans,details     HASH_JOIN
file:/tmp/temp-1902294536/tmp-1821682154,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1821682154"

Counters:
```

```
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1704337873_0089        ->
job_local1121388564_0090,job_local258449096_0091,
job_local1121388564_0090        ->      job_local258449096_0091,
job_local258449096_0091


2018-07-14 20:43:48,893 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,894 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,894 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,895 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,896 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,896 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,897 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,897 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,898 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:43:48,899 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:43:48,899 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:43:48,899 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:43:48,899 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:43:48,907 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:43:48,907 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,10000,2,102,500)
grunt> describe ans;
ans: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ans::EmpId: int,ans::sum: long}
grunt>
grunt> ans1 = foreach ans generate $0,$1,ans.sum;
grunt> dump ans1;
2018-07-14 20:44:56,376 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 20:44:56,398 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:56,398 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:44:56,398 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:44:56,399 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
```

```
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:44:56,402 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:44:56,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:44:56,403 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 5
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 20:44:56,404 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 4
2018-07-14 20:44:56,413 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:56,414 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:44:56,414 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:56,415 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:44:56,416 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:44:56,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:44:56,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:44:56,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:44:56,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:44:56,418 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:44:56,418 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
```

```
2018-07-14 20:44:56,418 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:44:56,418 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581296418-0
2018-07-14 20:44:56,429 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:44:56,430 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:56,439 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:44:56,441 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:44:56,441 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:56,442 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:56,442 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:44:56,461 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:44:56,474 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1248339603_0092
2018-07-14 20:44:56,539 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:44:56,539 [Thread-2568] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:44:56,543 [Thread-2568] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:44:56,543 [Thread-2568] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:44:56,543 [Thread-2568] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:56,543 [Thread-2568] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:44:56,543 [Thread-2568] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:56,543 [Thread-2568] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:44:56,545 [Thread-2568] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:44:56,545 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1248339603_0092_m_000000_0
2018-07-14 20:44:56,551 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:56,551 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
```

```
   Locations:

-----------------------

2018-07-14 20:44:56,553 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:44:56,553 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:44:56,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:44:56,645 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:44:56,645 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:44:56,645 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:44:56,645 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:44:56,645 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:44:56,646 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:56,646 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:44:56,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:44:56,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:44:56,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:44:56,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:44:56,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:44:56,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:44:56,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:44:56,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1248339603_0092_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:44:56,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:44:56,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1248339603_0092_m_000000_0' done.
2018-07-14 20:44:56,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1248339603_0092_m_000000_0
2018-07-14 20:44:56,656 [Thread-2568] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:44:56,656 [Thread-2568] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:44:56,656 [pool-258-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1248339603_0092_r_000000_0
```

```
2018-07-14 20:44:56,662 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:56,662 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@598eac2d
2018-07-14 20:44:56,663 [pool-258-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:44:56,665 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1248339603_0092_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:44:56,666 [localfetcher#74] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#74 about to shuffle
output of map attempt_local1248339603_0092_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:44:56,666 [localfetcher#74] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1248339603_0092_m_000000_0
2018-07-14 20:44:56,666 [localfetcher#74] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:44:56,666 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:44:56,667 [pool-258-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:44:56,667 [pool-258-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:44:56,668 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:44:56,668 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:44:56,669 [pool-258-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:44:56,669 [pool-258-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:44:56,669 [pool-258-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:44:56,669 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:44:56,669 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:44:56,669 [pool-258-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:44:56,675 [pool-258-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:56,675 [pool-258-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:44:56,677 [pool-258-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:44:56,678 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1248339603_0092_r_000000_0 is done. And is in the process of committing
2018-07-14 20:44:56,679 [pool-258-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:44:56,679 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local1248339603_0092_r_000000_0 is allowed to commit now
2018-07-14 20:44:56,680 [pool-258-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1248339603_0092_r_000000_0' to
file:/tmp/temp-1902294536/tmp609880736/_temporary/0/task_local1248339603_0092_r_000000
2018-07-14 20:44:56,680 [pool-258-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:44:56,680 [pool-258-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1248339603_0092_r_000000_0' done.
2018-07-14 20:44:56,680 [pool-258-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1248339603_0092_r_000000_0
2018-07-14 20:44:56,680 [Thread-2568] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:44:56,843 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1248339603_0092
2018-07-14 20:44:56,843 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:44:56,843 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:44:56,844 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 25%
complete
2018-07-14 20:44:56,845 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:56,845 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:56,846 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:56,850 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:44:56,851 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:44:56,851 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:44:56,851 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:44:56,852 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:44:56,864 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:44:56,866 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:56,868 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:56,868 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:56,874 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:44:56,882 [JobControl] INFO
```

```
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:56,882 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:56,882 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:44:56,904 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:44:56,920 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1273406386_0093
2018-07-14 20:44:56,981 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:44:56,981 [Thread-2597] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:44:56,985 [Thread-2597] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:44:56,985 [Thread-2597] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:44:56,985 [Thread-2597] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:56,985 [Thread-2597] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:44:56,985 [Thread-2597] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:56,985 [Thread-2597] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:44:56,987 [Thread-2597] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:44:56,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1273406386_0093_m_000000_0
2018-07-14 20:44:56,992 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:56,992 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:44:56,993 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp609880736/part-r-00000:0+62
2018-07-14 20:44:57,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:44:57,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:44:57,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:44:57,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:44:57,008 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:44:57,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:44:57,010 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:57,010 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:44:57,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:44:57,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:44:57,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:44:57,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 20:44:57,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:44:57,021 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:44:57,021 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1273406386_0093_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:44:57,022 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:44:57,022 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1273406386_0093_m_000000_0' done.
2018-07-14 20:44:57,022 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1273406386_0093_m_000000_0
2018-07-14 20:44:57,022 [Thread-2597] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:44:57,023 [Thread-2597] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:44:57,023 [pool-261-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1273406386_0093_r_000000_0
2018-07-14 20:44:57,027 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:57,027 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6f0ea258
2018-07-14 20:44:57,028 [pool-261-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:44:57,028 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1273406386_0093_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:44:57,030 [localfetcher#75] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#75 about to shuffle
output of map attempt_local1273406386_0093_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 20:44:57,030 [localfetcher#75] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1273406386_0093_m_000000_0
2018-07-14 20:44:57,030 [localfetcher#75] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
```

```
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 20:44:57,031 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:44:57,031 [pool-261-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:44:57,031 [pool-261-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:44:57,032 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:44:57,032 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:44:57,035 [pool-261-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 20:44:57,035 [pool-261-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 20:44:57,035 [pool-261-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:44:57,035 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:44:57,035 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 20:44:57,035 [pool-261-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:44:57,041 [pool-261-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:57,041 [pool-261-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,044 [pool-261-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[46,9],ex_m[39,7]
C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:44:57,045 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1273406386_0093_r_000000_0 is done. And is in the process of committing
2018-07-14 20:44:57,047 [pool-261-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:44:57,047 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1273406386_0093_r_000000_0 is allowed to commit now
2018-07-14 20:44:57,048 [pool-261-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1273406386_0093_r_000000_0' to
file:/tmp/temp-1902294536/tmp540182525/_temporary/0/task_local1273406386_0093_r_000000
2018-07-14 20:44:57,052 [pool-261-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:44:57,052 [pool-261-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1273406386_0093_r_000000_0' done.
2018-07-14 20:44:57,052 [pool-261-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1273406386_0093_r_000000_0
2018-07-14 20:44:57,052 [Thread-2597] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:44:57,240 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1273406386_0093
2018-07-14 20:44:57,240 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 20:44:57,240 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[46,9],ex_m[39,7] C: ex_max[46,9],ex_m[39,7] R: ex_max[46,9]
2018-07-14 20:44:57,243 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:44:57,255 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,261 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,263 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,268 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:44:57,268 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:44:57,269 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:44:57,269 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:44:57,270 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 20:44:57,270 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:44:57,272 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 20:44:57,285 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:44:57,287 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,290 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:57,290 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:57,297 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:44:57,298 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:44:57,298 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:57,298 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:57,299 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:44:57,307 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:57,307 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:57,307 [JobControl] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:44:57,325 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:44:57,340 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1291248604_0094
2018-07-14 20:44:57,398 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:44:57,399 [Thread-2626] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:44:57,403 [Thread-2626] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:44:57,403 [Thread-2626] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:44:57,403 [Thread-2626] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:57,403 [Thread-2626] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:44:57,403 [Thread-2626] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:57,403 [Thread-2626] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:44:57,406 [Thread-2626] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:44:57,406 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1291248604_0094_m_000000_0
2018-07-14 20:44:57,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:57,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:44:57,413 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:44:57,413 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:44:57,428 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:44:57,428 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:44:57,428 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:44:57,428 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:44:57,428 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:44:57,429 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
```

```
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:44:57,430 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:57,430 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,432 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[47,6],ans[48,6],details[8,10],details[-1,-1],ans[48,6] C:  R:
2018-07-14 20:44:57,433 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:44:57,433 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:44:57,433 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:44:57,433 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:44:57,433 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:44:57,434 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:44:57,435 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1291248604_0094_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:44:57,440 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:44:57,440 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1291248604_0094_m_000000_0' done.
2018-07-14 20:44:57,440 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1291248604_0094_m_000000_0
2018-07-14 20:44:57,440 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1291248604_0094_m_000001_0
2018-07-14 20:44:57,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:57,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:44:57,444 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp609880736/part-r-00000:0+62
2018-07-14 20:44:57,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:44:57,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:44:57,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:44:57,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:44:57,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:44:57,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
```

```
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:44:57,463 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:57,463 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,465 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[47,6],ans[48,6],details[8,10],details[-1,-1],ans[48,6] C:  R:
2018-07-14 20:44:57,466 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:57,474 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:57,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:44:57,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:44:57,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:44:57,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 12; bufvoid = 104857600
2018-07-14 20:44:57,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 20:44:57,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:44:57,476 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1291248604_0094_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:44:57,477 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:44:57,477 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1291248604_0094_m_000001_0' done.
2018-07-14 20:44:57,477 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1291248604_0094_m_000001_0
2018-07-14 20:44:57,477 [Thread-2626] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:44:57,477 [Thread-2626] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:44:57,478 [pool-264-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1291248604_0094_r_000000_0
2018-07-14 20:44:57,487 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:57,487 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@524b3453
2018-07-14 20:44:57,487 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:44:57,488 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1291248604_0094_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:44:57,491 [localfetcher#76] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#76 about to shuffle
output of map attempt_local1291248604_0094_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:44:57,491 [localfetcher#76] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1291248604_0094_m_000000_0
2018-07-14 20:44:57,491 [localfetcher#76] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 20:44:57,493 [localfetcher#76] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#76 about to shuffle
output of map attempt_local1291248604_0094_m_000001_0 decomp: 16 len: 20 to MEMORY
2018-07-14 20:44:57,493 [localfetcher#76] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 16 bytes from map-output
for attempt_local1291248604_0094_m_000001_0
2018-07-14 20:44:57,493 [localfetcher#76] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 16, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->368
2018-07-14 20:44:57,494 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:44:57,494 [pool-264-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:44:57,494 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:44:57,495 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:44:57,495 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 352 bytes
2018-07-14 20:44:57,496 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 368 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:44:57,496 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 370 bytes from
disk
2018-07-14 20:44:57,496 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:44:57,496 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:44:57,496 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 358 bytes
2018-07-14 20:44:57,496 [pool-264-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:44:57,497 [pool-264-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:57,497 [pool-264-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,499 [pool-264-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[47,6],ans[48,6],details[8,10],details[-1,-1],ans[48,6] C:  R:
2018-07-14 20:44:57,514 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1291248604_0094_r_000000_0 is done. And is in the process of committing
2018-07-14 20:44:57,516 [pool-264-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:44:57,516 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1291248604_0094_r_000000_0 is allowed to commit now
2018-07-14 20:44:57,516 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1291248604_0094_r_000000_0' to
file:/tmp/temp-1902294536/tmp484711713/_temporary/0/task_local1291248604_0094_r_000000
2018-07-14 20:44:57,519 [pool-264-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
```

```
'attempt_local1291248604_0094_r_000000_0' to
file:/tmp/temp-1902294536/tmp2071819735/_temporary/0/task_local1291248604_0094_r_00000
0
2018-07-14 20:44:57,519 [pool-264-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:44:57,519 [pool-264-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1291248604_0094_r_000000_0' done.
2018-07-14 20:44:57,519 [pool-264-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1291248604_0094_r_000000_0
2018-07-14 20:44:57,519 [Thread-2626] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:44:57,703 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1291248604_0094
2018-07-14 20:44:57,703 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,details
2018-07-14 20:44:57,703 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[47,6],ans[48,6],details[8,10],details[-1,-1],ans[48,6] C:  R:
2018-07-14 20:44:57,709 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 75%
complete
2018-07-14 20:44:57,717 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,720 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,722 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,728 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:44:57,728 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:44:57,730 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:44:57,741 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:44:57,743 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:57,744 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:57,745 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:57,753 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:44:57,761 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:57,761 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:57,761 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:44:57,784 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
```

```
2018-07-14 20:44:57,798 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1348107978_0095
2018-07-14 20:44:57,861 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:44:57,862 [Thread-2662] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:44:57,865 [Thread-2662] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:44:57,865 [Thread-2662] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:44:57,866 [Thread-2662] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:57,866 [Thread-2662] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:44:57,866 [Thread-2662] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:44:57,868 [Thread-2662] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:44:57,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1348107978_0095_m_000000_0
2018-07-14 20:44:57,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:44:57,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 27
Input split[0]:
   Length = 27
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:44:57,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp2071819735/part-r-00000:0+27
2018-07-14 20:44:57,881 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:44:57,881 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,882 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: ans1[50,7] C:  R:
2018-07-14 20:44:57,882 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:44:57,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:57,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:44:57,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:44:57,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1348107978_0095_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:44:57,892 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:44:57,892 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1348107978_0095_m_000000_0 is allowed
to commit now
2018-07-14 20:44:57,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1348107978_0095_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1945414275/_temporary/0/task_local1348107978_0095_m_0000
00
2018-07-14 20:44:57,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:44:57,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1348107978_0095_m_000000_0' done.
2018-07-14 20:44:57,893 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1348107978_0095_m_000000_0
2018-07-14 20:44:57,893 [Thread-2662] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:44:58,063 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1348107978_0095
2018-07-14 20:44:58,064 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans1
2018-07-14 20:44:58,064 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans1[50,7] C:  R:
2018-07-14 20:44:58,077 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,079 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,081 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,084 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:44:58,087 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:44:56     2018-07-14 20:44:58
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1248339603_0092        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1273406386_0093        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local1291248604_0094        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans,details     HASH_JOIN,MULTI_QUERY
job_local1348107978_0095        1       0       n/a     n/a     n/a     n/a     0
0       0       0       ans1    MAP_ONLY
file:/tmp/temp-1902294536/tmp-1945414275,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"
```

```
Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1902294536/tmp-1945414275"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1248339603_0092        ->
job_local1273406386_0093,job_local1291248604_0094,
job_local1273406386_0093        ->      job_local1291248604_0094,
job_local1291248604_0094        ->      job_local1348107978_0095,
job_local1348107978_0095


2018-07-14 20:44:58,089 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,090 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,090 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,092 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,093 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,093 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,094 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,095 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,096 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,097 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,097 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,098 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:44:58,099 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:44:58,099 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:44:58,100 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:44:58,100 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:44:58,108 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:44:58,108 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,500)
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> dump ex_g;
```

```
2018-07-14 20:47:53,396 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:47:53,414 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:47:53,414 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:47:53,414 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:47:53,414 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:47:53,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:47:53,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:47:53,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:47:53,423 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:47:53,423 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:47:53,423 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,424 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:47:53,424 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:47:53,424 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:47:53,424 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:47:53,424 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:47:53,424 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:47:53,425 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:47:53,425 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:47:53,425 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:47:53,425 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581473425-0
2018-07-14 20:47:53,431 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:47:53,431 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:47:53,438 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:47:53,440 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:47:53,440 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:47:53,440 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:47:53,440 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:47:53,459 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:47:53,473 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1190480721_0096
2018-07-14 20:47:53,539 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:47:53,540 [Thread-2687] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:47:53,545 [Thread-2687] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:47:53,545 [Thread-2687] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:47:53,545 [Thread-2687] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:47:53,545 [Thread-2687] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:47:53,545 [Thread-2687] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:47:53,545 [Thread-2687] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:47:53,547 [Thread-2687] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:47:53,547 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1190480721_0096_m_000000_0
2018-07-14 20:47:53,552 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:47:53,552 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:47:53,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:47:53,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:47:53,569 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:47:53,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:47:53,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:47:53,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:47:53,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:47:53,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:47:53,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:47:53,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:47:53,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:47:53,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:47:53,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:47:53,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:47:53,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 20:47:53,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:47:53,573 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:47:53,573 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1190480721_0096_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:47:53,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:47:53,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1190480721_0096_m_000000_0' done.
2018-07-14 20:47:53,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1190480721_0096_m_000000_0
2018-07-14 20:47:53,574 [Thread-2687] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:47:53,574 [Thread-2687] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:47:53,575 [pool-269-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1190480721_0096_r_000000_0
2018-07-14 20:47:53,582 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:47:53,582 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@107fc669
2018-07-14 20:47:53,583 [pool-269-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:47:53,584 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1190480721_0096_r_000000_0 Thread started: EventFetcher for fetching Map
```

```
Completion Events
2018-07-14 20:47:53,585 [localfetcher#77] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#77 about to shuffle
output of map attempt_local1190480721_0096_m_000000_0 decomp: 125 len: 129 to MEMORY
2018-07-14 20:47:53,585 [localfetcher#77] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local1190480721_0096_m_000000_0
2018-07-14 20:47:53,585 [localfetcher#77] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 20:47:53,586 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:47:53,586 [pool-269-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:47:53,586 [pool-269-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:47:53,587 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:47:53,587 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 20:47:53,588 [pool-269-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:47:53,588 [pool-269-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 20:47:53,588 [pool-269-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:47:53,588 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:47:53,588 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 20:47:53,588 [pool-269-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:47:53,594 [pool-269-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:47:53,594 [pool-269-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:47:53,595 [pool-269-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:47:53,596 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1190480721_0096_r_000000_0 is done. And is in the process of committing
2018-07-14 20:47:53,597 [pool-269-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:47:53,597 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1190480721_0096_r_000000_0 is allowed to commit now
2018-07-14 20:47:53,597 [pool-269-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1190480721_0096_r_000000_0' to
file:/tmp/temp-1902294536/tmp1323421184/_temporary/0/task_local1190480721_0096_r_00000
0
2018-07-14 20:47:53,598 [pool-269-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:47:53,598 [pool-269-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1190480721_0096_r_000000_0' done.
2018-07-14 20:47:53,598 [pool-269-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
```

```
attempt_local1190480721_0096_r_000000_0
2018-07-14 20:47:53,598 [Thread-2687] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:47:53,741 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1190480721_0096
2018-07-14 20:47:53,741 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,expenses
2018-07-14 20:47:53,741 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:47:53,747 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,749 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,749 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,758 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:47:53,759 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:47:53     2018-07-14 20:47:53
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1190480721_0096        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,expenses   GROUP_BY
file:/tmp/temp-1902294536/tmp1323421184,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp1323421184"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1190480721_0096


2018-07-14 20:47:53,761 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,762 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,763 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:47:53,764 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
Success!
2018-07-14 20:47:53,765 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:47:53,765 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:47:53,766 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:47:53,780 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:47:53,780 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> ans3 = JOIN details BY EmpId, ex_g BY EmpId;
2018-07-14 20:48:23,936 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 51, column 38> Invalid field projection. Projected field [EmpId] does not exist in
schema: group:int,expenses:bag{:tuple(EmpId:int,Expense:int)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans3 = JOIN details BY EmpId, ex_g BY expenses.EmpId;
2018-07-14 20:48:43,929 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1107: Cannot
merge join keys, incompatible types. Outer: int; inner: bag
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans3 = JOIN details BY details.EmpId, ex_g BY expenses.EmpId;
2018-07-14 20:49:01,741 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1107: Cannot
merge join keys, incompatible types. Outer: int; inner: bag
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt>
grunt> ans3 = JOIN details BY EmpId, ex_g BY EmpId;
2018-07-14 20:49:22,658 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 51, column 38> Invalid field projection. Projected field [EmpId] does not exist in
schema: group:int,expenses:bag{:tuple(EmpId:int,Expense:int)}.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans3 = JOIN details BY EmpId, ex_g BY expenses.EmpId;
2018-07-14 20:49:38,467 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1107: Cannot
merge join keys, incompatible types. Outer: int; inner: bag
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt> describe details;
details: {EmpId: int,Name: chararray,Salary: int,Dept: int}
grunt> describe ex_g;;
ex_g: {group: int,expenses: {(EmpId: int,Expense: int)}}
grunt> dump ex_g;
2018-07-14 20:50:09,394 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:50:09,412 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:09,413 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:09,413 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:50:09,413 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
```

```
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:50:09,416 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:50:09,416 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:50:09,416 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:50:09,424 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:09,424 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:09,424 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,428 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:50:09,428 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:50:09,428 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:50:09,428 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:50:09,428 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:50:09,428 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:50:09,429 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:50:09,429 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:50:09,429 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:50:09,429 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581609429-0
2018-07-14 20:50:09,436 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:50:09,437 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,444 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:50:09,445 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:50:09,445 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:09,445 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:50:09,446 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
```

```
(combined) to process : 1
2018-07-14 20:50:09,466 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:50:09,479 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1898157325_0097
2018-07-14 20:50:09,540 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:50:09,540 [Thread-2716] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:50:09,544 [Thread-2716] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:50:09,544 [Thread-2716] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:50:09,544 [Thread-2716] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:09,544 [Thread-2716] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:50:09,544 [Thread-2716] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:50:09,544 [Thread-2716] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:50:09,546 [Thread-2716] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:50:09,546 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1898157325_0097_m_000000_0
2018-07-14 20:50:09,551 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:09,552 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:50:09,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:50:09,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:50:09,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:50:09,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:50:09,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:50:09,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:50:09,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:50:09,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
```

```
2018-07-14 20:50:09,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:09,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:50:09,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:50:09,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:50:09,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:50:09,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:50:09,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 20:50:09,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:50:09,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:50:09,573 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1898157325_0097_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:50:09,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:50:09,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1898157325_0097_m_000000_0' done.
2018-07-14 20:50:09,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1898157325_0097_m_000000_0
2018-07-14 20:50:09,574 [Thread-2716] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:50:09,575 [Thread-2716] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:50:09,575 [pool-272-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1898157325_0097_r_000000_0
2018-07-14 20:50:09,582 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:09,583 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7cd57dbb
2018-07-14 20:50:09,583 [pool-272-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:50:09,584 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1898157325_0097_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:50:09,586 [localfetcher#78] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#78 about to shuffle
output of map attempt_local1898157325_0097_m_000000_0 decomp: 125 len: 129 to MEMORY
2018-07-14 20:50:09,586 [localfetcher#78] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local1898157325_0097_m_000000_0
2018-07-14 20:50:09,586 [localfetcher#78] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 20:50:09,590 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 20:50:09,591 [pool-272-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:09,591 [pool-272-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:50:09,592 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:09,592 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 20:50:09,592 [pool-272-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:50:09,592 [pool-272-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 20:50:09,592 [pool-272-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:50:09,593 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:09,593 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 20:50:09,593 [pool-272-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:09,600 [pool-272-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:09,600 [pool-272-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:50:09,600 [pool-272-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
2018-07-14 20:50:09,600 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1898157325_0097_r_000000_0 is done. And is in the process of committing
2018-07-14 20:50:09,601 [pool-272-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:09,601 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1898157325_0097_r_000000_0 is allowed to commit now
2018-07-14 20:50:09,602 [pool-272-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1898157325_0097_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1262551667/_temporary/0/task_local1898157325_0097_r_0000
00
2018-07-14 20:50:09,602 [pool-272-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:50:09,602 [pool-272-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1898157325_0097_r_000000_0' done.
2018-07-14 20:50:09,602 [pool-272-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1898157325_0097_r_000000_0
2018-07-14 20:50:09,602 [Thread-2716] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:50:09,742 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1898157325_0097
2018-07-14 20:50:09,742 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,expenses
2018-07-14 20:50:09,742 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R:
```

```
2018-07-14 20:50:09,744 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,744 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,745 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,746 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:50:09,747 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:50:09     2018-07-14 20:50:09
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1898157325_0097        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,expenses   GROUP_BY
file:/tmp/temp-1902294536/tmp-1262551667,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1262551667"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1898157325_0097


2018-07-14 20:50:09,747 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,748 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,748 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:09,751 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:50:09,751 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:09,752 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:09,752 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:50:09,762 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:09,762 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
```

```
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt> dump ex_sum;
2018-07-14 20:50:18,202 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:50:18,216 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:18,216 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:18,216 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:50:18,217 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:50:18,218 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:50:18,219 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:50:18,219 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:50:18,219 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:50:18,230 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:18,230 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:18,230 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,231 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:50:18,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:50:18,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:50:18,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:50:18,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:50:18,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:50:18,232 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:50:18,233 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
```

```
[pig.schematuple] is false, will not generate code.
2018-07-14 20:50:18,233 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:50:18,233 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581618232-0
2018-07-14 20:50:18,242 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:50:18,244 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,251 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:50:18,252 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:50:18,252 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:18,252 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:50:18,252 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:50:18,274 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:50:18,287 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local93845674_0098
2018-07-14 20:50:18,360 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:50:18,360 [Thread-2745] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:50:18,363 [Thread-2745] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:50:18,363 [Thread-2745] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:50:18,363 [Thread-2745] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:18,363 [Thread-2745] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:50:18,363 [Thread-2745] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:50:18,364 [Thread-2745] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:50:18,365 [Thread-2745] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:50:18,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local93845674_0098_m_000000_0
2018-07-14 20:50:18,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:18,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
```

```
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:50:18,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:50:18,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:50:18,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:50:18,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:50:18,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:50:18,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:50:18,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:50:18,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:50:18,395 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:18,395 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:50:18,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:50:18,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:50:18,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:50:18,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:50:18,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:50:18,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:50:18,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:50:18,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local93845674_0098_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:50:18,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:50:18,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local93845674_0098_m_000000_0' done.
2018-07-14 20:50:18,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local93845674_0098_m_000000_0
2018-07-14 20:50:18,409 [Thread-2745] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:50:18,410 [Thread-2745] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:50:18,410 [pool-275-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
```

```
attempt_local93845674_0098_r_000000_0
2018-07-14 20:50:18,415 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:18,415 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5c49d5b0
2018-07-14 20:50:18,415 [pool-275-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:50:18,416 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local93845674_0098_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:50:18,418 [localfetcher#79] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#79 about to shuffle
output of map attempt_local93845674_0098_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:50:18,418 [localfetcher#79] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local93845674_0098_m_000000_0
2018-07-14 20:50:18,418 [localfetcher#79] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:50:18,419 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:50:18,419 [pool-275-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:18,419 [pool-275-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:50:18,420 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:18,421 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:50:18,422 [pool-275-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:50:18,422 [pool-275-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:50:18,422 [pool-275-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:50:18,422 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:18,422 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:50:18,422 [pool-275-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:18,430 [pool-275-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:18,430 [pool-275-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:50:18,433 [pool-275-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:50:18,433 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local93845674_0098_r_000000_0 is done. And is in the process of committing
2018-07-14 20:50:18,435 [pool-275-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
```

```
2018-07-14 20:50:18,435 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local93845674_0098_r_000000_0 is allowed to commit now
2018-07-14 20:50:18,436 [pool-275-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local93845674_0098_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1526178509/_temporary/0/task_local93845674_0098_r_000000
2018-07-14 20:50:18,436 [pool-275-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:50:18,436 [pool-275-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local93845674_0098_r_000000_0' done.
2018-07-14 20:50:18,436 [pool-275-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local93845674_0098_r_000000_0
2018-07-14 20:50:18,436 [Thread-2745] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:50:18,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local93845674_0098
2018-07-14 20:50:18,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:50:18,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:50:18,562 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,562 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,562 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,564 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:50:18,564 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:50:18     2018-07-14 20:50:18
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local93845674_0098  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-1526178509,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1526178509"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
```

```
Job DAG:
job_local93845674_0098


2018-07-14 20:50:18,564 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,565 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,565 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:18,566 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:50:18,566 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:18,566 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:18,566 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:50:18,573 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:18,573 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> ans3 = JOIN details BY EmpId, ex_sum BY EmpId;;
grunt> dump ans3;
2018-07-14 20:50:38,573 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:50:38,586 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:38,586 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:38,586 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:50:38,587 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:50:38,588 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:50:38,589 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:50:38,589 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:50:38,590 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:50:38,590 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
```

```
2018-07-14 20:50:38,600 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:38,600 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:38,601 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:38,602 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:50:38,602 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:50:38,603 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:50:38,603 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:50:38,603 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:50:38,603 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:50:38,605 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:50:38,605 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:50:38,605 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:50:38,605 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581638605-0
2018-07-14 20:50:38,615 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:50:38,616 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:38,623 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:50:38,625 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:50:38,625 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:38,625 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:50:38,625 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:50:38,645 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:50:38,660 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local902174342_0099
2018-07-14 20:50:38,734 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:50:38,734 [Thread-2774] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:50:38,740 [Thread-2774] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
```

```
Instead, use mapreduce.jobtracker.address
2018-07-14 20:50:38,740 [Thread-2774] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:50:38,740 [Thread-2774] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:38,740 [Thread-2774] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:50:38,740 [Thread-2774] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:50:38,740 [Thread-2774] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:50:38,743 [Thread-2774] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:50:38,743 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local902174342_0099_m_000000_0
2018-07-14 20:50:38,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:38,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:50:38,751 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:50:38,752 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:50:38,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:50:38,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:50:38,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:50:38,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:50:38,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:50:38,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:50:38,779 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:38,779 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:50:38,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
```

```
2018-07-14 20:50:38,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:50:38,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:50:38,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:50:38,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:50:38,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:50:38,790 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:50:38,792 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local902174342_0099_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:50:38,793 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:50:38,793 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local902174342_0099_m_000000_0' done.
2018-07-14 20:50:38,793 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local902174342_0099_m_000000_0
2018-07-14 20:50:38,793 [Thread-2774] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:50:38,793 [Thread-2774] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:50:38,794 [pool-278-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local902174342_0099_r_000000_0
2018-07-14 20:50:38,799 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:38,799 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@700f478b
2018-07-14 20:50:38,800 [pool-278-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:50:38,801 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local902174342_0099_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:50:38,803 [localfetcher#80] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#80 about to shuffle
output of map attempt_local902174342_0099_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:50:38,803 [localfetcher#80] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local902174342_0099_m_000000_0
2018-07-14 20:50:38,803 [localfetcher#80] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:50:38,804 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:50:38,804 [pool-278-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:38,804 [pool-278-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:50:38,805 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:38,805 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
```

```
2018-07-14 20:50:38,806 [pool-278-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:50:38,806 [pool-278-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:50:38,806 [pool-278-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:50:38,807 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:38,807 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:50:38,807 [pool-278-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:38,815 [pool-278-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:38,815 [pool-278-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:50:38,819 [pool-278-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:50:38,820 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local902174342_0099_r_000000_0 is done. And is in the process of committing
2018-07-14 20:50:38,822 [pool-278-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:50:38,822 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local902174342_0099_r_000000_0 is allowed to commit now
2018-07-14 20:50:38,823 [pool-278-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local902174342_0099_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1061633667/_temporary/0/task_local902174342_0099_r_00000
0
2018-07-14 20:50:38,823 [pool-278-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:50:38,823 [pool-278-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local902174342_0099_r_000000_0' done.
2018-07-14 20:50:38,823 [pool-278-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local902174342_0099_r_000000_0
2018-07-14 20:50:38,823 [Thread-2774] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:50:38,935 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local902174342_0099
2018-07-14 20:50:38,935 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:50:38,935 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:50:38,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:50:38,940 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:38,941 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:38,942 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:38,945 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:50:38,945 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:50:38,945 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:50:38,945 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:50:38,946 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 20:50:38,946 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:50:38,948 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:50:38,957 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:50:38,959 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:38,961 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:38,961 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:50:38,966 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:50:38,967 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:50:38,967 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:38,967 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:50:38,968 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:50:38,975 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:38,975 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:50:38,975 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:50:38,996 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:50:39,008 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1760802496_0100
2018-07-14 20:50:39,073 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:50:39,073 [Thread-2803] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
```

```
2018-07-14 20:50:39,078 [Thread-2803] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:50:39,078 [Thread-2803] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:50:39,078 [Thread-2803] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:50:39,078 [Thread-2803] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:50:39,078 [Thread-2803] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:50:39,078 [Thread-2803] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:50:39,080 [Thread-2803] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:50:39,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1760802496_0100_m_000000_0
2018-07-14 20:50:39,084 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:39,085 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:50:39,086 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:50:39,086 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:50:39,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:50:39,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:50:39,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:50:39,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:50:39,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:50:39,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:50:39,104 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:39,104 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:50:39,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[51,7],details[8,10],details[-1,-1],ans3[51,7] C:  R:
```

```
2018-07-14 20:50:39,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:50:39,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:50:39,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:50:39,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:50:39,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:50:39,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:50:39,108 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1760802496_0100_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:50:39,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:50:39,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1760802496_0100_m_000000_0' done.
2018-07-14 20:50:39,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1760802496_0100_m_000000_0
2018-07-14 20:50:39,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1760802496_0100_m_000001_0
2018-07-14 20:50:39,111 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:39,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:50:39,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1061633667/part-r-00000:0+62
2018-07-14 20:50:39,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:50:39,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:50:39,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:50:39,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:50:39,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:50:39,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:50:39,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:39,239 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:50:39,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[51,7],details[8,10],details[-1,-1],ans3[51,7] C:  R:
```

```
2018-07-14 20:50:39,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:50:39,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:50:39,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:50:39,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 20:50:39,241 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:50:39,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:50:39,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1760802496_0100_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:50:39,244 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:50:39,244 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1760802496_0100_m_000001_0' done.
2018-07-14 20:50:39,244 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1760802496_0100_m_000001_0
2018-07-14 20:50:39,244 [Thread-2803] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:50:39,246 [Thread-2803] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:50:39,246 [pool-281-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1760802496_0100_r_000000_0
2018-07-14 20:50:39,251 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:50:39,251 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6710972c
2018-07-14 20:50:39,252 [pool-281-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:50:39,252 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1760802496_0100_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:50:39,253 [localfetcher#81] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#81 about to shuffle
output of map attempt_local1760802496_0100_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 20:50:39,253 [localfetcher#81] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1760802496_0100_m_000001_0
2018-07-14 20:50:39,253 [localfetcher#81] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 20:50:39,254 [localfetcher#81] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#81 about to shuffle
output of map attempt_local1760802496_0100_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:50:39,254 [localfetcher#81] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1760802496_0100_m_000000_0
2018-07-14 20:50:39,254 [localfetcher#81] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 20:50:39,255 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 20:50:39,257 [pool-281-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:50:39,257 [pool-281-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:50:39,258 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:50:39,258 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 20:50:39,260 [pool-281-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:50:39,260 [pool-281-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 20:50:39,260 [pool-281-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:50:39,260 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:50:39,260 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 20:50:39,260 [pool-281-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:50:39,266 [pool-281-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:50:39,266 [pool-281-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:50:39,267 [pool-281-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[51,7],details[8,10],details[-1,-1],ans3[51,7] C:  R:
2018-07-14 20:50:39,268 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1760802496_0100_r_000000_0 is done. And is in the process of committing
2018-07-14 20:50:39,269 [pool-281-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:50:39,269 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1760802496_0100_r_000000_0 is allowed to commit now
2018-07-14 20:50:39,270 [pool-281-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1760802496_0100_r_000000_0' to
file:/tmp/temp-1902294536/tmp978306037/_temporary/0/task_local1760802496_0100_r_000000
2018-07-14 20:50:39,271 [pool-281-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:50:39,271 [pool-281-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1760802496_0100_r_000000_0' done.
2018-07-14 20:50:39,271 [pool-281-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1760802496_0100_r_000000_0
2018-07-14 20:50:39,271 [Thread-2803] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:50:39,428 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1760802496_0100
2018-07-14 20:50:39,429 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,details
2018-07-14 20:50:39,429 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans3[51,7],details[8,10],details[-1,-1],ans3[51,7] C:  R:
2018-07-14 20:50:39,437 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,441 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,445 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,457 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:50:39,460 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:50:38     2018-07-14 20:50:39
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1760802496_0100        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans3,details    HASH_JOIN
file:/tmp/temp-1902294536/tmp978306037,
job_local902174342_0099 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 6 records in: "file:/tmp/temp-1902294536/tmp978306037"

Counters:
Total records written : 6
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local902174342_0099 ->      job_local1760802496_0100,
job_local1760802496_0100


2018-07-14 20:50:39,462 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,464 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,465 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,467 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,468 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,468 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:50:39,470 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:50:39,470 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 20:50:39,470 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:50:39,470 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:50:39,481 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:50:39,481 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(110,Priyanka,2000,5,110,400)
(114,Madhuri,2000,2,114,200)
grunt> ans3 = JOIN details BY EmpId, ex_sum BY EmpId;
grunt> ans4 = FILTER ans3 BY EmpId,Name;
2018-07-14 20:51:28,132 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
53, column 27>  mismatched input ',' expecting SEMI_COLON
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans4 = FILTER ans3 BY detaisl.EmpId,Name;
2018-07-14 20:51:40,005 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
53, column 35>  mismatched input ',' expecting SEMI_COLON
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe ans3
ans3: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::EmpId: int,ex_sum::sum: long}
grunt> ans4 = foreach ans3 GENERATE details.EmpId,Name;
grunt> dump ans4;
2018-07-14 20:52:38,064 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:52:38,081 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:52:38,081 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:52:38,081 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:52:38,081 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:52:38,083 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for expenses: $1
2018-07-14 20:52:38,085 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:52:38,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:52:38,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 4
2018-07-14 20:52:38,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 20:52:38,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 20:52:38,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
```

```
2018-07-14 20:52:38,093 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:52:38,093 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:52:38,095 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,096 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:52:38,097 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:52:38,098 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 20:52:38,099 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:52:38,099 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:52:38,099 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581758098-0
2018-07-14 20:52:38,105 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:52:38,105 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:52:38,105 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:52:38,105 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:52:38,105 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:52:38,105 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:52:38,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:52:38,106 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:52:38,106 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:52:38,106 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581758106-0
2018-07-14 20:52:38,114 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 2
map-reduce job(s) waiting for submission.
2018-07-14 20:52:38,115 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,122 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:52:38,125 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:52:38,126 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
```

```
2018-07-14 20:52:38,126 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:52:38,126 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:52:38,148 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:52:38,161 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local91752469_0101
2018-07-14 20:52:38,224 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:52:38,225 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,226 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:52:38,226 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:52:38,226 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:52:38,226 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:52:38,227 [Thread-2833] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:52:38,238 [Thread-2833] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:52:38,245 [Thread-2833] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:52:38,245 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local91752469_0101_m_000000_0
2018-07-14 20:52:38,254 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:52:38,255 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:52:38,255 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:52:38,255 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:52:38,255 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:52:38,263 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:52:38,264 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:52:38,271 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:52:38,271 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:52:38,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:52:38,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:52:38,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],details[-1,-1]
C:  R:
2018-07-14 20:52:38,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:52:38,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local91752469_0101_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:52:38,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:52:38,308 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local91752469_0101_m_000000_0 is allowed to
commit now
2018-07-14 20:52:38,309 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local91752469_0101_m_000000_0' to
file:/tmp/temp-1902294536/tmp-1804414535/_temporary/0/task_local91752469_0101_m_000000
2018-07-14 20:52:38,310 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local91752469_0101_m_000000_0' to
file:/tmp/temp-1902294536/tmp677625872/_temporary/0/task_local91752469_0101_m_000000
2018-07-14 20:52:38,311 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:52:38,311 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local91752469_0101_m_000000_0' done.
2018-07-14 20:52:38,311 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local91752469_0101_m_000000_0
2018-07-14 20:52:38,311 [Thread-2833] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:52:38,323 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:52:38,361 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1307687529_0102
2018-07-14 20:52:38,594 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:52:38,595 [Thread-2858] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:52:38,601 [Thread-2858] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:52:38,601 [Thread-2858] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:52:38,601 [Thread-2858] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:52:38,601 [Thread-2858] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:52:38,601 [Thread-2858] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:52:38,601 [Thread-2858] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:52:38,604 [Thread-2858] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:52:38,604 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1307687529_0102_m_000000_0
2018-07-14 20:52:38,611 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:52:38,612 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:52:38,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:52:38,615 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:52:38,616 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local91752469_0101
2018-07-14 20:52:38,616 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 20:52:38,616 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1] C:  R:
2018-07-14 20:52:38,617 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1307687529_0102
2018-07-14 20:52:38,617 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:52:38,617 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:52:38,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:52:38,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:52:38,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:52:38,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:52:38,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:52:38,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 16%
complete
2018-07-14 20:52:38,632 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:52:38,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
Running jobs are [job_local91752469_0101,job_local1307687529_0102]
2018-07-14 20:52:38,634 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:52:38,634 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:52:38,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:52:38,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:52:38,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:52:38,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:52:38,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 81; bufvoid = 104857600
2018-07-14 20:52:38,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:52:38,637 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:52:38,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1307687529_0102_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:52:38,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:52:38,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1307687529_0102_m_000000_0' done.
2018-07-14 20:52:38,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1307687529_0102_m_000000_0
2018-07-14 20:52:38,638 [Thread-2858] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:52:38,639 [Thread-2858] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:52:38,639 [pool-286-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1307687529_0102_r_000000_0
2018-07-14 20:52:38,643 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:52:38,643 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@681c030b
2018-07-14 20:52:38,644 [pool-286-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:52:38,644 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1307687529_0102_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:52:38,646 [localfetcher#82] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#82 about to shuffle
output of map attempt_local1307687529_0102_m_000000_0 decomp: 101 len: 105 to MEMORY
2018-07-14 20:52:38,646 [localfetcher#82] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 101 bytes from
map-output for attempt_local1307687529_0102_m_000000_0
2018-07-14 20:52:38,646 [localfetcher#82] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 101, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->101
2018-07-14 20:52:38,648 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 20:52:38,648 [pool-286-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:52:38,648 [pool-286-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:52:38,650 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:52:38,650 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:52:38,650 [pool-286-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 101 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:52:38,651 [pool-286-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 105 bytes from
disk
2018-07-14 20:52:38,651 [pool-286-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:52:38,651 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:52:38,651 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:52:38,651 [pool-286-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:52:38,656 [pool-286-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:52:38,656 [pool-286-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:52:38,657 [pool-286-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:52:38,657 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1307687529_0102_r_000000_0 is done. And is in the process of committing
2018-07-14 20:52:38,658 [pool-286-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:52:38,658 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1307687529_0102_r_000000_0 is allowed to commit now
2018-07-14 20:52:38,659 [pool-286-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1307687529_0102_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1732771438/_temporary/0/task_local1307687529_0102_r_0000
00
2018-07-14 20:52:38,659 [pool-286-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:52:38,659 [pool-286-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1307687529_0102_r_000000_0' done.
2018-07-14 20:52:38,659 [pool-286-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1307687529_0102_r_000000_0
2018-07-14 20:52:38,659 [Thread-2858] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:52:38,798 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:52:38,803 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,806 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,808 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:52:38,816 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,819 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,821 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,827 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:52:38,828 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:52:38,829 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:52:38,830 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:52:38,835 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=322
2018-07-14 20:52:38,835 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:52:38,844 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:52:38,867 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:52:38,870 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:38,872 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:52:38,872 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:52:38,880 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:52:38,889 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:52:38,889 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:52:38,890 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:52:38,897 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:52:38,897 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:52:38,898 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:52:38,920 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:52:38,934 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local284759631_0103
2018-07-14 20:52:38,998 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
```

```
track the job: http://localhost:8080/
2018-07-14 20:52:38,998 [Thread-2889] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:52:39,002 [Thread-2889] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:52:39,002 [Thread-2889] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:52:39,002 [Thread-2889] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:52:39,002 [Thread-2889] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:52:39,002 [Thread-2889] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:52:39,002 [Thread-2889] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:52:39,003 [Thread-2889] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:52:39,003 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local284759631_0103_m_000000_0
2018-07-14 20:52:39,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:52:39,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 280
Input split[0]:
   Length = 280
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:52:39,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp677625872/part-m-00000:0+280
2018-07-14 20:52:39,024 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:52:39,024 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:52:39,024 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:52:39,024 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:52:39,024 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:52:39,025 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:52:39,026 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:52:39,026 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:52:39,027 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
```

```
details[-1,-1],ans3[52,7],ans3[52,7] C:  R: ans4[53,7]
2018-07-14 20:52:39,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:52:39,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:52:39,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:52:39,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 20:52:39,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:52:39,030 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:52:39,031 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local284759631_0103_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:52:39,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:52:39,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local284759631_0103_m_000000_0' done.
2018-07-14 20:52:39,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local284759631_0103_m_000000_0
2018-07-14 20:52:39,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local284759631_0103_m_000001_0
2018-07-14 20:52:39,034 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:52:39,035 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 42
Input split[0]:
   Length = 42
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:52:39,036 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1732771438/part-r-00000:0+42
2018-07-14 20:52:39,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:52:39,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:52:39,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:52:39,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:52:39,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:52:39,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:52:39,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:52:39,231 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:52:39,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
```

```
details[-1,-1],ans3[52,7],ans3[52,7] C:  R: ans4[53,7]
2018-07-14 20:52:39,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:52:39,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:52:39,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:52:39,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 63; bufvoid = 104857600
2018-07-14 20:52:39,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:52:39,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:52:39,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local284759631_0103_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:52:39,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:52:39,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local284759631_0103_m_000001_0' done.
2018-07-14 20:52:39,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local284759631_0103_m_000001_0
2018-07-14 20:52:39,235 [Thread-2889] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:52:39,235 [Thread-2889] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:52:39,236 [pool-289-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local284759631_0103_r_000000_0
2018-07-14 20:52:39,241 [pool-289-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:52:39,241 [pool-289-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7b12db51
2018-07-14 20:52:39,242 [pool-289-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:52:39,242 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local284759631_0103_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:52:39,243 [localfetcher#83] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#83 about to shuffle
output of map attempt_local284759631_0103_m_000001_0 decomp: 79 len: 83 to MEMORY
2018-07-14 20:52:39,243 [localfetcher#83] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 79 bytes from map-output
for attempt_local284759631_0103_m_000001_0
2018-07-14 20:52:39,243 [localfetcher#83] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 79, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->79
2018-07-14 20:52:39,244 [localfetcher#83] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#83 about to shuffle
output of map attempt_local284759631_0103_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 20:52:39,244 [localfetcher#83] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local284759631_0103_m_000000_0
2018-07-14 20:52:39,244 [localfetcher#83] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 2, commitMemory -> 79, usedMemory
->365
2018-07-14 20:52:39,244 [EventFetcher for fetching Map Completion Events] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:52:39,245 [pool-289-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:52:39,245 [pool-289-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:52:39,246 [pool-289-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:52:39,246 [pool-289-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 349 bytes
2018-07-14 20:52:39,247 [pool-289-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 365 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:52:39,247 [pool-289-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 367 bytes from
disk
2018-07-14 20:52:39,247 [pool-289-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:52:39,247 [pool-289-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:52:39,247 [pool-289-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 355 bytes
2018-07-14 20:52:39,247 [pool-289-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:52:39,256 [pool-289-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:52:39,256 [pool-289-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:52:39,257 [pool-289-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[-1,-1],ans3[52,7],ans3[52,7] C:  R: ans4[53,7]
2018-07-14 20:52:39,258 [pool-289-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:52:39,265 [pool-289-thread-1] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:52:39,265 [pool-289-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:52:39,266 [Thread-2889] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:52:39,267 [Thread-2889] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local284759631_0103
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,Amitabh,20000,1), 2nd
:(102,Shahrukh,10000,2) (common cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should
be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:529)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,Amitabh,20000,1), 2nd :(102,Shahrukh,10000,2) (common
cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextInteger(POUserFunc.java:375)
```

```
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:347)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.processPlan(POForEach.java:404)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:321)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.runPipeline(PigGenericMapReduce.java:465)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.processOnePackageOutput(PigGenericMapReduce.java:433)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:405)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:262)
        at org.apache.hadoop.mapreduce.Reducer.run(Reducer.java:171)
        at org.apache.hadoop.mapred.ReduceTask.runNewReducer(ReduceTask.java:627)
        at org.apache.hadoop.mapred.ReduceTask.run(ReduceTask.java:389)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$ReduceTaskRunnable.run(LocalJobRunner.java
:319)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
2018-07-14 20:52:39,369 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local284759631_0103
2018-07-14 20:52:39,369 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,ans4,details
2018-07-14 20:52:39,369 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[-1,-1],ans3[52,7],ans3[52,7] C:  R: ans4[53,7]
2018-07-14 20:52:39,372 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 20:52:39,372 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local284759631_0103 has failed! Stop running all dependent jobs
2018-07-14 20:52:39,372 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:52:39,375 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,376 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,376 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 20:52:39,377 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:52:38     2018-07-14 20:52:39
```

```
HASH_JOIN,GROUP_BY

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1307687529_0102        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY
job_local91752469_0101  1       0       n/a     n/a     n/a     n/a     0       0
0       0       details MULTI_QUERY,MAP_ONLY

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local284759631_0103 ans3,ans4,details       HASH_JOIN       Message: Job failed!
file:/tmp/temp-1902294536/tmp-1229096654,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp-1229096654"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local91752469_0101  ->      job_local284759631_0103,
job_local1307687529_0102        ->      job_local284759631_0103,
job_local284759631_0103


2018-07-14 20:52:39,379 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,380 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,382 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,384 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,385 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,385 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:52:39,387 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 20:52:39,388 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias ans4
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> 2018-07-14 20:52:45,243 [communication thread] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce

grunt>
grunt> describe ans3;;
2018-07-14 20:53:09,874 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 20:53:09,874 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
ans3: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::EmpId: int,ex_sum::sum: long}
grunt>
grunt> ans4 = foreach ans3 GENERATE details.EmpId,details.Name;
grunt> dump ans4;
2018-07-14 20:53:22,069 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:53:22,088 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,088 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:53:22,088 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:53:22,089 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:53:22,090 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for expenses: $1
2018-07-14 20:53:22,092 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:53:22,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:53:22,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 4
2018-07-14 20:53:22,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 20:53:22,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 20:53:22,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 20:53:22,101 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,102 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:53:22,102 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,103 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:53:22,103 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:53:22,104 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 20:53:22,104 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:53:22,104 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:53:22,104 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581802104-0
```

```
2018-07-14 20:53:22,111 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:53:22,112 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:53:22,112 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:53:22,112 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:53:22,112 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:53:22,112 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:53:22,113 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:53:22,113 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:53:22,113 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:53:22,113 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581802113-0
2018-07-14 20:53:22,120 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 2
map-reduce job(s) waiting for submission.
2018-07-14 20:53:22,122 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,131 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:53:22,133 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:53:22,133 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:22,133 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:22,133 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:22,153 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:53:22,167 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local294871590_0104
2018-07-14 20:53:22,230 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:53:22,230 [Thread-2917] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:53:22,230 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,232 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:53:22,232 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
```

```
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:53:22,232 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,232 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:22,236 [Thread-2917] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:53:22,239 [Thread-2917] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:53:22,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local294871590_0104_m_000000_0
2018-07-14 20:53:22,243 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:53:22,243 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:53:22,244 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:22,244 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:22,244 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:22,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:22,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:22,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:53:22,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:53:22,249 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:22,251 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:53:22,253 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[8,10],details[-1,-1]
C:  R:
2018-07-14 20:53:22,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:22,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local294871590_0104_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:53:22,278 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:22,278 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task attempt_local294871590_0104_m_000000_0 is allowed to
commit now
2018-07-14 20:53:22,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local294871590_0104_m_000000_0' to
file:/tmp/temp-1902294536/tmp-430276577/_temporary/0/task_local294871590_0104_m_000000
2018-07-14 20:53:22,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local294871590_0104_m_000000_0' to
file:/tmp/temp-1902294536/tmp-226648363/_temporary/0/task_local294871590_0104_m_000000
2018-07-14 20:53:22,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:22,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local294871590_0104_m_000000_0' done.
2018-07-14 20:53:22,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local294871590_0104_m_000000_0
2018-07-14 20:53:22,281 [Thread-2917] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:53:22,298 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:53:22,333 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,333 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:22,336 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local423246280_0105
2018-07-14 20:53:22,396 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:53:22,396 [Thread-2942] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:53:22,400 [Thread-2942] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:53:22,400 [Thread-2942] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:53:22,400 [Thread-2942] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,400 [Thread-2942] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:53:22,400 [Thread-2942] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:22,400 [Thread-2942] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:53:22,402 [Thread-2942] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:53:22,402 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local423246280_0105_m_000000_0
2018-07-14 20:53:22,406 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:22,407 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
```

```
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:22,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:53:22,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:53:22,507 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:53:22,507 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:53:22,507 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:53:22,507 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:53:22,507 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:53:22,508 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:53:22,508 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:22,509 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:53:22,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:53:22,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:22,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:53:22,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:53:22,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 81; bufvoid = 104857600
2018-07-14 20:53:22,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:53:22,511 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:53:22,511 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local423246280_0105_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:53:22,512 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:22,512 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local423246280_0105_m_000000_0' done.
2018-07-14 20:53:22,512 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local423246280_0105_m_000000_0
2018-07-14 20:53:22,512 [Thread-2942] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:53:22,512 [Thread-2942] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:53:22,513 [pool-294-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local423246280_0105_r_000000_0
```

```
2018-07-14 20:53:22,518 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:22,518 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7b0b6e64
2018-07-14 20:53:22,518 [pool-294-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:53:22,519 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local423246280_0105_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:53:22,523 [localfetcher#84] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#84 about to shuffle
output of map attempt_local423246280_0105_m_000000_0 decomp: 101 len: 105 to MEMORY
2018-07-14 20:53:22,524 [localfetcher#84] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 101 bytes from
map-output for attempt_local423246280_0105_m_000000_0
2018-07-14 20:53:22,524 [localfetcher#84] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 101, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->101
2018-07-14 20:53:22,524 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:53:22,525 [pool-294-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:53:22,525 [pool-294-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:53:22,525 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:53:22,525 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:53:22,526 [pool-294-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 101 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:53:22,526 [pool-294-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 105 bytes from
disk
2018-07-14 20:53:22,526 [pool-294-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:53:22,526 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:53:22,526 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:53:22,526 [pool-294-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:53:22,532 [pool-294-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:22,532 [pool-294-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:53:22,533 [pool-294-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:53:22,533 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local423246280_0105_r_000000_0 is done. And is in the process of committing
2018-07-14 20:53:22,534 [pool-294-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:53:22,534 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local423246280_0105_r_000000_0 is allowed to commit now
```

```
2018-07-14 20:53:22,535 [pool-294-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local423246280_0105_r_000000_0' to
file:/tmp/temp-1902294536/tmp1622207236/_temporary/0/task_local423246280_0105_r_000000
2018-07-14 20:53:22,535 [pool-294-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:53:22,536 [pool-294-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local423246280_0105_r_000000_0' done.
2018-07-14 20:53:22,536 [pool-294-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local423246280_0105_r_000000_0
2018-07-14 20:53:22,536 [Thread-2942] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:53:22,621 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local294871590_0104
2018-07-14 20:53:22,621 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 20:53:22,621 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1] C:  R:
2018-07-14 20:53:22,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local423246280_0105
2018-07-14 20:53:22,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:53:22,625 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:53:22,628 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 20:53:22,633 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,634 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,636 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,640 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,641 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,643 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,645 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:53:22,645 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:53:22,645 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:53:22,645 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:53:22,647 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=322
2018-07-14 20:53:22,647 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:53:22,648 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:53:22,659 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:53:22,660 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:22,661 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,661 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:22,668 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:53:22,680 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:22,680 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:22,680 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:22,686 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:22,686 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:22,686 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:22,707 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:53:22,720 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local965001028_0106
2018-07-14 20:53:22,778 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:53:22,778 [Thread-2973] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:53:22,782 [Thread-2973] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:53:22,782 [Thread-2973] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:53:22,782 [Thread-2973] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:22,782 [Thread-2973] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:53:22,782 [Thread-2973] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:22,782 [Thread-2973] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:53:22,784 [Thread-2973] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for map tasks
2018-07-14 20:53:22,784 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local965001028_0106_m_000000_0
2018-07-14 20:53:22,789 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:22,789 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 280
Input split[0]:
   Length = 280
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:22,790 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-226648363/part-m-00000:0+280
2018-07-14 20:53:22,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:53:22,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:53:22,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:53:22,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:53:22,806 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:53:22,807 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:53:22,809 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:22,809 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:53:22,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[-1,-1],ans3[52,7] C:  R: ans4[54,7]
2018-07-14 20:53:22,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:22,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:53:22,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:53:22,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 126; bufvoid = 104857600
2018-07-14 20:53:22,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:53:22,812 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:53:22,813 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local965001028_0106_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:53:22,813 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:22,813 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local965001028_0106_m_000000_0' done.
2018-07-14 20:53:22,813 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
```

```
attempt_local965001028_0106_m_000000_0
2018-07-14 20:53:22,813 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local965001028_0106_m_000001_0
2018-07-14 20:53:22,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:22,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 42
Input split[0]:
   Length = 42
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:22,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1622207236/part-r-00000:0+42
2018-07-14 20:53:22,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:53:22,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:53:22,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:53:22,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:53:22,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:53:22,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:53:22,952 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:22,952 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:53:22,953 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[-1,-1],ans3[52,7] C:  R: ans4[54,7]
2018-07-14 20:53:22,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:22,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:53:22,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:53:22,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 63; bufvoid = 104857600
2018-07-14 20:53:22,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:53:22,954 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:53:22,956 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local965001028_0106_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:53:22,956 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:22,956 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local965001028_0106_m_000001_0' done.
2018-07-14 20:53:22,956 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
```

```
attempt_local965001028_0106_m_000001_0
2018-07-14 20:53:22,956 [Thread-2973] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:53:22,956 [Thread-2973] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:53:22,957 [pool-297-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local965001028_0106_r_000000_0
2018-07-14 20:53:22,962 [pool-297-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:22,962 [pool-297-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3ca11baf
2018-07-14 20:53:22,962 [pool-297-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:53:22,963 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local965001028_0106_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:53:22,967 [localfetcher#85] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#85 about to shuffle
output of map attempt_local965001028_0106_m_000001_0 decomp: 79 len: 83 to MEMORY
2018-07-14 20:53:22,967 [localfetcher#85] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 79 bytes from map-output
for attempt_local965001028_0106_m_000001_0
2018-07-14 20:53:22,967 [localfetcher#85] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 79, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->79
2018-07-14 20:53:22,968 [localfetcher#85] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#85 about to shuffle
output of map attempt_local965001028_0106_m_000000_0 decomp: 156 len: 160 to MEMORY
2018-07-14 20:53:22,968 [localfetcher#85] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 156 bytes from
map-output for attempt_local965001028_0106_m_000000_0
2018-07-14 20:53:22,968 [localfetcher#85] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 156, inMemoryMapOutputs.size() -> 2, commitMemory -> 79, usedMemory
->235
2018-07-14 20:53:22,968 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:53:22,969 [pool-297-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:53:22,969 [pool-297-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:53:22,970 [pool-297-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:53:22,970 [pool-297-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 219 bytes
2018-07-14 20:53:22,970 [pool-297-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 235 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:53:22,970 [pool-297-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 237 bytes from
disk
2018-07-14 20:53:22,970 [pool-297-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:53:22,970 [pool-297-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:53:22,971 [pool-297-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
```

```
to the last merge-pass, with 1 segments left of total size: 225 bytes
2018-07-14 20:53:22,971 [pool-297-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:53:22,976 [pool-297-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:22,976 [pool-297-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:53:22,979 [pool-297-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[-1,-1],ans3[52,7] C:  R: ans4[54,7]
2018-07-14 20:53:22,979 [pool-297-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:53:22,986 [pool-297-thread-1] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:22,986 [pool-297-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:22,987 [Thread-2973] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:53:22,990 [Thread-2973] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local965001028_0106
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,Amitabh,20000,1), 2nd
:(102,Shahrukh,10000,2) (common cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should
be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:529)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,Amitabh,20000,1), 2nd :(102,Shahrukh,10000,2) (common
cause: "JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextInteger(POUserFunc.java:375)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:347)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.processPlan(POForEach.java:404)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:321)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.runPipeline(PigGenericMapReduce.java:465)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.processOnePackageOutput(PigGenericMapReduce.java:433)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:405)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Reduc
e.reduce(PigGenericMapReduce.java:262)
        at org.apache.hadoop.mapreduce.Reducer.run(Reducer.java:171)
        at org.apache.hadoop.mapred.ReduceTask.runNewReducer(ReduceTask.java:627)
```

```
        at org.apache.hadoop.mapred.ReduceTask.run(ReduceTask.java:389)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$ReduceTaskRunnable.run(LocalJobRunner.java
:319)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
2018-07-14 20:53:23,137 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local965001028_0106
2018-07-14 20:53:23,137 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,ans4,details
2018-07-14 20:53:23,137 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans3[52,7],details[-1,-1],ans3[52,7] C:  R: ans4[54,7]
2018-07-14 20:53:23,139 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 20:53:23,139 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local965001028_0106 has failed! Stop running all dependent jobs
2018-07-14 20:53:23,139 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:53:23,142 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,142 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,142 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 20:53:23,143 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:53:22     2018-07-14 20:53:23
HASH_JOIN,GROUP_BY

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local294871590_0104 1       0       n/a     n/a     n/a     n/a     0       0
0       0       details MULTI_QUERY,MAP_ONLY
job_local423246280_0105 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local965001028_0106 ans3,ans4,details       HASH_JOIN       Message: Job failed!
file:/tmp/temp-1902294536/tmp781690606,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
```

```
Failed to produce result in "file:/tmp/temp-1902294536/tmp781690606"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local294871590_0104 ->      job_local965001028_0106,
job_local423246280_0105 ->      job_local965001028_0106,
job_local965001028_0106


2018-07-14 20:53:23,145 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,146 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,146 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,148 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,149 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,150 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:23,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 20:53:23,152 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias ans4
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt> describe ans3;;
2018-07-14 20:53:26,797 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:26,797 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
ans3: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::EmpId: int,ex_sum::sum: long}
grunt> ans4 = foreach ans3 GENERATE details.EmpId,details.Name;2018-07-14 20:53:28,962
[communication thread] INFO  org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce

grunt> ans4 = foreach ans3 GENERATE $0,$1;
grunt> dump ans4;
2018-07-14 20:53:47,155 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:53:47,174 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:47,174 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:53:47,175 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:53:47,175 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:53:47,175 [main] INFO
```

```
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 20:53:47,176 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for expenses: $1
2018-07-14 20:53:47,177 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:53:47,179 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:53:47,179 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:53:47,179 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:53:47,188 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:47,188 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:53:47,189 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:47,190 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:53:47,190 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:53:47,190 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:53:47,190 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:53:47,190 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:53:47,190 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:53:47,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:53:47,191 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:53:47,191 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:53:47,191 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581827191-0
2018-07-14 20:53:47,200 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:53:47,201 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:47,208 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:53:47,209 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:53:47,210 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
```

```
2018-07-14 20:53:47,210 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:47,210 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:47,229 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:53:47,244 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local392100616_0107
2018-07-14 20:53:47,308 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:53:47,309 [Thread-3001] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:53:47,312 [Thread-3001] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:53:47,312 [Thread-3001] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:53:47,312 [Thread-3001] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:47,312 [Thread-3001] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:53:47,312 [Thread-3001] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:47,313 [Thread-3001] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:53:47,315 [Thread-3001] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:53:47,315 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local392100616_0107_m_000000_0
2018-07-14 20:53:47,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:47,321 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:47,322 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:53:47,322 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:53:47,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:53:47,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:53:47,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:53:47,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
```

```
2018-07-14 20:53:47,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:53:47,361 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:53:47,361 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:47,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:53:47,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:53:47,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:47,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:53:47,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:53:47,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 81; bufvoid = 104857600
2018-07-14 20:53:47,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:53:47,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:53:47,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local392100616_0107_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:53:47,367 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:47,367 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local392100616_0107_m_000000_0' done.
2018-07-14 20:53:47,367 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local392100616_0107_m_000000_0
2018-07-14 20:53:47,367 [Thread-3001] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:53:47,367 [Thread-3001] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:53:47,368 [pool-300-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local392100616_0107_r_000000_0
2018-07-14 20:53:47,372 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:47,372 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2449280c
2018-07-14 20:53:47,373 [pool-300-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:53:47,377 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local392100616_0107_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:53:47,378 [localfetcher#86] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#86 about to shuffle
output of map attempt_local392100616_0107_m_000000_0 decomp: 101 len: 105 to MEMORY
2018-07-14 20:53:47,378 [localfetcher#86] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 101 bytes from
map-output for attempt_local392100616_0107_m_000000_0
```

```
2018-07-14 20:53:47,379 [localfetcher#86] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 101, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->101
2018-07-14 20:53:47,379 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:53:47,379 [pool-300-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:53:47,379 [pool-300-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:53:47,380 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:53:47,380 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:53:47,381 [pool-300-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 101 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:53:47,382 [pool-300-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 105 bytes from
disk
2018-07-14 20:53:47,382 [pool-300-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:53:47,382 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:53:47,382 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:53:47,382 [pool-300-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:53:47,389 [pool-300-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:47,389 [pool-300-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:53:47,391 [pool-300-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:53:47,393 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local392100616_0107_r_000000_0 is done. And is in the process of committing
2018-07-14 20:53:47,395 [pool-300-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:53:47,395 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local392100616_0107_r_000000_0 is allowed to commit now
2018-07-14 20:53:47,395 [pool-300-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local392100616_0107_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1871502772/_temporary/0/task_local392100616_0107_r_00000
0
2018-07-14 20:53:47,396 [pool-300-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:53:47,396 [pool-300-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local392100616_0107_r_000000_0' done.
2018-07-14 20:53:47,396 [pool-300-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local392100616_0107_r_000000_0
2018-07-14 20:53:47,396 [Thread-3001] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:53:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local392100616_0107
2018-07-14 20:53:47,509 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:53:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:53:47,510 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:53:47,512 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:47,513 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:47,513 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:47,516 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:53:47,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:53:47,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:53:47,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:53:47,518 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=317
2018-07-14 20:53:47,518 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:53:47,519 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:53:47,527 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:53:47,530 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:47,531 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:47,531 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:47,537 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:53:47,545 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:47,545 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:47,545 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:47,547 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:53:47,547 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:47,547 [JobControl] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:53:47,547 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:53:47,570 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:53:47,585 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1756321985_0108
2018-07-14 20:53:47,645 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:53:47,646 [Thread-3030] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:53:47,649 [Thread-3030] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:53:47,649 [Thread-3030] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:53:47,649 [Thread-3030] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:47,649 [Thread-3030] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:53:47,649 [Thread-3030] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:53:47,649 [Thread-3030] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:53:47,651 [Thread-3030] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:53:47,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1756321985_0108_m_000000_0
2018-07-14 20:53:47,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:47,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:47,659 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:53:47,659 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:53:47,674 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:53:47,674 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:53:47,674 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:53:47,674 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:53:47,674 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:53:47,675 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:53:47,677 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:47,677 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:53:47,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R: ans4[56,7]
2018-07-14 20:53:47,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:47,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:53:47,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:53:47,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 20:53:47,679 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:53:47,680 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:53:47,681 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1756321985_0108_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:53:47,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:47,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1756321985_0108_m_000000_0' done.
2018-07-14 20:53:47,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1756321985_0108_m_000000_0
2018-07-14 20:53:47,682 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1756321985_0108_m_000001_0
2018-07-14 20:53:47,683 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:47,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 42
Input split[0]:
   Length = 42
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:53:47,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1871502772/part-r-00000:0+42
2018-07-14 20:53:47,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:53:47,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:53:47,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:53:47,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:53:47,817 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:53:47,818 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:53:47,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:47,820 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R: ans4[56,7]
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 63; bufvoid = 104857600
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:53:47,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:53:47,825 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1756321985_0108_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:53:47,825 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:53:47,825 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1756321985_0108_m_000001_0' done.
2018-07-14 20:53:47,825 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1756321985_0108_m_000001_0
2018-07-14 20:53:47,825 [Thread-3030] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:53:47,826 [Thread-3030] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:53:47,826 [pool-303-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1756321985_0108_r_000000_0
2018-07-14 20:53:47,831 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:53:47,831 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1458f337
2018-07-14 20:53:47,831 [pool-303-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:53:47,832 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1756321985_0108_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:53:47,833 [localfetcher#87] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#87 about to shuffle
output of map attempt_local1756321985_0108_m_000001_0 decomp: 79 len: 83 to MEMORY
2018-07-14 20:53:47,833 [localfetcher#87] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 79 bytes from map-output
for attempt_local1756321985_0108_m_000001_0
2018-07-14 20:53:47,833 [localfetcher#87] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
```

```
map-output of size: 79, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->79
2018-07-14 20:53:47,834 [localfetcher#87] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#87 about to shuffle
output of map attempt_local1756321985_0108_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 20:53:47,834 [localfetcher#87] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local1756321985_0108_m_000000_0
2018-07-14 20:53:47,834 [localfetcher#87] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 2, commitMemory -> 79, usedMemory
->365
2018-07-14 20:53:47,834 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:53:47,835 [pool-303-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:53:47,835 [pool-303-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:53:47,835 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:53:47,835 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 349 bytes
2018-07-14 20:53:47,835 [pool-303-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 365 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:53:47,836 [pool-303-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 367 bytes from
disk
2018-07-14 20:53:47,836 [pool-303-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:53:47,836 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:53:47,836 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 355 bytes
2018-07-14 20:53:47,836 [pool-303-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:53:47,842 [pool-303-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:53:47,842 [pool-303-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:53:47,844 [pool-303-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R: ans4[56,7]
2018-07-14 20:53:47,845 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1756321985_0108_r_000000_0 is done. And is in the process of committing
2018-07-14 20:53:47,846 [pool-303-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:53:47,846 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1756321985_0108_r_000000_0 is allowed to commit now
2018-07-14 20:53:47,847 [pool-303-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1756321985_0108_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1020400330/_temporary/0/task_local1756321985_0108_r_0000
00
2018-07-14 20:53:47,847 [pool-303-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:53:47,847 [pool-303-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1756321985_0108_r_000000_0' done.
2018-07-14 20:53:47,847 [pool-303-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1756321985_0108_r_000000_0
2018-07-14 20:53:47,847 [Thread-3030] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:53:48,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1756321985_0108
2018-07-14 20:53:48,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,ans4,details
2018-07-14 20:53:48,007 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R:
ans4[56,7]
2018-07-14 20:53:48,011 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,013 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,014 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,018 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:53:48,019 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:53:47     2018-07-14 20:53:48
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1756321985_0108        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans3,ans4,details       HASH_JOIN
file:/tmp/temp-1902294536/tmp-1020400330,
job_local392100616_0107 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 6 records in: "file:/tmp/temp-1902294536/tmp-1020400330"

Counters:
Total records written : 6
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local392100616_0107 ->      job_local1756321985_0108,
job_local1756321985_0108


2018-07-14 20:53:48,020 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:53:48,022 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,023 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,026 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,027 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,028 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:53:48,030 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:53:48,030 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:53:48,031 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:53:48,031 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:53:48,045 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:53:48,045 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh)
(102,Shahrukh)
(104,Anubhav)
(105,Pawan)
(110,Priyanka)
(114,Madhuri)
grunt> describe ans3;;
ans3: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::EmpId: int,ex_sum::sum: long}
grunt> ans4 = foreach ans3 GENERATE $0,$1;
grunt> dump ans4;
2018-07-14 20:54:54,653 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:54:54,667 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:54:54,668 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:54:54,668 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:54:54,668 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:54:54,669 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 20:54:54,670 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for expenses: $1
2018-07-14 20:54:54,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:54:54,672 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 20:54:54,672 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
```

```
2018-07-14 20:54:54,672 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:54:54,679 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:54:54,680 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:54:54,680 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:54,682 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:54:54,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:54:54,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:54:54,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:54:54,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:54:54,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:54:54,685 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:54:54,686 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:54:54,686 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:54:54,686 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531581894686-0
2018-07-14 20:54:54,692 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:54:54,693 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:54,699 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:54:54,701 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:54:54,702 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:54:54,702 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:54:54,702 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:54:54,722 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:54:54,735 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local867678939_0109
2018-07-14 20:54:54,797 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:54:54,797 [Thread-3060] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter set in config null
2018-07-14 20:54:54,801 [Thread-3060] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:54:54,801 [Thread-3060] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:54:54,801 [Thread-3060] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:54:54,801 [Thread-3060] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:54:54,801 [Thread-3060] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:54:54,801 [Thread-3060] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:54:54,803 [Thread-3060] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:54:54,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local867678939_0109_m_000000_0
2018-07-14 20:54:54,807 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:54:54,808 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:54:54,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:54:54,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:54:54,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:54:54,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:54:54,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:54:54,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:54:54,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:54:54,848 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:54:54,849 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:54:54,850 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:54:54,853 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
```

```
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:54:54,858 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:54:54,858 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:54:54,858 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:54:54,858 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 81; bufvoid = 104857600
2018-07-14 20:54:54,858 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:54:54,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:54:54,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local867678939_0109_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:54:54,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:54:54,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local867678939_0109_m_000000_0' done.
2018-07-14 20:54:54,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local867678939_0109_m_000000_0
2018-07-14 20:54:54,860 [Thread-3060] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:54:54,861 [Thread-3060] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:54:54,861 [pool-306-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local867678939_0109_r_000000_0
2018-07-14 20:54:54,866 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:54:54,866 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@47e08a60
2018-07-14 20:54:54,868 [pool-306-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:54:54,868 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local867678939_0109_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:54:54,869 [localfetcher#88] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#88 about to shuffle
output of map attempt_local867678939_0109_m_000000_0 decomp: 101 len: 105 to MEMORY
2018-07-14 20:54:54,869 [localfetcher#88] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 101 bytes from
map-output for attempt_local867678939_0109_m_000000_0
2018-07-14 20:54:54,869 [localfetcher#88] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 101, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->101
2018-07-14 20:54:54,870 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:54:54,870 [pool-306-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:54:54,870 [pool-306-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:54:54,871 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
```

```
2018-07-14 20:54:54,871 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:54:54,872 [pool-306-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 101 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:54:54,872 [pool-306-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 105 bytes from
disk
2018-07-14 20:54:54,872 [pool-306-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:54:54,872 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:54:54,872 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 20:54:54,872 [pool-306-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:54:54,878 [pool-306-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:54:54,878 [pool-306-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:54:54,879 [pool-306-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:54:54,880 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local867678939_0109_r_000000_0 is done. And is in the process of committing
2018-07-14 20:54:54,881 [pool-306-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:54:54,881 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local867678939_0109_r_000000_0 is allowed to commit now
2018-07-14 20:54:54,882 [pool-306-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local867678939_0109_r_000000_0' to
file:/tmp/temp-1902294536/tmp535435342/_temporary/0/task_local867678939_0109_r_000000
2018-07-14 20:54:54,883 [pool-306-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:54:54,883 [pool-306-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local867678939_0109_r_000000_0' done.
2018-07-14 20:54:54,883 [pool-306-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local867678939_0109_r_000000_0
2018-07-14 20:54:54,883 [Thread-3060] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:54:54,998 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local867678939_0109
2018-07-14 20:54:54,998 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:54:54,998 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_g[37,6] C:  R: ex_sum[-1,-1]
2018-07-14 20:54:54,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:54:54,999 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,000 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,000 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:54:55,001 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:54:55,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:54:55,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:54:55,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:54:55,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=317
2018-07-14 20:54:55,002 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:54:55,002 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:54:55,009 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:54:55,010 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,011 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:54:55,011 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:54:55,017 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:54:55,026 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:54:55,026 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:54:55,026 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:54:55,027 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:54:55,027 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:54:55,027 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:54:55,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:54:55,048 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:54:55,064 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local522992374_0110
2018-07-14 20:54:55,124 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:54:55,125 [Thread-3089] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:54:55,128 [Thread-3089] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:54:55,128 [Thread-3089] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:54:55,128 [Thread-3089] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:54:55,128 [Thread-3089] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:54:55,128 [Thread-3089] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:54:55,128 [Thread-3089] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:54:55,130 [Thread-3089] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:54:55,131 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local522992374_0110_m_000000_0
2018-07-14 20:54:55,135 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:54:55,135 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:54:55,137 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:54:55,137 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:54:55,152 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:54:55,152 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:54:55,152 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:54:55,152 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:54:55,152 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:54:55,153 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:54:55,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:54:55,154 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:54:55,156 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R: ans4[57,7]
2018-07-14 20:54:55,159 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:54:55,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:54:55,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:54:55,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 20:54:55,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:54:55,161 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:54:55,163 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local522992374_0110_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:54:55,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:54:55,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local522992374_0110_m_000000_0' done.
2018-07-14 20:54:55,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local522992374_0110_m_000000_0
2018-07-14 20:54:55,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local522992374_0110_m_000001_0
2018-07-14 20:54:55,166 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:54:55,166 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 42
Input split[0]:
   Length = 42
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:54:55,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp535435342/part-r-00000:0+42
2018-07-14 20:54:55,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:54:55,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:54:55,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:54:55,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:54:55,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:54:55,284 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:54:55,285 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:54:55,285 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:54:55,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R: ans4[57,7]
2018-07-14 20:54:55,287 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:54:55,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:54:55,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:54:55,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 63; bufvoid = 104857600
2018-07-14 20:54:55,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:54:55,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:54:55,289 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local522992374_0110_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:54:55,289 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:54:55,289 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local522992374_0110_m_000001_0' done.
2018-07-14 20:54:55,289 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local522992374_0110_m_000001_0
2018-07-14 20:54:55,289 [Thread-3089] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:54:55,290 [Thread-3089] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:54:55,290 [pool-309-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local522992374_0110_r_000000_0
2018-07-14 20:54:55,294 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:54:55,294 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@745c6839
2018-07-14 20:54:55,295 [pool-309-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:54:55,296 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local522992374_0110_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:54:55,297 [localfetcher#89] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#89 about to shuffle
output of map attempt_local522992374_0110_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 20:54:55,297 [localfetcher#89] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local522992374_0110_m_000000_0
2018-07-14 20:54:55,297 [localfetcher#89] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->286
2018-07-14 20:54:55,298 [localfetcher#89] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#89 about to shuffle
output of map attempt_local522992374_0110_m_000001_0 decomp: 79 len: 83 to MEMORY
2018-07-14 20:54:55,298 [localfetcher#89] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 79 bytes from map-output
for attempt_local522992374_0110_m_000001_0
2018-07-14 20:54:55,298 [localfetcher#89] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 79, inMemoryMapOutputs.size() -> 2, commitMemory -> 286, usedMemory
->365
2018-07-14 20:54:55,298 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
```

```
2018-07-14 20:54:55,299 [pool-309-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:54:55,299 [pool-309-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:54:55,300 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:54:55,300 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 349 bytes
2018-07-14 20:54:55,301 [pool-309-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 365 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:54:55,301 [pool-309-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 367 bytes from
disk
2018-07-14 20:54:55,301 [pool-309-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:54:55,301 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:54:55,301 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 355 bytes
2018-07-14 20:54:55,301 [pool-309-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:54:55,307 [pool-309-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:54:55,307 [pool-309-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:54:55,309 [pool-309-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R: ans4[57,7]
2018-07-14 20:54:55,310 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local522992374_0110_r_000000_0 is done. And is in the process of committing
2018-07-14 20:54:55,311 [pool-309-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:54:55,311 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local522992374_0110_r_000000_0 is allowed to commit now
2018-07-14 20:54:55,312 [pool-309-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local522992374_0110_r_000000_0' to
file:/tmp/temp-1902294536/tmp1824748787/_temporary/0/task_local522992374_0110_r_000000
2018-07-14 20:54:55,312 [pool-309-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:54:55,312 [pool-309-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local522992374_0110_r_000000_0' done.
2018-07-14 20:54:55,312 [pool-309-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local522992374_0110_r_000000_0
2018-07-14 20:54:55,312 [Thread-3089] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:54:55,472 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local522992374_0110
2018-07-14 20:54:55,472 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,ans4,details
2018-07-14 20:54:55,472 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans3[52,7],details[8,10],details[-1,-1],ans3[52,7] C:  R:
ans4[57,7]
2018-07-14 20:54:55,479 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,480 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,481 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,484 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:54:55,485 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:54:54     2018-07-14 20:54:55
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local522992374_0110 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans3,ans4,details       HASH_JOIN
file:/tmp/temp-1902294536/tmp1824748787,
job_local867678939_0109 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 6 records in: "file:/tmp/temp-1902294536/tmp1824748787"

Counters:
Total records written : 6
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local867678939_0109 ->      job_local522992374_0110,
job_local522992374_0110


2018-07-14 20:54:55,486 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,487 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,488 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,490 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,491 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,491 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:54:55,492 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:54:55,492 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 20:54:55,493 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:54:55,493 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:54:55,501 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:54:55,501 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh)
(102,Shahrukh)
(104,Anubhav)
(105,Pawan)
(110,Priyanka)
(114,Madhuri)
grunt> ans5 = JOIN details BY RIGHT OUTER, ex_sum BY EmpId
>> ;
2018-07-14 20:57:11,374 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
58, column 29>  mismatched input 'OUTER' expecting LEFT_PAREN
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans5 = JOIN details BY EmpId RIGHT OUTER, ex_sum BY EmpId
>> ;
grunt> dump ans5;
2018-07-14 20:57:35,774 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:57:35,793 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:35,794 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:57:35,794 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:57:35,794 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:57:35,798 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:57:35,799 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:57:35,799 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:57:35,799 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:57:35,807 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:35,807 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:57:35,808 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:35,809 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:57:35,810 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:57:35,810 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:57:35,810 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:57:35,810 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:57:35,810 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:57:35,812 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:57:35,813 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:57:35,813 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:57:35,813 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582055812-0
2018-07-14 20:57:35,822 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:57:35,823 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:35,832 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:57:35,833 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:57:35,834 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:57:35,834 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:57:35,834 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:57:35,854 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:57:35,869 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local74919530_0111
2018-07-14 20:57:35,934 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:57:35,935 [Thread-3119] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:57:35,939 [Thread-3119] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:57:35,939 [Thread-3119] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:57:35,939 [Thread-3119] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:35,939 [Thread-3119] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:57:35,939 [Thread-3119] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 20:57:35,939 [Thread-3119] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:57:35,941 [Thread-3119] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:57:35,941 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local74919530_0111_m_000000_0
2018-07-14 20:57:35,946 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:35,946 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:57:35,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:57:35,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:57:36,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:57:36,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:57:36,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:57:36,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:57:36,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:57:36,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:57:36,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:36,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:57:36,010 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:57:36,011 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:57:36,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:57:36,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:57:36,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:57:36,012 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:57:36,020 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:57:36,021 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task:attempt_local74919530_0111_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:57:36,022 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:57:36,022 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local74919530_0111_m_000000_0' done.
2018-07-14 20:57:36,022 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local74919530_0111_m_000000_0
2018-07-14 20:57:36,022 [Thread-3119] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:57:36,022 [Thread-3119] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:57:36,022 [pool-312-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local74919530_0111_r_000000_0
2018-07-14 20:57:36,028 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:36,028 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4b439ae3
2018-07-14 20:57:36,029 [pool-312-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:57:36,030 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local74919530_0111_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:57:36,031 [localfetcher#90] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#90 about to shuffle
output of map attempt_local74919530_0111_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:57:36,031 [localfetcher#90] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local74919530_0111_m_000000_0
2018-07-14 20:57:36,031 [localfetcher#90] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:57:36,032 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:57:36,032 [pool-312-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:57:36,032 [pool-312-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:57:36,033 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:57:36,033 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:57:36,034 [pool-312-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:57:36,034 [pool-312-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:57:36,034 [pool-312-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:57:36,034 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:57:36,034 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:57:36,034 [pool-312-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:57:36,040 [pool-312-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:36,041 [pool-312-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:57:36,043 [pool-312-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:57:36,044 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local74919530_0111_r_000000_0 is done. And is in the process of committing
2018-07-14 20:57:36,045 [pool-312-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:57:36,045 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local74919530_0111_r_000000_0 is allowed to commit now
2018-07-14 20:57:36,046 [pool-312-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local74919530_0111_r_000000_0' to
file:/tmp/temp-1902294536/tmp858120994/_temporary/0/task_local74919530_0111_r_000000
2018-07-14 20:57:36,047 [pool-312-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:57:36,047 [pool-312-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local74919530_0111_r_000000_0' done.
2018-07-14 20:57:36,047 [pool-312-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local74919530_0111_r_000000_0
2018-07-14 20:57:36,053 [Thread-3119] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:57:36,237 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local74919530_0111
2018-07-14 20:57:36,237 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:57:36,237 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:57:36,238 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:57:36,238 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,239 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,239 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,240 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:57:36,240 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:57:36,241 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:57:36,241 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:57:36,241 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 20:57:36,241 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:57:36,241 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:57:36,251 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:57:36,253 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,254 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:36,254 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:57:36,260 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:57:36,268 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:57:36,268 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:57:36,268 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:57:36,269 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:57:36,269 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:57:36,269 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:57:36,270 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:57:36,291 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:57:36,305 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1782003613_0112
2018-07-14 20:57:36,379 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:57:36,380 [Thread-3148] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:57:36,386 [Thread-3148] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:57:36,386 [Thread-3148] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:57:36,386 [Thread-3148] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:36,386 [Thread-3148] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:57:36,386 [Thread-3148] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
```

```
Instead, use fs.defaultFS
2018-07-14 20:57:36,386 [Thread-3148] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:57:36,389 [Thread-3148] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:57:36,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1782003613_0112_m_000000_0
2018-07-14 20:57:36,396 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:36,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:57:36,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:57:36,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:57:36,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:57:36,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:57:36,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:57:36,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:57:36,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:57:36,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:57:36,422 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:36,422 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:57:36,424 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[58,7],details[8,10],details[-1,-1],ans5[58,7] C:  R: ans5[58,7]
2018-07-14 20:57:36,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:57:36,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:57:36,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:57:36,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:57:36,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:57:36,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:57:36,426 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1782003613_0112_m_000000_0 is done. And
```

```
is in the process of committing
2018-07-14 20:57:36,427 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:57:36,427 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1782003613_0112_m_000000_0' done.
2018-07-14 20:57:36,427 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1782003613_0112_m_000000_0
2018-07-14 20:57:36,427 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1782003613_0112_m_000001_0
2018-07-14 20:57:36,428 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:36,429 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:57:36,430 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp858120994/part-r-00000:0+62
2018-07-14 20:57:36,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:57:36,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:57:36,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:57:36,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:57:36,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:57:36,584 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:57:36,585 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:36,585 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:57:36,586 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[58,7],details[8,10],details[-1,-1],ans5[58,7] C:  R: ans5[58,7]
2018-07-14 20:57:36,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:57:36,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:57:36,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:57:36,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 20:57:36,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:57:36,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:57:36,588 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1782003613_0112_m_000001_0 is done. And
```

```
is in the process of committing
2018-07-14 20:57:36,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:57:36,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1782003613_0112_m_000001_0' done.
2018-07-14 20:57:36,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1782003613_0112_m_000001_0
2018-07-14 20:57:36,589 [Thread-3148] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:57:36,589 [Thread-3148] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:57:36,589 [pool-315-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1782003613_0112_r_000000_0
2018-07-14 20:57:36,595 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:36,595 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@d170dfc
2018-07-14 20:57:36,595 [pool-315-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:57:36,596 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1782003613_0112_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:57:36,597 [localfetcher#91] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#91 about to shuffle
output of map attempt_local1782003613_0112_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:57:36,597 [localfetcher#91] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1782003613_0112_m_000000_0
2018-07-14 20:57:36,597 [localfetcher#91] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 20:57:36,598 [localfetcher#91] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#91 about to shuffle
output of map attempt_local1782003613_0112_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 20:57:36,598 [localfetcher#91] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1782003613_0112_m_000001_0
2018-07-14 20:57:36,598 [localfetcher#91] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 20:57:36,598 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:57:36,598 [pool-315-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:57:36,599 [pool-315-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:57:36,599 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:57:36,599 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 20:57:36,600 [pool-315-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:57:36,600 [pool-315-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
```

```
disk
2018-07-14 20:57:36,600 [pool-315-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:57:36,600 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:57:36,600 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 20:57:36,600 [pool-315-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:57:36,607 [pool-315-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:36,608 [pool-315-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:57:36,610 [pool-315-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[58,7],details[8,10],details[-1,-1],ans5[58,7] C:  R: ans5[58,7]
2018-07-14 20:57:36,611 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1782003613_0112_r_000000_0 is done. And is in the process of committing
2018-07-14 20:57:36,613 [pool-315-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:57:36,613 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1782003613_0112_r_000000_0 is allowed to commit now
2018-07-14 20:57:36,613 [pool-315-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1782003613_0112_r_000000_0' to
file:/tmp/temp-1902294536/tmp1744388008/_temporary/0/task_local1782003613_0112_r_00000
0
2018-07-14 20:57:36,614 [pool-315-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:57:36,614 [pool-315-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1782003613_0112_r_000000_0' done.
2018-07-14 20:57:36,614 [pool-315-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1782003613_0112_r_000000_0
2018-07-14 20:57:36,614 [Thread-3148] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:57:36,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1782003613_0112
2018-07-14 20:57:36,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 20:57:36,752 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[58,7],details[8,10],details[-1,-1],ans5[58,7] C:  R:
ans5[58,7]
2018-07-14 20:57:36,754 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,754 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,755 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,757 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:57:36,758 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:57:35     2018-07-14 20:57:36
```

```
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1782003613_0112        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp-1902294536/tmp1744388008,
job_local74919530_0111  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp1744388008"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local74919530_0111  ->      job_local1782003613_0112,
job_local1782003613_0112


2018-07-14 20:57:36,758 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,759 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,759 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,761 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,761 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,762 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:36,764 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:57:36,764 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:36,764 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:57:36,764 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:57:36,774 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:57:36,774 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
```

```
(110,Priyanka,2000,5,110,400)
(114,Madhuri,2000,2,114,200)
(,,,,119,200)
grunt> dump ex_sum;
2018-07-14 20:57:56,460 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:57:56,474 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:56,475 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:57:56,476 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:57:56,476 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:57:56,477 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:57:56,478 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:57:56,480 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:57:56,480 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:57:56,487 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:56,488 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:57:56,488 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:56,489 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:57:56,489 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:57:56,489 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:57:56,489 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:57:56,489 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:57:56,489 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:57:56,490 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:57:56,490 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:57:56,490 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:57:56,490 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
```

```
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582076490-0
2018-07-14 20:57:56,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:57:56,500 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:56,507 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:57:56,508 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:57:56,509 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:57:56,509 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:57:56,509 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:57:56,531 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:57:56,545 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local547305171_0113
2018-07-14 20:57:56,608 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:57:56,608 [Thread-3178] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:57:56,612 [Thread-3178] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:57:56,612 [Thread-3178] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:57:56,612 [Thread-3178] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:56,612 [Thread-3178] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:57:56,612 [Thread-3178] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:57:56,612 [Thread-3178] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:57:56,614 [Thread-3178] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:57:56,615 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local547305171_0113_m_000000_0
2018-07-14 20:57:56,620 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:56,621 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------
```

```
2018-07-14 20:57:56,622 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:57:56,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:57:56,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:57:56,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:57:56,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:57:56,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:57:56,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:57:56,639 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:57:56,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:56,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:57:56,641 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:57:56,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:57:56,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:57:56,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:57:56,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:57:56,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:57:56,652 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:57:56,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local547305171_0113_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:57:56,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:57:56,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local547305171_0113_m_000000_0' done.
2018-07-14 20:57:56,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local547305171_0113_m_000000_0
2018-07-14 20:57:56,654 [Thread-3178] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:57:56,654 [Thread-3178] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:57:56,655 [pool-318-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local547305171_0113_r_000000_0
2018-07-14 20:57:56,662 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:57:56,662 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
```

```
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3f77423a
2018-07-14 20:57:56,663 [pool-318-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:57:56,663 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local547305171_0113_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:57:56,664 [localfetcher#92] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#92 about to shuffle
output of map attempt_local547305171_0113_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:57:56,664 [localfetcher#92] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local547305171_0113_m_000000_0
2018-07-14 20:57:56,664 [localfetcher#92] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:57:56,665 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:57:56,667 [pool-318-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:57:56,667 [pool-318-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:57:56,667 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:57:56,667 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:57:56,668 [pool-318-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:57:56,668 [pool-318-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:57:56,668 [pool-318-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:57:56,668 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:57:56,668 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:57:56,668 [pool-318-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:57:56,674 [pool-318-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:57:56,674 [pool-318-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:57:56,812 [pool-318-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:57:56,813 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local547305171_0113_r_000000_0 is done. And is in the process of committing
2018-07-14 20:57:56,814 [pool-318-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:57:56,814 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local547305171_0113_r_000000_0 is allowed to commit now
2018-07-14 20:57:56,815 [pool-318-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
```

```
'attempt_local547305171_0113_r_000000_0' to
file:/tmp/temp-1902294536/tmp-519387951/_temporary/0/task_local547305171_0113_r_000000
2018-07-14 20:57:56,815 [pool-318-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:57:56,816 [pool-318-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local547305171_0113_r_000000_0' done.
2018-07-14 20:57:56,816 [pool-318-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local547305171_0113_r_000000_0
2018-07-14 20:57:56,816 [Thread-3178] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:57:56,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local547305171_0113
2018-07-14 20:57:56,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:57:57,000 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:57:57,009 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:57,012 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:57,014 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:57,021 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:57:57,026 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:57:56     2018-07-14 20:57:57
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local547305171_0113 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-519387951,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-519387951"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local547305171_0113
```

```
2018-07-14 20:57:57,030 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:57,035 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:57,036 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:57:57,040 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:57:57,040 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:57:57,041 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:57:57,041 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:57:57,054 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:57:57,055 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> ans5 = JOIN details BY EmpId LEFT OUTER, ex_sum BY EmpId
>> ;
grunt> dump ans5;
2018-07-14 20:58:12,797 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 20:58:12,810 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:12,810 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:58:12,810 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:58:12,810 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:58:12,812 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:58:12,813 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:58:12,814 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 20:58:12,814 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 20:58:12,822 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:12,823 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:58:12,823 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 20:58:12,824 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:58:12,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:58:12,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:58:12,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:58:12,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:58:12,824 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:58:12,825 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:58:12,825 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:58:12,825 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:58:12,825 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582092825-0
2018-07-14 20:58:12,834 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:58:12,835 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:12,842 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:58:12,843 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:58:12,844 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:58:12,844 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:58:12,844 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:58:12,864 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:58:12,878 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local836177631_0114
2018-07-14 20:58:12,949 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:58:12,950 [Thread-3207] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:58:12,954 [Thread-3207] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:58:12,954 [Thread-3207] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:58:12,955 [Thread-3207] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:12,955 [Thread-3207] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:58:12,955 [Thread-3207] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:58:12,955 [Thread-3207] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:58:12,957 [Thread-3207] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:58:12,957 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local836177631_0114_m_000000_0
2018-07-14 20:58:12,961 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:12,962 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:58:12,964 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:58:12,964 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:58:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:58:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:58:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:58:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:58:12,979 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:58:12,980 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:58:12,980 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:12,981 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:58:12,983 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:58:12,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:58:12,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:58:12,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
```

```
2018-07-14 20:58:12,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:58:12,984 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:58:12,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:58:12,995 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local836177631_0114_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:58:12,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:58:12,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local836177631_0114_m_000000_0' done.
2018-07-14 20:58:12,996 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local836177631_0114_m_000000_0
2018-07-14 20:58:12,996 [Thread-3207] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:58:12,996 [Thread-3207] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:58:12,997 [pool-321-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local836177631_0114_r_000000_0
2018-07-14 20:58:13,002 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:13,002 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7f511cc8
2018-07-14 20:58:13,003 [pool-321-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:58:13,004 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local836177631_0114_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:58:13,005 [localfetcher#93] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#93 about to shuffle
output of map attempt_local836177631_0114_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:58:13,005 [localfetcher#93] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local836177631_0114_m_000000_0
2018-07-14 20:58:13,005 [localfetcher#93] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:58:13,005 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:58:13,005 [pool-321-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:58:13,006 [pool-321-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
```

```
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:58:13,007 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:58:13,008 [pool-321-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:58:13,012 [pool-321-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:13,013 [pool-321-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:58:13,014 [pool-321-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:58:13,015 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local836177631_0114_r_000000_0 is done. And is in the process of committing
2018-07-14 20:58:13,015 [pool-321-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:58:13,015 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local836177631_0114_r_000000_0 is allowed to commit now
2018-07-14 20:58:13,017 [pool-321-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local836177631_0114_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1152963760/_temporary/0/task_local836177631_0114_r_00000
0
2018-07-14 20:58:13,017 [pool-321-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:58:13,017 [pool-321-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local836177631_0114_r_000000_0' done.
2018-07-14 20:58:13,017 [pool-321-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local836177631_0114_r_000000_0
2018-07-14 20:58:13,017 [Thread-3207] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:58:13,152 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local836177631_0114
2018-07-14 20:58:13,153 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:58:13,153 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:58:13,156 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 20:58:13,160 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,163 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,166 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,180 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:58:13,181 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:58:13,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:58:13,182 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:58:13,185 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 20:58:13,185 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:58:13,187 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:58:13,202 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:58:13,204 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,205 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:13,205 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:58:13,212 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:58:13,213 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:58:13,213 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:58:13,214 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:58:13,214 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:58:13,221 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:58:13,221 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:58:13,221 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:58:13,242 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 20:58:13,257 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2001968175_0115
2018-07-14 20:58:13,313 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:58:13,314 [Thread-3236] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:58:13,319 [Thread-3236] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:58:13,319 [Thread-3236] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
```

```
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:58:13,319 [Thread-3236] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:13,319 [Thread-3236] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:58:13,319 [Thread-3236] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:58:13,319 [Thread-3236] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 20:58:13,321 [Thread-3236] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:58:13,324 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2001968175_0115_m_000000_0
2018-07-14 20:58:13,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:13,328 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:58:13,330 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:58:13,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 20:58:13,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:58:13,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:58:13,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:58:13,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:58:13,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:58:13,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:58:13,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:13,365 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:58:13,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[60,7],details[8,10],details[-1,-1],ans5[60,7] C:  R: ans5[60,7]
2018-07-14 20:58:13,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:58:13,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:58:13,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
```

```
2018-07-14 20:58:13,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 20:58:13,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 20:58:13,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:58:13,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2001968175_0115_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:58:13,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:58:13,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2001968175_0115_m_000000_0' done.
2018-07-14 20:58:13,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2001968175_0115_m_000000_0
2018-07-14 20:58:13,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2001968175_0115_m_000001_0
2018-07-14 20:58:13,373 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:13,373 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:58:13,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1152963760/part-r-00000:0+62
2018-07-14 20:58:13,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:58:13,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:58:13,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:58:13,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:58:13,389 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:58:13,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:58:13,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:13,391 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 20:58:13,392 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[60,7],details[8,10],details[-1,-1],ans5[60,7] C:  R: ans5[60,7]
2018-07-14 20:58:13,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:58:13,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:58:13,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
```

```
2018-07-14 20:58:13,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 20:58:13,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 20:58:13,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:58:13,395 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2001968175_0115_m_000001_0 is done. And
is in the process of committing
2018-07-14 20:58:13,396 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:58:13,396 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2001968175_0115_m_000001_0' done.
2018-07-14 20:58:13,396 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2001968175_0115_m_000001_0
2018-07-14 20:58:13,396 [Thread-3236] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:58:13,397 [Thread-3236] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:58:13,397 [pool-324-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2001968175_0115_r_000000_0
2018-07-14 20:58:13,404 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:13,405 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5dac40a0
2018-07-14 20:58:13,405 [pool-324-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:58:13,406 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2001968175_0115_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:58:13,407 [localfetcher#94] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#94 about to shuffle
output of map attempt_local2001968175_0115_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 20:58:13,407 [localfetcher#94] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local2001968175_0115_m_000001_0
2018-07-14 20:58:13,408 [localfetcher#94] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 20:58:13,408 [localfetcher#94] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#94 about to shuffle
output of map attempt_local2001968175_0115_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 20:58:13,408 [localfetcher#94] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local2001968175_0115_m_000000_0
2018-07-14 20:58:13,408 [localfetcher#94] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 20:58:13,409 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:58:13,409 [pool-324-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:58:13,409 [pool-324-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
```

```
2018-07-14 20:58:13,410 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 20:58:13,410 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 20:58:13,412 [pool-324-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:58:13,412 [pool-324-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 20:58:13,412 [pool-324-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:58:13,412 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:58:13,412 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 20:58:13,412 [pool-324-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:58:13,419 [pool-324-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:13,419 [pool-324-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:58:13,421 [pool-324-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[60,7],details[8,10],details[-1,-1],ans5[60,7] C:  R: ans5[60,7]
2018-07-14 20:58:13,422 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2001968175_0115_r_000000_0 is done. And is in the process of committing
2018-07-14 20:58:13,423 [pool-324-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 20:58:13,423 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local2001968175_0115_r_000000_0 is allowed to commit now
2018-07-14 20:58:13,424 [pool-324-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2001968175_0115_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1017842867/_temporary/0/task_local2001968175_0115_r_0000
00
2018-07-14 20:58:13,424 [pool-324-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:58:13,424 [pool-324-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2001968175_0115_r_000000_0' done.
2018-07-14 20:58:13,424 [pool-324-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2001968175_0115_r_000000_0
2018-07-14 20:58:13,425 [Thread-3236] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:58:13,618 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2001968175_0115
2018-07-14 20:58:13,618 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 20:58:13,618 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[60,7],details[8,10],details[-1,-1],ans5[60,7] C:  R:
ans5[60,7]
2018-07-14 20:58:13,635 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,638 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,641 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,650 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:58:13,655 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:58:12     2018-07-14 20:58:13
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local2001968175_0115        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp-1902294536/tmp-1017842867,
job_local836177631_0114 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp-1017842867"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local836177631_0114 ->      job_local2001968175_0115,
job_local2001968175_0115


2018-07-14 20:58:13,658 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,659 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,660 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,669 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,671 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,672 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:13,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:58:13,677 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:13,678 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:58:13,678 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
```

```
2018-07-14 20:58:13,689 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:58:13,689 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(103,Akshay,11000,3,,)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(106,Aamir,25000,1,,)
(107,Salman,17500,2,,)
(108,Ranbir,14000,3,,)
(109,Katrina,1000,4,,)
(110,Priyanka,2000,5,110,400)
(111,Tushar,500,1,,)
(112,Ajay,5000,2,,)
(113,Jubeen,1000,1,,)
(114,Madhuri,2000,2,114,200)
grunt> dump ex_sum;
2018-07-14 20:58:55,208 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 20:58:55,227 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:55,227 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:58:55,227 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 20:58:55,227 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 20:58:55,229 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 20:58:55,230 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 20:58:55,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 20:58:55,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 20:58:55,241 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:55,241 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:58:55,242 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,243 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 20:58:55,243 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 20:58:55,243 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 20:58:55,243 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
```

```
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 20:58:55,244 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 20:58:55,244 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 20:58:55,245 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 20:58:55,245 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 20:58:55,245 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 20:58:55,245 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582135245-0
2018-07-14 20:58:55,256 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 20:58:55,257 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,266 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 20:58:55,267 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 20:58:55,267 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:58:55,267 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 20:58:55,268 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 20:58:55,288 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 20:58:55,303 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local844924555_0116
2018-07-14 20:58:55,365 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 20:58:55,366 [Thread-3266] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 20:58:55,369 [Thread-3266] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 20:58:55,369 [Thread-3266] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 20:58:55,369 [Thread-3266] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:55,369 [Thread-3266] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 20:58:55,369 [Thread-3266] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 20:58:55,370 [Thread-3266] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
```

```
2018-07-14 20:58:55,371 [Thread-3266] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 20:58:55,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local844924555_0116_m_000000_0
2018-07-14 20:58:55,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:55,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 20:58:55,378 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 20:58:55,378 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 20:58:55,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 20:58:55,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 20:58:55,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 20:58:55,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 20:58:55,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 20:58:55,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 20:58:55,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:55,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 20:58:55,401 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:58:55,401 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 20:58:55,401 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 20:58:55,401 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 20:58:55,401 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 20:58:55,401 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 20:58:55,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 20:58:55,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local844924555_0116_m_000000_0 is done. And
is in the process of committing
2018-07-14 20:58:55,411 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 20:58:55,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local844924555_0116_m_000000_0' done.
2018-07-14 20:58:55,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local844924555_0116_m_000000_0
2018-07-14 20:58:55,412 [Thread-3266] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 20:58:55,412 [Thread-3266] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 20:58:55,412 [pool-327-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local844924555_0116_r_000000_0
2018-07-14 20:58:55,416 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 20:58:55,417 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3626545a
2018-07-14 20:58:55,417 [pool-327-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 20:58:55,418 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local844924555_0116_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 20:58:55,419 [localfetcher#95] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#95 about to shuffle
output of map attempt_local844924555_0116_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 20:58:55,420 [localfetcher#95] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local844924555_0116_m_000000_0
2018-07-14 20:58:55,420 [localfetcher#95] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 20:58:55,420 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 20:58:55,421 [pool-327-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:58:55,421 [pool-327-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 20:58:55,422 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:58:55,422 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:58:55,423 [pool-327-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 20:58:55,423 [pool-327-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 20:58:55,423 [pool-327-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 20:58:55,423 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 20:58:55,423 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 20:58:55,423 [pool-327-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:58:55,430 [pool-327-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
```

```
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 20:58:55,430 [pool-327-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 20:58:55,431 [pool-327-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 20:58:55,432 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local844924555_0116_r_000000_0 is done. And is in the process of committing
2018-07-14 20:58:55,433 [pool-327-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 20:58:55,433 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local844924555_0116_r_000000_0 is allowed to commit now
2018-07-14 20:58:55,434 [pool-327-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local844924555_0116_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1143097085/_temporary/0/task_local844924555_0116_r_00000
0
2018-07-14 20:58:55,434 [pool-327-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 20:58:55,434 [pool-327-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local844924555_0116_r_000000_0' done.
2018-07-14 20:58:55,434 [pool-327-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local844924555_0116_r_000000_0
2018-07-14 20:58:55,434 [Thread-3266] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 20:58:55,569 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local844924555_0116
2018-07-14 20:58:55,569 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 20:58:55,569 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 20:58:55,577 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,581 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,584 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,589 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 20:58:55,589 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 20:58:55     2018-07-14 20:58:55
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local844924555_0116 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp-1902294536/tmp-1143097085,
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1143097085"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local844924555_0116


2018-07-14 20:58:55,590 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,592 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,593 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 20:58:55,595 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 20:58:55,595 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 20:58:55,596 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 20:58:55,596 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 20:58:55,609 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 20:58:55,609 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt> ans5 = foreach details GENERATE EmpId==expenses.EmpId;
2018-07-14 21:01:22,911 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
62, column 32>  Syntax error, unexpected symbol at or near 'EmpId'
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans5 = JOIN ex_sum BY EmpID LEFT OUTER, details By EmpId;
2018-07-14 21:02:54,936 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 62, column 22> Invalid field projection. Projected field [EmpID] does not exist in
schema: EmpId:int,sum:double.
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans5 = JOIN ex_sum BY EmpId LEFT OUTER, details By EmpId;
grunt>
grunt>
grunt> dump ans5;
2018-07-14 21:03:04,860 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:03:04,877 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:04,877 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:03:04,877 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:03:04,877 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:03:04,878 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:03:04,879 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:03:04,880 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:03:04,880 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:03:04,889 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:04,890 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:03:04,890 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:04,891 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:03:04,891 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:03:04,891 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:03:04,891 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:03:04,891 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:03:04,891 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:03:04,892 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:03:04,893 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:03:04,894 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:03:04,894 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582384893-0
2018-07-14 21:03:04,902 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:03:04,903 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:04,911 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
```

```
not be found. See Job or Job#setJar(String).
2018-07-14 21:03:04,913 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:03:04,913 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:04,913 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:03:04,913 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:03:04,934 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:03:04,948 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local870472728_0117
2018-07-14 21:03:05,008 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:03:05,008 [Thread-3295] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:03:05,012 [Thread-3295] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:03:05,012 [Thread-3295] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:03:05,013 [Thread-3295] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:05,013 [Thread-3295] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:03:05,013 [Thread-3295] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:03:05,013 [Thread-3295] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:03:05,015 [Thread-3295] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:03:05,015 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local870472728_0117_m_000000_0
2018-07-14 21:03:05,019 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:05,020 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:03:05,021 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:03:05,021 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:03:05,048 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:03:05,048 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:03:05,048 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:03:05,048 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:03:05,048 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:03:05,049 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:03:05,050 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:05,050 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:03:05,053 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:03:05,053 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:03:05,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:03:05,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:03:05,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:03:05,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:03:05,060 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:03:05,061 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local870472728_0117_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:03:05,062 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:03:05,062 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local870472728_0117_m_000000_0' done.
2018-07-14 21:03:05,062 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local870472728_0117_m_000000_0
2018-07-14 21:03:05,062 [Thread-3295] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:03:05,062 [Thread-3295] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:03:05,062 [pool-330-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local870472728_0117_r_000000_0
2018-07-14 21:03:05,067 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:05,067 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3298bd7f
2018-07-14 21:03:05,067 [pool-330-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:03:05,068 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local870472728_0117_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
```

```
2018-07-14 21:03:05,072 [localfetcher#96] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#96 about to shuffle
output of map attempt_local870472728_0117_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:03:05,072 [localfetcher#96] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local870472728_0117_m_000000_0
2018-07-14 21:03:05,072 [localfetcher#96] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:03:05,073 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:03:05,073 [pool-330-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:03:05,073 [pool-330-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:03:05,074 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:03:05,074 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:03:05,075 [pool-330-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:03:05,075 [pool-330-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:03:05,075 [pool-330-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:03:05,075 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:03:05,075 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:03:05,075 [pool-330-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:03:05,082 [pool-330-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:05,082 [pool-330-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:03:05,085 [pool-330-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:03:05,086 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local870472728_0117_r_000000_0 is done. And is in the process of committing
2018-07-14 21:03:05,087 [pool-330-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:03:05,087 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local870472728_0117_r_000000_0 is allowed to commit now
2018-07-14 21:03:05,088 [pool-330-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local870472728_0117_r_000000_0' to
file:/tmp/temp-1902294536/tmp-392756560/_temporary/0/task_local870472728_0117_r_000000
2018-07-14 21:03:05,088 [pool-330-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:03:05,088 [pool-330-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local870472728_0117_r_000000_0' done.
2018-07-14 21:03:05,088 [pool-330-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local870472728_0117_r_000000_0
```

```
2018-07-14 21:03:05,088 [Thread-3295] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:03:05,209 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local870472728_0117
2018-07-14 21:03:05,209 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:03:05,209 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:03:05,211 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:03:05,212 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,213 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,214 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,224 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:03:05,224 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:03:05,224 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:03:05,224 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:03:05,225 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:03:05,225 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:03:05,227 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:03:05,237 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:03:05,238 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,240 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:05,240 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:03:05,248 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:03:05,249 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:03:05,249 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:05,249 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
2018-07-14 21:03:05,249 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:03:05,263 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:05,263 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:03:05,263 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:03:05,281 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:03:05,297 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1166491083_0118
2018-07-14 21:03:05,357 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:03:05,358 [Thread-3324] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:03:05,361 [Thread-3324] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:03:05,361 [Thread-3324] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:03:05,362 [Thread-3324] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:05,362 [Thread-3324] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:03:05,362 [Thread-3324] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:03:05,362 [Thread-3324] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:03:05,363 [Thread-3324] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:03:05,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1166491083_0118_m_000000_0
2018-07-14 21:03:05,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:05,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:03:05,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:03:05,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:03:05,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:03:05,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:03:05,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:03:05,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:03:05,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:03:05,387 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:03:05,388 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:05,389 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:03:05,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[62,7],details[8,10],details[-1,-1],ans5[62,7] C:  R: ans5[62,7]
2018-07-14 21:03:05,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:03:05,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:03:05,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:03:05,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:03:05,390 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:03:05,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:03:05,393 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1166491083_0118_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:03:05,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:03:05,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1166491083_0118_m_000000_0' done.
2018-07-14 21:03:05,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1166491083_0118_m_000000_0
2018-07-14 21:03:05,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1166491083_0118_m_000001_0
2018-07-14 21:03:05,397 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:05,398 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:03:05,399 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-392756560/part-r-00000:0+62
2018-07-14 21:03:05,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:03:05,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:03:05,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:03:05,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:03:05,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:03:05,439 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:03:05,440 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:05,440 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:03:05,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[62,7],details[8,10],details[-1,-1],ans5[62,7] C:  R: ans5[62,7]
2018-07-14 21:03:05,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:03:05,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:03:05,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:03:05,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:03:05,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:03:05,444 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:03:05,445 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1166491083_0118_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:03:05,446 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:03:05,446 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1166491083_0118_m_000001_0' done.
2018-07-14 21:03:05,446 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1166491083_0118_m_000001_0
2018-07-14 21:03:05,446 [Thread-3324] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:03:05,446 [Thread-3324] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:03:05,446 [pool-333-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1166491083_0118_r_000000_0
2018-07-14 21:03:05,451 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:05,451 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@419555c8
2018-07-14 21:03:05,452 [pool-333-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:03:05,453 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1166491083_0118_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:03:05,454 [localfetcher#97] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#97 about to shuffle
output of map attempt_local1166491083_0118_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:03:05,454 [localfetcher#97] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1166491083_0118_m_000001_0
2018-07-14 21:03:05,454 [localfetcher#97] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 21:03:05,454 [localfetcher#97] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#97 about to shuffle
output of map attempt_local1166491083_0118_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:03:05,455 [localfetcher#97] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1166491083_0118_m_000000_0
2018-07-14 21:03:05,455 [localfetcher#97] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 21:03:05,455 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:03:05,455 [pool-333-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:03:05,455 [pool-333-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:03:05,457 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:03:05,457 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:03:05,459 [pool-333-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:03:05,459 [pool-333-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:03:05,459 [pool-333-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:03:05,459 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:03:05,459 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:03:05,459 [pool-333-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:03:05,466 [pool-333-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:05,466 [pool-333-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:03:05,467 [pool-333-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[62,7],details[8,10],details[-1,-1],ans5[62,7] C:  R: ans5[62,7]
2018-07-14 21:03:05,468 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1166491083_0118_r_000000_0 is done. And is in the process of committing
2018-07-14 21:03:05,469 [pool-333-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:03:05,470 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1166491083_0118_r_000000_0 is allowed to commit now
2018-07-14 21:03:05,470 [pool-333-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1166491083_0118_r_000000_0' to
```

```
file:/tmp/temp-1902294536/tmp-1058159502/_temporary/0/task_local1166491083_0118_r_0000
00
2018-07-14 21:03:05,473 [pool-333-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:03:05,473 [pool-333-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1166491083_0118_r_000000_0' done.
2018-07-14 21:03:05,473 [pool-333-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1166491083_0118_r_000000_0
2018-07-14 21:03:05,473 [Thread-3324] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:03:05,659 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1166491083_0118
2018-07-14 21:03:05,660 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 21:03:05,660 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[62,7],details[8,10],details[-1,-1],ans5[62,7] C:  R:
ans5[62,7]
2018-07-14 21:03:05,666 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,668 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,668 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,673 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:03:05,674 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:03:04     2018-07-14 21:03:05
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1166491083_0118        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp-1902294536/tmp-1058159502,
job_local870472728_0117 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp-1902294536/tmp-1058159502"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0
```

```
Job DAG:
job_local870472728_0117 ->      job_local1166491083_0118,
job_local1166491083_0118


2018-07-14 21:03:05,675 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,676 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,677 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,678 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,679 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,680 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:05,681 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:03:05,681 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:05,682 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:03:05,682 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:03:05,690 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:05,690 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300,101,Amitabh,20000,1)
(102,500,102,Shahrukh,10000,2)
(104,300,104,Anubhav,5000,4)
(105,100,105,Pawan,2500,5)
(110,400,110,Priyanka,2000,5)
(114,200,114,Madhuri,2000,2)
(119,200,,,,)
grunt> ans5 = JOIN details BY EmpId LEFT OUTER, ex_sum BY EmpId;
grunt> dump ans5;
2018-07-14 21:03:47,164 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:03:47,177 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:47,177 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:03:47,177 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:03:47,177 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:03:47,178 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:03:47,179 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:03:47,180 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
```

```
2018-07-14 21:03:47,180 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:03:47,190 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:47,190 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:03:47,190 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,191 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:03:47,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:03:47,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:03:47,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:03:47,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:03:47,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:03:47,192 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:03:47,192 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:03:47,192 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:03:47,192 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582427192-0
2018-07-14 21:03:47,200 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:03:47,201 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,208 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:03:47,210 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:03:47,210 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:47,210 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:03:47,210 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:03:47,230 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:03:47,244 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1081690599_0119
2018-07-14 21:03:47,303 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:03:47,304 [Thread-3354] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter set in config null
2018-07-14 21:03:47,307 [Thread-3354] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:03:47,308 [Thread-3354] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:03:47,308 [Thread-3354] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:47,308 [Thread-3354] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:03:47,308 [Thread-3354] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:03:47,308 [Thread-3354] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:03:47,310 [Thread-3354] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:03:47,310 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1081690599_0119_m_000000_0
2018-07-14 21:03:47,314 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:47,315 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:03:47,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:03:47,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:03:47,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:03:47,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:03:47,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:03:47,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:03:47,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:03:47,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:03:47,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:47,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:03:47,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
```

```
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:03:47,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:03:47,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:03:47,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:03:47,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:03:47,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:03:47,350 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:03:47,351 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1081690599_0119_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:03:47,352 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:03:47,352 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1081690599_0119_m_000000_0' done.
2018-07-14 21:03:47,352 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1081690599_0119_m_000000_0
2018-07-14 21:03:47,352 [Thread-3354] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:03:47,353 [Thread-3354] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:03:47,353 [pool-336-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1081690599_0119_r_000000_0
2018-07-14 21:03:47,358 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:47,358 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@60a9ca5d
2018-07-14 21:03:47,358 [pool-336-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:03:47,359 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1081690599_0119_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:03:47,360 [localfetcher#98] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#98 about to shuffle
output of map attempt_local1081690599_0119_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:03:47,360 [localfetcher#98] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1081690599_0119_m_000000_0
2018-07-14 21:03:47,360 [localfetcher#98] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:03:47,360 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:03:47,361 [pool-336-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:03:47,361 [pool-336-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:03:47,362 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Merger -
```

```
Merging 1 sorted segments
2018-07-14 21:03:47,362 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:03:47,363 [pool-336-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:03:47,363 [pool-336-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:03:47,363 [pool-336-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:03:47,363 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:03:47,363 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:03:47,363 [pool-336-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:03:47,369 [pool-336-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:47,370 [pool-336-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:03:47,371 [pool-336-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:03:47,372 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1081690599_0119_r_000000_0 is done. And is in the process of committing
2018-07-14 21:03:47,373 [pool-336-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:03:47,373 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1081690599_0119_r_000000_0 is allowed to commit now
2018-07-14 21:03:47,374 [pool-336-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1081690599_0119_r_000000_0' to
file:/tmp/temp-1902294536/tmp1603471986/_temporary/0/task_local1081690599_0119_r_00000
0
2018-07-14 21:03:47,374 [pool-336-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:03:47,374 [pool-336-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1081690599_0119_r_000000_0' done.
2018-07-14 21:03:47,374 [pool-336-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1081690599_0119_r_000000_0
2018-07-14 21:03:47,374 [Thread-3354] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:03:47,505 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1081690599_0119
2018-07-14 21:03:47,505 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:03:47,505 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:03:47,506 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:03:47,506 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 21:03:47,507 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,507 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,509 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:03:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:03:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:03:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:03:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:03:47,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:03:47,510 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:03:47,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:03:47,518 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,520 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:47,520 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:03:47,525 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:03:47,526 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:03:47,527 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:47,527 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:03:47,527 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:03:47,534 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:47,534 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:03:47,534 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:03:47,555 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:03:47,568 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local988613361_0120
2018-07-14 21:03:47,627 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
```

```
track the job: http://localhost:8080/
2018-07-14 21:03:47,628 [Thread-3383] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:03:47,631 [Thread-3383] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:03:47,631 [Thread-3383] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:03:47,631 [Thread-3383] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:47,631 [Thread-3383] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:03:47,631 [Thread-3383] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:03:47,631 [Thread-3383] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:03:47,633 [Thread-3383] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:03:47,633 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local988613361_0120_m_000000_0
2018-07-14 21:03:47,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:47,638 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:03:47,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:03:47,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:03:47,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:03:47,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:03:47,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:03:47,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:03:47,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:03:47,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:03:47,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:47,658 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:03:47,659 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:03:47,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:03:47,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:03:47,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:03:47,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:03:47,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:03:47,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:03:47,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local988613361_0120_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:03:47,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:03:47,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local988613361_0120_m_000000_0' done.
2018-07-14 21:03:47,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local988613361_0120_m_000000_0
2018-07-14 21:03:47,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local988613361_0120_m_000001_0
2018-07-14 21:03:47,664 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:47,664 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:03:47,666 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1603471986/part-r-00000:0+62
2018-07-14 21:03:47,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:03:47,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:03:47,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:03:47,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:03:47,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:03:47,783 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:03:47,784 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:47,784 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:03:47,786 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:03:47,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:03:47,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:03:47,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:03:47,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:03:47,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:03:47,788 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:03:47,789 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local988613361_0120_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:03:47,790 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:03:47,790 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local988613361_0120_m_000001_0' done.
2018-07-14 21:03:47,790 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local988613361_0120_m_000001_0
2018-07-14 21:03:47,790 [Thread-3383] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:03:47,790 [Thread-3383] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:03:47,791 [pool-339-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local988613361_0120_r_000000_0
2018-07-14 21:03:47,795 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:03:47,795 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@429d63f
2018-07-14 21:03:47,796 [pool-339-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:03:47,797 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local988613361_0120_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:03:47,798 [localfetcher#99] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#99 about to shuffle
output of map attempt_local988613361_0120_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:03:47,798 [localfetcher#99] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local988613361_0120_m_000001_0
2018-07-14 21:03:47,798 [localfetcher#99] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 21:03:47,799 [localfetcher#99] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#99 about to shuffle
output of map attempt_local988613361_0120_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:03:47,799 [localfetcher#99] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local988613361_0120_m_000000_0
2018-07-14 21:03:47,799 [localfetcher#99] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
```

```
->451
2018-07-14 21:03:47,799 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:03:47,800 [pool-339-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:03:47,800 [pool-339-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:03:47,801 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:03:47,801 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:03:47,801 [pool-339-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:03:47,802 [pool-339-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:03:47,802 [pool-339-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:03:47,802 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:03:47,802 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:03:47,802 [pool-339-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:03:47,809 [pool-339-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:03:47,809 [pool-339-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:03:47,812 [pool-339-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:03:47,815 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local988613361_0120_r_000000_0 is done. And is in the process of committing
2018-07-14 21:03:47,816 [pool-339-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:03:47,816 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local988613361_0120_r_000000_0 is allowed to commit now
2018-07-14 21:03:47,817 [pool-339-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local988613361_0120_r_000000_0' to
file:/tmp/temp-1902294536/tmp243546928/_temporary/0/task_local988613361_0120_r_000000
2018-07-14 21:03:47,818 [pool-339-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:03:47,818 [pool-339-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local988613361_0120_r_000000_0' done.
2018-07-14 21:03:47,818 [pool-339-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local988613361_0120_r_000000_0
2018-07-14 21:03:47,818 [Thread-3383] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:03:47,971 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local988613361_0120
2018-07-14 21:03:47,971 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 21:03:47,971 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7]
2018-07-14 21:03:47,973 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,973 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,974 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:03:47,978 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:03:47     2018-07-14 21:03:47
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1081690599_0119        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local988613361_0120 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp-1902294536/tmp243546928,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp243546928"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1081690599_0119        ->      job_local988613361_0120,
job_local988613361_0120


2018-07-14 21:03:47,979 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,979 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,979 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,982 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,983 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,983 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:03:47,984 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:03:47,984 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:03:47,984 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:03:47,984 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:03:47,992 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:03:47,992 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(103,Akshay,11000,3,,)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(106,Aamir,25000,1,,)
(107,Salman,17500,2,,)
(108,Ranbir,14000,3,,)
(109,Katrina,1000,4,,)
(110,Priyanka,2000,5,110,400)
(111,Tushar,500,1,,)
(112,Ajay,5000,2,,)
(113,Jubeen,1000,1,,)
(114,Madhuri,2000,2,114,200)
grunt> ans6 = foreach ans5 generate sum==NULL;
2018-07-14 21:05:23,860 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
64, column 29>  Syntax error, unexpected symbol at or near 'sum'
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> describe ans5;
ans5: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::EmpId: int,ex_sum::sum: long}
grunt>
grunt>
grunt> ans6 = foreach ans5 generate sum==0;
2018-07-14 21:05:47,649 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
64, column 29>  Syntax error, unexpected symbol at or near 'sum'
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans6 = foreach ans5 generate NOT sum;
2018-07-14 21:05:56,730 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
64, column 29>  Syntax error, unexpected symbol at or near 'NOT'
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans6 = foreach ans5 generate NOT ex_sum.sum;
2018-07-14 21:06:02,732 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
64, column 29>  Syntax error, unexpected symbol at or near 'NOT'
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans6 = foreach ans5 generate ex_sum.sum!=0;
2018-07-14 21:06:17,093 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
64, column 39>  mismatched input '!=' expecting SEMI_COLON
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans6 = foreach ans5 generate ex_sum.sum==0;
2018-07-14 21:06:24,397 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: <line
64, column 39>  mismatched input '==' expecting SEMI_COLON
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans6 = filter ans5 BY sum==0;
2018-07-14 21:06:51,688 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 1 time(s).
grunt> ans6 = filter ans5 BY ex_sum.sum==0;
2018-07-14 21:07:01,980 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 2 time(s).
grunt> ans6 = filter ans5 BY ex_sum.sum;
```

```
2018-07-14 21:07:10,317 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1058:
<line 66, column 7> Filter's condition must evaluate to boolean. Found: long
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> ans6 = filter ans5 BY ex_sum.sum==NULL;
2018-07-14 21:07:14,659 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 3 time(s).
grunt> dump ans6;
2018-07-14 21:07:20,036 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 1 time(s).
2018-07-14 21:07:20,038 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:07:20,053 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:20,053 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:07:20,053 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:07:20,054 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:07:20,070 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:20,071 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:07:20,074 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:07:20,075 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:07:20,076 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 21:07:20,076 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 21:07:20,076 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 21:07:20,076 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:07:20,087 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:20,087 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:07:20,087 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:20,089 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:07:20,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:07:20,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:07:20,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:07:20,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:07:20,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:07:20,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 21:07:20,091 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:07:20,091 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:07:20,091 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582640091-0
2018-07-14 21:07:20,103 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:07:20,106 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:20,114 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:07:20,116 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:07:20,118 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:07:20,118 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:07:20,118 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:07:20,140 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:07:20,155 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1888978724_0121
2018-07-14 21:07:20,218 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:07:20,219 [Thread-3413] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:07:20,222 [Thread-3413] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:07:20,222 [Thread-3413] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:07:20,222 [Thread-3413] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:20,222 [Thread-3413] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:07:20,222 [Thread-3413] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:07:20,223 [Thread-3413] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:07:20,227 [Thread-3413] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for map tasks
2018-07-14 21:07:20,227 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1888978724_0121_m_000000_0
2018-07-14 21:07:20,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:07:20,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:07:20,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:07:20,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:07:20,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:07:20,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:07:20,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:07:20,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:07:20,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:07:20,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:07:20,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:07:20,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:07:20,278 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:07:20,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:07:20,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:07:20,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:07:20,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:07:20,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:07:20,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:07:20,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1888978724_0121_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:07:20,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
```

```
2018-07-14 21:07:20,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1888978724_0121_m_000000_0' done.
2018-07-14 21:07:20,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1888978724_0121_m_000000_0
2018-07-14 21:07:20,287 [Thread-3413] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:07:20,288 [Thread-3413] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:07:20,288 [pool-342-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1888978724_0121_r_000000_0
2018-07-14 21:07:20,293 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:07:20,293 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@451f3029
2018-07-14 21:07:20,299 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:07:20,302 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1888978724_0121_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:07:20,303 [localfetcher#100] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#100 about to shuffle
output of map attempt_local1888978724_0121_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:07:20,303 [localfetcher#100] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1888978724_0121_m_000000_0
2018-07-14 21:07:20,303 [localfetcher#100] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:07:20,304 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:07:20,304 [pool-342-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:07:20,304 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:07:20,305 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:07:20,305 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:07:20,306 [pool-342-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 21:07:20,306 [pool-342-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:07:20,309 [pool-342-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:07:20,324 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1888978724_0121_r_000000_0 is done. And is in the process of committing
2018-07-14 21:07:20,326 [pool-342-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:07:20,326 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1888978724_0121_r_000000_0 is allowed to commit now
2018-07-14 21:07:20,326 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1888978724_0121_r_000000_0' to
file:/tmp/temp-1902294536/tmp26425204/_temporary/0/task_local1888978724_0121_r_000000
2018-07-14 21:07:20,328 [pool-342-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1888978724_0121_r_000000_0' to
file:/tmp/temp-1902294536/tmp-336027020/_temporary/0/task_local1888978724_0121_r_00000
0
2018-07-14 21:07:20,329 [pool-342-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:07:20,329 [pool-342-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1888978724_0121_r_000000_0' done.
2018-07-14 21:07:20,329 [pool-342-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1888978724_0121_r_000000_0
2018-07-14 21:07:20,329 [Thread-3413] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:07:20,521 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1888978724_0121
2018-07-14 21:07:20,521 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:07:20,521 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:07:20,525 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:07:20,531 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:20,533 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:20,537 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:20,546 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:07:20,551 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:07:20,551 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:07:20,552 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
```

```
2018-07-14 21:07:20,554 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:07:20,555 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:07:20,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:07:20,584 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:07:20,586 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:20,587 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:20,587 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:07:20,593 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:07:20,594 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:07:20,594 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:07:20,594 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:07:20,595 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:07:20,604 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:07:20,604 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:07:20,604 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:07:20,624 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:07:20,639 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local300505448_0122
2018-07-14 21:07:20,710 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:07:20,711 [Thread-3446] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:07:20,716 [Thread-3446] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:07:20,716 [Thread-3446] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:07:20,716 [Thread-3446] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:20,716 [Thread-3446] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:07:20,716 [Thread-3446] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:07:20,717 [Thread-3446] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:07:20,720 [Thread-3446] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:07:20,720 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local300505448_0122_m_000000_0
2018-07-14 21:07:20,726 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:07:20,726 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:07:20,728 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:07:20,729 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:07:20,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:07:20,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:07:20,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:07:20,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:07:20,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:07:20,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:07:20,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:07:20,750 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:07:20,753 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],ans6[66,7],details[-1,-1],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:07:20,753 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:07:20,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:07:20,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:07:20,765 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:07:20,768 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local300505448_0122_m_000001_0
2018-07-14 21:07:20,770 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
```

```
2018-07-14 21:07:20,770 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:07:20,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-336027020/part-r-00000:0+62
2018-07-14 21:07:20,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:07:20,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:07:20,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:07:20,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:07:20,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:07:20,924 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:07:20,924 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:07:20,924 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:07:20,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],ans6[66,7],details[-1,-1],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:07:20,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:07:20,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:07:20,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:07:20,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:07:20,926 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:07:20,927 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:07:20,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local300505448_0122_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:07:20,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:07:20,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local300505448_0122_m_000001_0' done.
2018-07-14 21:07:20,928 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local300505448_0122_m_000001_0
2018-07-14 21:07:20,929 [Thread-3446] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:07:20,930 [Thread-3446] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local300505448_0122
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
```

```
Exception while executing (Name: ans5: Local Rearrange[tuple]{int}(false) - scope-3399
Operator Key: scope-3399): org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Exception while executing (Name: details: New For Each(false,false,false,false)[bag] -
scope-3392 Operator Key: scope-3392):
org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more than one
row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN" then "FOREACH ...
GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:522)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception while
executing (Name: ans5: Local Rearrange[tuple]{int}(false) - scope-3399 Operator Key:
scope-3399): org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception
while executing (Name: details: New For Each(false,false,false,false)[bag] - scope-3392
Operator Key: scope-3392): org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause:
"JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:314)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POLoca
lRearrange.getNextTuple(POLocalRearrange.java:287)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POUnio
n.getNextTuple(POUnion.java:167)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.runPipe
line(PigGenericMapBase.java:280)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:275)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:65)
        at org.apache.hadoop.mapreduce.Mapper.run(Mapper.java:145)
        at org.apache.hadoop.mapred.MapTask.runNewMapper(MapTask.java:784)
        at org.apache.hadoop.mapred.MapTask.run(MapTask.java:341)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$MapTaskRunnable.run(LocalJobRunner.java:24
3)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception while
executing (Name: details: New For Each(false,false,false,false)[bag] - scope-3392 Operator
Key: scope-3392): org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar
has more than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN"
then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:314)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:256)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:305)
        ... 14 more
```

```
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN" then
"FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextLong(POUserFunc.java:410)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:349)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.EqualT
oExpr.getNextBoolean(EqualToExpr.java:83)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POFilt
er.getNextTuple(POFilter.java:143)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:305)
        ... 16 more
2018-07-14 21:07:21,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local300505448_0122
2018-07-14 21:07:21,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,details
2018-07-14 21:07:21,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[63,7],details[8,10],ans6[66,7],details[-1,-1],ans5[63,7] C:
R: ans5[63,7]
2018-07-14 21:07:21,090 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 21:07:21,090 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local300505448_0122 has failed! Stop running all dependent jobs
2018-07-14 21:07:21,090 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:07:21,094 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:21,096 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:21,097 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 21:07:21,098 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:07:20     2018-07-14 21:07:21
HASH_JOIN,GROUP_BY,FILTER

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1888978724_0121        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,MULTI_QUERY,COMBINER
```

```
Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local300505448_0122 ans5,ans6,details       HASH_JOIN       Message: Job failed!
file:/tmp/temp-1902294536/tmp609163429,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Failed to read data from "file:///home/acadgild/employee_details.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp609163429"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1888978724_0121        ->      job_local300505448_0122,
job_local300505448_0122


2018-07-14 21:07:21,102 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:21,104 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:21,105 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:07:21,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 21:07:21,111 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias ans6
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt>
grunt>
grunt>
grunt> ans6 = filter ans5 BY ex_sum.sum==NULL2018-07-14 21:07:26,729 [communication thread]
INFO  org.apache.hadoop.mapred.LocalJo
grunt>
grunt>
grunt>
grunt>
grunt> ans6 = filter ans5 BY ex_sum.sum>0;
2018-07-14 21:07:41,484 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:07:41,484 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:07:41,511 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
grunt> describe ans5;
ans5: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::EmpId: int,ex_sum::sum: long}
grunt>
grunt> ans6 = filter ans5 BY ex_sum.sum is not null;
2018-07-14 21:08:54,023 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
grunt> dump ans6;
2018-07-14 21:08:58,637 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
```

```
2018-07-14 21:08:58,655 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:08:58,656 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:08:58,656 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:08:58,656 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:08:58,659 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:08:58,660 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:08:58,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 21:08:58,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 21:08:58,665 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 21:08:58,666 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:08:58,675 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:08:58,675 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:08:58,675 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:58,676 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:08:58,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:08:58,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:08:58,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:08:58,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:08:58,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:08:58,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 21:08:58,679 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:08:58,679 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:08:58,679 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
```

```
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582738678-0
2018-07-14 21:08:58,691 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:08:58,692 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:58,701 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:08:58,702 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:08:58,702 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:08:58,702 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:08:58,702 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:08:58,722 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:08:58,740 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1973793369_0123
2018-07-14 21:08:58,818 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:08:58,818 [Thread-3467] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:08:58,822 [Thread-3467] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:08:58,822 [Thread-3467] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:08:58,822 [Thread-3467] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:08:58,822 [Thread-3467] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:08:58,822 [Thread-3467] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:08:58,823 [Thread-3467] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:08:58,827 [Thread-3467] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:08:58,828 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1973793369_0123_m_000000_0
2018-07-14 21:08:58,839 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:08:58,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------
```

```
2018-07-14 21:08:58,844 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:08:58,844 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:08:58,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:08:58,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:08:58,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:08:58,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:08:58,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:08:58,906 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:08:58,907 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:08:58,907 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:08:58,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:08:58,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:08:58,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:08:58,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:08:58,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:08:58,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:08:58,918 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:08:58,918 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1973793369_0123_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:08:58,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:08:58,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1973793369_0123_m_000000_0' done.
2018-07-14 21:08:58,919 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1973793369_0123_m_000000_0
2018-07-14 21:08:58,919 [Thread-3467] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:08:58,920 [Thread-3467] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:08:58,920 [pool-347-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1973793369_0123_r_000000_0
2018-07-14 21:08:58,927 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:08:58,927 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
```

```
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@dd6a4d0
2018-07-14 21:08:58,928 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:08:58,928 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1973793369_0123_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:08:58,930 [localfetcher#101] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#101 about to shuffle
output of map attempt_local1973793369_0123_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:08:58,930 [localfetcher#101] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1973793369_0123_m_000000_0
2018-07-14 21:08:58,930 [localfetcher#101] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:08:58,930 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:08:58,931 [pool-347-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:08:58,931 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:08:58,931 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:08:58,931 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:08:58,933 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:08:58,933 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:08:58,933 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:08:58,933 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:08:58,934 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:08:58,934 [pool-347-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:08:58,934 [pool-347-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:08:58,934 [pool-347-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:08:58,937 [pool-347-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:08:58,955 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1973793369_0123_r_000000_0 is done. And is in the process of committing
2018-07-14 21:08:58,957 [pool-347-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:08:58,957 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1973793369_0123_r_000000_0 is allowed to commit now
2018-07-14 21:08:58,958 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
```

```
'attempt_local1973793369_0123_r_000000_0' to
file:/tmp/temp-1902294536/tmp1716613548/_temporary/0/task_local1973793369_0123_r_00000
0
2018-07-14 21:08:58,959 [pool-347-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1973793369_0123_r_000000_0' to
file:/tmp/temp-1902294536/tmp-42774628/_temporary/0/task_local1973793369_0123_r_000000
2018-07-14 21:08:58,959 [pool-347-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:08:58,959 [pool-347-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1973793369_0123_r_000000_0' done.
2018-07-14 21:08:58,959 [pool-347-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1973793369_0123_r_000000_0
2018-07-14 21:08:58,959 [Thread-3467] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:08:59,120 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1973793369_0123
2018-07-14 21:08:59,120 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:08:59,120 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:08:59,124 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:08:59,128 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,130 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,132 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,139 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:08:59,139 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:08:59,141 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:08:59,141 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:08:59,148 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:08:59,148 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:08:59,152 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:08:59,196 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:08:59,199 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,202 [JobControl] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:08:59,202 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:08:59,210 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:08:59,212 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:08:59,213 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:08:59,213 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:08:59,213 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:08:59,221 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:08:59,221 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:08:59,221 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:08:59,240 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:08:59,254 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local577049818_0124
2018-07-14 21:08:59,314 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:08:59,314 [Thread-3500] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:08:59,319 [Thread-3500] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:08:59,319 [Thread-3500] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:08:59,319 [Thread-3500] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:08:59,319 [Thread-3500] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:08:59,319 [Thread-3500] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:08:59,319 [Thread-3500] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:08:59,321 [Thread-3500] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:08:59,321 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local577049818_0124_m_000000_0
2018-07-14 21:08:59,325 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:08:59,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
```

```
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:08:59,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:08:59,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:08:59,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:08:59,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:08:59,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:08:59,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:08:59,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:08:59,343 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:08:59,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:08:59,346 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:08:59,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],ans6[68,7],details[-1,-1],ans5[63,7],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:08:59,350 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:08:59,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:08:59,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:08:59,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:08:59,361 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local577049818_0124_m_000001_0
2018-07-14 21:08:59,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:08:59,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:08:59,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-42774628/part-r-00000:0+62
2018-07-14 21:08:59,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:08:59,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:08:59,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:08:59,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:08:59,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:08:59,480 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:08:59,481 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:08:59,481 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:08:59,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],ans6[68,7],details[-1,-1],ans5[63,7],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:08:59,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:08:59,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:08:59,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:08:59,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:08:59,484 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:08:59,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:08:59,486 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local577049818_0124_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:08:59,486 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:08:59,486 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local577049818_0124_m_000001_0' done.
2018-07-14 21:08:59,486 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local577049818_0124_m_000001_0
2018-07-14 21:08:59,486 [Thread-3500] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:08:59,487 [Thread-3500] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local577049818_0124
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Exception while executing (Name: ans5: Local Rearrange[tuple]{int}(false) - scope-3489
Operator Key: scope-3489): org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Exception while executing (Name: details: New For Each(false,false,false,false)[bag] -
scope-3482 Operator Key: scope-3482):
org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more than one
row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN" then "FOREACH ...
GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:522)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception while
executing (Name: ans5: Local Rearrange[tuple]{int}(false) - scope-3489 Operator Key:
scope-3489): org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception
while executing (Name: details: New For Each(false,false,false,false)[bag] - scope-3482
Operator Key: scope-3482): org.apache.pig.backend.executionengine.ExecException: ERROR 0:
```

```
Scalar has more than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause:
"JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:314)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POLoca
lRearrange.getNextTuple(POLocalRearrange.java:287)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POUnio
n.getNextTuple(POUnion.java:167)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.runPipe
line(PigGenericMapBase.java:280)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:275)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:65)
        at org.apache.hadoop.mapreduce.Mapper.run(Mapper.java:145)
        at org.apache.hadoop.mapred.MapTask.runNewMapper(MapTask.java:784)
        at org.apache.hadoop.mapred.MapTask.run(MapTask.java:341)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$MapTaskRunnable.run(LocalJobRunner.java:24
3)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception while
executing (Name: details: New For Each(false,false,false,false)[bag] - scope-3482 Operator
Key: scope-3482): org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar
has more than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN"
then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:314)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:256)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:305)
        ... 14 more
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN" then
"FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextLong(POUserFunc.java:410)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:349)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POIsNu
```

```
ll.getNextBoolean(POIsNull.java:78)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.PONot.
getNextBoolean(PONot.java:71)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POFilt
er.getNextTuple(POFilter.java:143)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:305)
        ... 16 more
2018-07-14 21:08:59,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local577049818_0124
2018-07-14 21:08:59,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,details
2018-07-14 21:08:59,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],ans6[68,7],details[-1,-1],ans5[63,7],ans5[63,7] C:
R: ans5[63,7]
2018-07-14 21:08:59,681 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 21:08:59,681 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local577049818_0124 has failed! Stop running all dependent jobs
2018-07-14 21:08:59,682 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:08:59,688 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,690 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,690 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 21:08:59,691 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:08:58     2018-07-14 21:08:59
HASH_JOIN,GROUP_BY,FILTER

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1973793369_0123        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,MULTI_QUERY,COMBINER

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local577049818_0124 ans5,ans6,details       HASH_JOIN       Message: Job failed!
file:/tmp/temp-1902294536/tmp-2109823017,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Failed to read data from "file:///home/acadgild/employee_details.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp-2109823017"
```

```
Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1973793369_0123        ->      job_local577049818_0124,
job_local577049818_0124


2018-07-14 21:08:59,693 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,693 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,695 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:08:59,696 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 21:08:59,697 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias ans6
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> 2018-07-14 21:09:05,323 [communication thread] INFO
org.apache.hadoop.mapred.LocalJobRunner - map > sort

grunt>
grunt>
grunt> dump ans5;
2018-07-14 21:09:23,078 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:09:23,078 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:09:23,088 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:09:23,101 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:09:23,101 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:09:23,101 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:09:23,101 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:09:23,104 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:09:23,106 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:09:23,108 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:09:23,108 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:09:23,115 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 21:09:23,115 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:09:23,121 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,122 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:09:23,122 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:09:23,122 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:09:23,123 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:09:23,124 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:09:23,124 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:09:23,125 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:09:23,125 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:09:23,125 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:09:23,125 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582763125-0
2018-07-14 21:09:23,136 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:09:23,137 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,145 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:09:23,147 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:09:23,147 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:09:23,147 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:09:23,147 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:09:23,169 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:09:23,182 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local302694688_0125
2018-07-14 21:09:23,244 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:09:23,245 [Thread-3521] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:09:23,249 [Thread-3521] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:09:23,249 [Thread-3521] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:09:23,249 [Thread-3521] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:09:23,249 [Thread-3521] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:09:23,249 [Thread-3521] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:09:23,249 [Thread-3521] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:09:23,253 [Thread-3521] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:09:23,253 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local302694688_0125_m_000000_0
2018-07-14 21:09:23,258 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:09:23,259 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:09:23,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:09:23,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:09:23,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:09:23,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:09:23,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:09:23,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:09:23,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:09:23,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:09:23,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:09:23,277 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:09:23,279 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:09:23,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
```

```
2018-07-14 21:09:23,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:09:23,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:09:23,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:09:23,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:09:23,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:09:23,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local302694688_0125_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:09:23,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:09:23,288 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local302694688_0125_m_000000_0' done.
2018-07-14 21:09:23,288 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local302694688_0125_m_000000_0
2018-07-14 21:09:23,288 [Thread-3521] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:09:23,288 [Thread-3521] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:09:23,289 [pool-352-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local302694688_0125_r_000000_0
2018-07-14 21:09:23,295 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:09:23,296 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@693855bb
2018-07-14 21:09:23,296 [pool-352-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:09:23,297 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local302694688_0125_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:09:23,298 [localfetcher#102] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#102 about to shuffle
output of map attempt_local302694688_0125_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:09:23,298 [localfetcher#102] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local302694688_0125_m_000000_0
2018-07-14 21:09:23,298 [localfetcher#102] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:09:23,300 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:09:23,301 [pool-352-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:09:23,301 [pool-352-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:09:23,302 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:09:23,302 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:09:23,302 [pool-352-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
```

```
to disk to satisfy reduce memory limit
2018-07-14 21:09:23,303 [pool-352-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:09:23,303 [pool-352-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:09:23,303 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:09:23,303 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:09:23,303 [pool-352-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:09:23,310 [pool-352-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:09:23,310 [pool-352-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:09:23,312 [pool-352-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:09:23,313 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local302694688_0125_r_000000_0 is done. And is in the process of committing
2018-07-14 21:09:23,315 [pool-352-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:09:23,315 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local302694688_0125_r_000000_0 is allowed to commit now
2018-07-14 21:09:23,316 [pool-352-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local302694688_0125_r_000000_0' to
file:/tmp/temp-1902294536/tmp436256989/_temporary/0/task_local302694688_0125_r_000000
2018-07-14 21:09:23,317 [pool-352-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:09:23,317 [pool-352-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local302694688_0125_r_000000_0' done.
2018-07-14 21:09:23,317 [pool-352-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local302694688_0125_r_000000_0
2018-07-14 21:09:23,317 [Thread-3521] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:09:23,445 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local302694688_0125
2018-07-14 21:09:23,445 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:09:23,445 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:09:23,446 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:09:23,447 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,447 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,447 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,449 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
```

```
the job
2018-07-14 21:09:23,449 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:09:23,449 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:09:23,449 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:09:23,450 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:09:23,450 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:09:23,451 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:09:23,460 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:09:23,461 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,463 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:09:23,463 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:09:23,469 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:09:23,470 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:09:23,470 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:09:23,470 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:09:23,470 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:09:23,477 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:09:23,477 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:09:23,477 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:09:23,519 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:09:23,533 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1492476764_0126
2018-07-14 21:09:23,599 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:09:23,600 [Thread-3550] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:09:23,604 [Thread-3550] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
```

```
2018-07-14 21:09:23,604 [Thread-3550] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:09:23,604 [Thread-3550] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:09:23,604 [Thread-3550] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:09:23,604 [Thread-3550] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:09:23,605 [Thread-3550] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:09:23,607 [Thread-3550] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:09:23,607 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1492476764_0126_m_000000_0
2018-07-14 21:09:23,612 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:09:23,612 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:09:23,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:09:23,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:09:23,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:09:23,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:09:23,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:09:23,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:09:23,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:09:23,630 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:09:23,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:09:23,631 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:09:23,633 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans5[63,7],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:09:23,633 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:09:23,633 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:09:23,633 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:09:23,633 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:09:23,633 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:09:23,635 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:09:23,636 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1492476764_0126_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:09:23,637 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:09:23,637 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1492476764_0126_m_000000_0' done.
2018-07-14 21:09:23,637 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1492476764_0126_m_000000_0
2018-07-14 21:09:23,637 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1492476764_0126_m_000001_0
2018-07-14 21:09:23,639 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:09:23,641 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:09:23,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp436256989/part-r-00000:0+62
2018-07-14 21:09:23,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:09:23,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:09:23,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:09:23,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:09:23,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:09:23,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:09:23,659 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:09:23,659 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:09:23,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans5[63,7],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:09:23,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:09:23,661 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:09:23,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:09:23,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:09:23,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:09:23,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:09:23,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1492476764_0126_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:09:23,663 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:09:23,663 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1492476764_0126_m_000001_0' done.
2018-07-14 21:09:23,663 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1492476764_0126_m_000001_0
2018-07-14 21:09:23,663 [Thread-3550] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:09:23,664 [Thread-3550] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:09:23,664 [pool-355-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1492476764_0126_r_000000_0
2018-07-14 21:09:23,669 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:09:23,669 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@a8c888b
2018-07-14 21:09:23,670 [pool-355-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:09:23,674 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1492476764_0126_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:09:23,675 [localfetcher#103] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#103 about to shuffle
output of map attempt_local1492476764_0126_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:09:23,675 [localfetcher#103] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1492476764_0126_m_000000_0
2018-07-14 21:09:23,675 [localfetcher#103] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 21:09:23,676 [localfetcher#103] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#103 about to shuffle
output of map attempt_local1492476764_0126_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:09:23,676 [localfetcher#103] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1492476764_0126_m_000001_0
2018-07-14 21:09:23,676 [localfetcher#103] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 21:09:23,678 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:09:23,678 [pool-355-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
```

```
2018-07-14 21:09:23,679 [pool-355-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:09:23,680 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:09:23,680 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:09:23,682 [pool-355-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:09:23,682 [pool-355-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:09:23,682 [pool-355-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:09:23,682 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:09:23,683 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:09:23,683 [pool-355-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:09:23,689 [pool-355-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:09:23,689 [pool-355-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:09:23,692 [pool-355-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],details[-1,-1],ans5[63,7],ans5[63,7] C:  R: ans5[63,7]
2018-07-14 21:09:23,693 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1492476764_0126_r_000000_0 is done. And is in the process of committing
2018-07-14 21:09:23,694 [pool-355-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:09:23,694 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1492476764_0126_r_000000_0 is allowed to commit now
2018-07-14 21:09:23,695 [pool-355-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1492476764_0126_r_000000_0' to
file:/tmp/temp-1902294536/tmp-29779757/_temporary/0/task_local1492476764_0126_r_000000
2018-07-14 21:09:23,695 [pool-355-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:09:23,695 [pool-355-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1492476764_0126_r_000000_0' done.
2018-07-14 21:09:23,695 [pool-355-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1492476764_0126_r_000000_0
2018-07-14 21:09:23,695 [Thread-3550] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:09:23,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1492476764_0126
2018-07-14 21:09:23,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 21:09:23,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],details[-1,-1],ans5[63,7],ans5[63,7] C:  R:
ans5[63,7]
2018-07-14 21:09:23,908 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,909 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,910 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:09:23,917 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:09:23     2018-07-14 21:09:23
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1492476764_0126        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp-1902294536/tmp-29779757,
job_local302694688_0125 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp-29779757"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local302694688_0125 ->      job_local1492476764_0126,
job_local1492476764_0126


2018-07-14 21:09:23,919 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,920 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,922 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,926 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,927 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,928 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:09:23,930 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:09:23,933 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:09:23,934 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2018-07-14 21:09:23,934 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:09:23,943 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:09:23,943 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(103,Akshay,11000,3,,)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(106,Aamir,25000,1,,)
(107,Salman,17500,2,,)
(108,Ranbir,14000,3,,)
(109,Katrina,1000,4,,)
(110,Priyanka,2000,5,110,400)
(111,Tushar,500,1,,)
(112,Ajay,5000,2,,)
(113,Jubeen,1000,1,,)
(114,Madhuri,2000,2,114,200)
grunt> ans6 = foreach ans5 generate ..$3,$5;
2018-07-14 21:11:10,887 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
grunt> dump ans6;
2018-07-14 21:11:20,595 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:11:20,611 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:11:20,612 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:11:20,612 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:11:20,612 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:11:20,614 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:11:20,615 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:11:20,616 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:11:20,616 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:11:20,633 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:11:20,633 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:11:20,633 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:20,635 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:11:20,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
```

```
2018-07-14 21:11:20,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:11:20,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:11:20,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:11:20,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:11:20,636 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:11:20,637 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:11:20,637 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:11:20,637 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582880636-0
2018-07-14 21:11:20,646 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:11:20,647 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:20,662 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:11:20,663 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:11:20,663 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:11:20,663 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:11:20,664 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:11:20,683 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:11:20,698 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local991984607_0127
2018-07-14 21:11:20,767 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:11:20,768 [Thread-3580] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:11:20,772 [Thread-3580] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:11:20,772 [Thread-3580] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:11:20,772 [Thread-3580] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:11:20,772 [Thread-3580] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:11:20,772 [Thread-3580] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:11:20,772 [Thread-3580] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:11:20,775 [Thread-3580] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:11:20,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local991984607_0127_m_000000_0
2018-07-14 21:11:20,779 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:11:20,780 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:11:20,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:11:20,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:11:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:11:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:11:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:11:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:11:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:11:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:11:20,799 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:11:20,800 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:11:20,801 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:11:20,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:11:20,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:11:20,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:11:20,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:11:20,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:11:20,812 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:11:20,814 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local991984607_0127_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:11:20,814 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:11:20,814 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local991984607_0127_m_000000_0' done.
2018-07-14 21:11:20,814 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local991984607_0127_m_000000_0
2018-07-14 21:11:20,814 [Thread-3580] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:11:20,815 [Thread-3580] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:11:20,815 [pool-358-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local991984607_0127_r_000000_0
2018-07-14 21:11:20,820 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:11:20,820 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@a4a3ea
2018-07-14 21:11:20,822 [pool-358-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:11:20,823 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local991984607_0127_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:11:20,824 [localfetcher#104] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#104 about to shuffle
output of map attempt_local991984607_0127_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:11:20,825 [localfetcher#104] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local991984607_0127_m_000000_0
2018-07-14 21:11:20,826 [localfetcher#104] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:11:20,826 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:11:20,827 [pool-358-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:11:20,827 [pool-358-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:11:20,828 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:11:20,828 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:11:20,829 [pool-358-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:11:20,829 [pool-358-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:11:20,829 [pool-358-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:11:20,829 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:11:20,829 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
```

```
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:11:20,829 [pool-358-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:11:20,835 [pool-358-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:11:20,835 [pool-358-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:11:20,837 [pool-358-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:11:20,838 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local991984607_0127_r_000000_0 is done. And is in the process of committing
2018-07-14 21:11:20,839 [pool-358-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:11:20,839 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local991984607_0127_r_000000_0 is allowed to commit now
2018-07-14 21:11:20,840 [pool-358-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local991984607_0127_r_000000_0' to
file:/tmp/temp-1902294536/tmp-196044910/_temporary/0/task_local991984607_0127_r_000000
2018-07-14 21:11:20,842 [pool-358-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:11:20,842 [pool-358-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local991984607_0127_r_000000_0' done.
2018-07-14 21:11:20,842 [pool-358-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local991984607_0127_r_000000_0
2018-07-14 21:11:20,842 [Thread-3580] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:11:20,970 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local991984607_0127
2018-07-14 21:11:20,970 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:11:20,970 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:11:20,971 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:11:20,971 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:20,972 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:20,972 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:20,974 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:11:20,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:11:20,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:11:20,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:11:20,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:11:20,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:11:20,977 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:11:20,987 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:11:20,988 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:20,995 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:11:20,995 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:11:21,002 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:11:21,011 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:11:21,011 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:11:21,011 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:11:21,013 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:11:21,013 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:11:21,013 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:11:21,014 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:11:21,037 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:11:21,058 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local669564054_0128
2018-07-14 21:11:21,152 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:11:21,153 [Thread-3609] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:11:21,158 [Thread-3609] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:11:21,158 [Thread-3609] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:11:21,159 [Thread-3609] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:11:21,159 [Thread-3609] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
```

```
2018-07-14 21:11:21,159 [Thread-3609] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:11:21,159 [Thread-3609] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:11:21,164 [Thread-3609] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:11:21,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local669564054_0128_m_000000_0
2018-07-14 21:11:21,170 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:11:21,171 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:11:21,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:11:21,174 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:11:21,200 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:11:21,200 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:11:21,200 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:11:21,200 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:11:21,200 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:11:21,200 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:11:21,203 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:11:21,203 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:11:21,206 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7],ans6[69,7]
2018-07-14 21:11:21,211 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:11:21,211 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:11:21,211 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:11:21,211 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:11:21,211 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:11:21,216 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
```

```
2018-07-14 21:11:21,217 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local669564054_0128_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:11:21,218 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:11:21,218 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local669564054_0128_m_000000_0' done.
2018-07-14 21:11:21,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local669564054_0128_m_000000_0
2018-07-14 21:11:21,219 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local669564054_0128_m_000001_0
2018-07-14 21:11:21,220 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:11:21,221 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:11:21,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-196044910/part-r-00000:0+62
2018-07-14 21:11:21,237 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:11:21,237 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:11:21,237 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:11:21,237 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:11:21,237 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:11:21,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:11:21,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:11:21,239 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:11:21,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7],ans6[69,7]
2018-07-14 21:11:21,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:11:21,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:11:21,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:11:21,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:11:21,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:11:21,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
```

```
2018-07-14 21:11:21,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local669564054_0128_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:11:21,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:11:21,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local669564054_0128_m_000001_0' done.
2018-07-14 21:11:21,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local669564054_0128_m_000001_0
2018-07-14 21:11:21,247 [Thread-3609] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:11:21,248 [Thread-3609] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:11:21,248 [pool-361-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local669564054_0128_r_000000_0
2018-07-14 21:11:21,253 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:11:21,253 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7d328e88
2018-07-14 21:11:21,254 [pool-361-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:11:21,258 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local669564054_0128_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:11:21,260 [localfetcher#105] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#105 about to shuffle
output of map attempt_local669564054_0128_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:11:21,260 [localfetcher#105] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local669564054_0128_m_000001_0
2018-07-14 21:11:21,260 [localfetcher#105] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 21:11:21,260 [localfetcher#105] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#105 about to shuffle
output of map attempt_local669564054_0128_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:11:21,261 [localfetcher#105] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local669564054_0128_m_000000_0
2018-07-14 21:11:21,261 [localfetcher#105] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 21:11:21,262 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:11:21,262 [pool-361-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:11:21,262 [pool-361-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:11:21,262 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:11:21,262 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:11:21,264 [pool-361-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
```

```
2018-07-14 21:11:21,264 [pool-361-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:11:21,264 [pool-361-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:11:21,264 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:11:21,264 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:11:21,264 [pool-361-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:11:21,270 [pool-361-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:11:21,270 [pool-361-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:11:21,272 [pool-361-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7],ans6[69,7]
2018-07-14 21:11:21,274 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local669564054_0128_r_000000_0 is done. And is in the process of committing
2018-07-14 21:11:21,276 [pool-361-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:11:21,276 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local669564054_0128_r_000000_0 is allowed to commit now
2018-07-14 21:11:21,278 [pool-361-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local669564054_0128_r_000000_0' to
file:/tmp/temp-1902294536/tmp750940741/_temporary/0/task_local669564054_0128_r_000000
2018-07-14 21:11:21,278 [pool-361-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:11:21,278 [pool-361-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local669564054_0128_r_000000_0' done.
2018-07-14 21:11:21,278 [pool-361-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local669564054_0128_r_000000_0
2018-07-14 21:11:21,278 [Thread-3609] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:11:21,455 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local669564054_0128
2018-07-14 21:11:21,456 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,details
2018-07-14 21:11:21,456 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7],ans6[69,7]
2018-07-14 21:11:21,468 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,471 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,474 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,488 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:11:21,489 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
```

```
2.6.5   0.16.0  acadgild        2018-07-14 21:11:20     2018-07-14 21:11:21
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local669564054_0128 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans5,ans6,details       HASH_JOIN
file:/tmp/temp-1902294536/tmp750940741,
job_local991984607_0127 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp750940741"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local991984607_0127 ->      job_local669564054_0128,
job_local669564054_0128


2018-07-14 21:11:21,490 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,491 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,492 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,494 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,494 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,495 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:11:21,499 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:11:21,499 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:11:21,499 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:11:21,499 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:11:21,509 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:11:21,509 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,300)
(102,Shahrukh,10000,2,500)
(103,Akshay,11000,3,)
```

```
(104,Anubhav,5000,4,300)
(105,Pawan,2500,5,100)
(106,Aamir,25000,1,)
(107,Salman,17500,2,)
(108,Ranbir,14000,3,)
(109,Katrina,1000,4,)
(110,Priyanka,2000,5,400)
(111,Tushar,500,1,)
(112,Ajay,5000,2,)
(113,Jubeen,1000,1,)
(114,Madhuri,2000,2,200)
grunt> describe ans6;
ans6: {details::EmpId: int,details::Name: chararray,details::Salary: int,details::Dept:
int,ex_sum::sum: long}
grunt> ans7 = filter ans6 BY sum is null;
2018-07-14 21:12:18,242 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
grunt> dump ans7;
2018-07-14 21:12:23,013 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:12:23,031 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:12:23,032 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:12:23,032 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:12:23,032 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:12:23,034 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:12:23,039 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:12:23,040 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:12:23,040 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:12:23,056 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:12:23,056 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:12:23,057 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
```

```
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:12:23,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:12:23,059 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:12:23,059 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:12:23,059 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531582943059-0
2018-07-14 21:12:23,071 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:12:23,072 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,080 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:12:23,082 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:12:23,082 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:12:23,082 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:12:23,082 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:12:23,102 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:12:23,118 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local8978727_0129
2018-07-14 21:12:23,213 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:12:23,214 [Thread-3639] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:12:23,220 [Thread-3639] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:12:23,220 [Thread-3639] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:12:23,220 [Thread-3639] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:12:23,221 [Thread-3639] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:12:23,221 [Thread-3639] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:12:23,221 [Thread-3639] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:12:23,223 [Thread-3639] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
```

```
2018-07-14 21:12:23,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local8978727_0129_m_000000_0
2018-07-14 21:12:23,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:12:23,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:12:23,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:12:23,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:12:23,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:12:23,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:12:23,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:12:23,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:12:23,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:12:23,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:12:23,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:12:23,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:12:23,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:12:23,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:12:23,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:12:23,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:12:23,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:12:23,252 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:12:23,261 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:12:23,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local8978727_0129_m_000000_0 is done. And is
in the process of committing
2018-07-14 21:12:23,263 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:12:23,263 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task 'attempt_local8978727_0129_m_000000_0' done.
2018-07-14 21:12:23,263 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local8978727_0129_m_000000_0
2018-07-14 21:12:23,263 [Thread-3639] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:12:23,264 [Thread-3639] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:12:23,264 [pool-364-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local8978727_0129_r_000000_0
2018-07-14 21:12:23,272 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:12:23,272 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2f8052be
2018-07-14 21:12:23,273 [pool-364-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:12:23,275 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local8978727_0129_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:12:23,276 [localfetcher#106] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#106 about to shuffle
output of map attempt_local8978727_0129_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:12:23,276 [localfetcher#106] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local8978727_0129_m_000000_0
2018-07-14 21:12:23,276 [localfetcher#106] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:12:23,277 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:12:23,278 [pool-364-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:12:23,278 [pool-364-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:12:23,278 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:12:23,278 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:12:23,280 [pool-364-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:12:23,280 [pool-364-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:12:23,280 [pool-364-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:12:23,280 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:12:23,280 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:12:23,280 [pool-364-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:12:23,290 [pool-364-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:12:23,290 [pool-364-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
```

```
- SchemaTupleBackend has already been initialized
2018-07-14 21:12:23,291 [pool-364-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:12:23,292 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local8978727_0129_r_000000_0 is done. And is in the process of committing
2018-07-14 21:12:23,293 [pool-364-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:12:23,293 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local8978727_0129_r_000000_0 is allowed to commit now
2018-07-14 21:12:23,294 [pool-364-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local8978727_0129_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1158071767/_temporary/0/task_local8978727_0129_r_000000
2018-07-14 21:12:23,294 [pool-364-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:12:23,294 [pool-364-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local8978727_0129_r_000000_0' done.
2018-07-14 21:12:23,294 [pool-364-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local8978727_0129_r_000000_0
2018-07-14 21:12:23,294 [Thread-3639] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:12:23,414 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local8978727_0129
2018-07-14 21:12:23,414 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:12:23,414 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:12:23,415 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:12:23,415 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,416 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,416 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,417 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:12:23,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:12:23,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:12:23,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:12:23,418 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:12:23,418 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
```

```
2018-07-14 21:12:23,419 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:12:23,427 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:12:23,429 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,430 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:12:23,430 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:12:23,435 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:12:23,443 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:12:23,443 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:12:23,443 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:12:23,444 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:12:23,445 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:12:23,445 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:12:23,445 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:12:23,465 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:12:23,480 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local590882194_0130
2018-07-14 21:12:23,541 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:12:23,542 [Thread-3668] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:12:23,545 [Thread-3668] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:12:23,545 [Thread-3668] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:12:23,545 [Thread-3668] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:12:23,545 [Thread-3668] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:12:23,545 [Thread-3668] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:12:23,546 [Thread-3668] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:12:23,547 [Thread-3668] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for map tasks
2018-07-14 21:12:23,547 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local590882194_0130_m_000000_0
2018-07-14 21:12:23,552 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:12:23,552 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:12:23,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:12:23,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:12:23,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:12:23,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:12:23,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:12:23,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:12:23,570 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:12:23,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:12:23,573 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:12:23,573 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:12:23,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7],ans7[70,7],ans6[69,7]
2018-07-14 21:12:23,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:12:23,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:12:23,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:12:23,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:12:23,574 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:12:23,575 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:12:23,576 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local590882194_0130_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:12:23,577 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:12:23,577 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task 'attempt_local590882194_0130_m_000000_0' done.
2018-07-14 21:12:23,577 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local590882194_0130_m_000000_0
2018-07-14 21:12:23,577 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local590882194_0130_m_000001_0
2018-07-14 21:12:23,579 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:12:23,579 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:12:23,580 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp-1158071767/part-r-00000:0+62
2018-07-14 21:12:23,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:12:23,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:12:23,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:12:23,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:12:23,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:12:23,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:12:23,713 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:12:23,713 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:12:23,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7],ans7[70,7],ans6[69,7]
2018-07-14 21:12:23,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:12:23,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:12:23,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:12:23,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:12:23,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:12:23,716 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:12:23,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local590882194_0130_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:12:23,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
```

```
2018-07-14 21:12:23,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local590882194_0130_m_000001_0' done.
2018-07-14 21:12:23,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local590882194_0130_m_000001_0
2018-07-14 21:12:23,718 [Thread-3668] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:12:23,718 [Thread-3668] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:12:23,718 [pool-367-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local590882194_0130_r_000000_0
2018-07-14 21:12:23,725 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:12:23,725 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@d77a183
2018-07-14 21:12:23,726 [pool-367-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:12:23,730 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local590882194_0130_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:12:23,732 [localfetcher#107] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#107 about to shuffle
output of map attempt_local590882194_0130_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:12:23,732 [localfetcher#107] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local590882194_0130_m_000000_0
2018-07-14 21:12:23,732 [localfetcher#107] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 21:12:23,734 [localfetcher#107] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#107 about to shuffle
output of map attempt_local590882194_0130_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:12:23,734 [localfetcher#107] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local590882194_0130_m_000001_0
2018-07-14 21:12:23,734 [localfetcher#107] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 21:12:23,735 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:12:23,735 [pool-367-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:12:23,735 [pool-367-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:12:23,736 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:12:23,736 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:12:23,737 [pool-367-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:12:23,737 [pool-367-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:12:23,737 [pool-367-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
```

```
from memory into reduce
2018-07-14 21:12:23,737 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:12:23,738 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:12:23,738 [pool-367-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:12:23,745 [pool-367-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:12:23,745 [pool-367-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:12:23,748 [pool-367-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7],ans7[70,7],ans6[69,7]
2018-07-14 21:12:23,750 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local590882194_0130_r_000000_0 is done. And is in the process of committing
2018-07-14 21:12:23,751 [pool-367-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:12:23,751 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local590882194_0130_r_000000_0 is allowed to commit now
2018-07-14 21:12:23,752 [pool-367-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local590882194_0130_r_000000_0' to
file:/tmp/temp-1902294536/tmp-1155886549/_temporary/0/task_local590882194_0130_r_00000
0
2018-07-14 21:12:23,752 [pool-367-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:12:23,752 [pool-367-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local590882194_0130_r_000000_0' done.
2018-07-14 21:12:23,752 [pool-367-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local590882194_0130_r_000000_0
2018-07-14 21:12:23,752 [Thread-3668] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:12:23,898 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local590882194_0130
2018-07-14 21:12:23,898 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,ans7,details
2018-07-14 21:12:23,898 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7],ans7[70,7],ans6[69,7]
2018-07-14 21:12:23,902 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,907 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,908 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:12:23,916 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:12:23     2018-07-14 21:12:23
HASH_JOIN,GROUP_BY,FILTER
```

```
Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local590882194_0130 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans5,ans6,ans7,details  HASH_JOIN
file:/tmp/temp-1902294536/tmp-1155886549,
job_local8978727_0129   1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 8 records in: "file:/tmp/temp-1902294536/tmp-1155886549"

Counters:
Total records written : 8
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local8978727_0129   ->      job_local590882194_0130,
job_local590882194_0130


2018-07-14 21:12:23,918 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,919 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,920 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,922 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,923 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,923 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:12:23,925 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:12:23,925 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:12:23,926 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:12:23,926 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:12:23,937 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:12:23,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(103,Akshay,11000,3,)
(106,Aamir,25000,1,)
(107,Salman,17500,2,)
(108,Ranbir,14000,3,)
(109,Katrina,1000,4,)
(111,Tushar,500,1,)
```

```
(112,Ajay,5000,2,)
(113,Jubeen,1000,1,)
grunt> ans6 = filter ans5 BY ex_sum.sum is not null;
2018-07-14 21:13:45,586 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
grunt> ans6 = foreach ans5 generate ..$3,$5;
2018-07-14 21:13:46,359 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
grunt>
grunt>
grunt> dump ans6;
2018-07-14 21:13:50,112 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:13:50,128 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:13:50,129 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:13:50,129 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:13:50,129 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:13:50,131 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:13:50,132 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:13:50,132 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:13:50,132 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:13:50,143 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:13:50,143 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:13:50,143 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,145 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:13:50,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:13:50,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:13:50,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:13:50,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:13:50,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:13:50,146 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:13:50,147 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:13:50,147 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:13:50,147 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583030147-0
2018-07-14 21:13:50,155 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:13:50,158 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,166 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:13:50,167 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:13:50,168 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:13:50,168 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:13:50,168 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:13:50,188 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:13:50,202 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local327896914_0131
2018-07-14 21:13:50,262 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:13:50,263 [Thread-3698] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:13:50,267 [Thread-3698] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:13:50,267 [Thread-3698] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:13:50,267 [Thread-3698] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:13:50,267 [Thread-3698] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:13:50,267 [Thread-3698] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:13:50,267 [Thread-3698] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:13:50,272 [Thread-3698] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:13:50,272 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local327896914_0131_m_000000_0
2018-07-14 21:13:50,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:13:50,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
```

```
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:13:50,278 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:13:50,278 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:13:50,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:13:50,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:13:50,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:13:50,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:13:50,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:13:50,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:13:50,318 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:13:50,318 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:13:50,319 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:13:50,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:13:50,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:13:50,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:13:50,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:13:50,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:13:50,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:13:50,328 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local327896914_0131_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:13:50,328 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:13:50,328 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local327896914_0131_m_000000_0' done.
2018-07-14 21:13:50,328 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local327896914_0131_m_000000_0
2018-07-14 21:13:50,328 [Thread-3698] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:13:50,329 [Thread-3698] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for reduce tasks
2018-07-14 21:13:50,329 [pool-370-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local327896914_0131_r_000000_0
2018-07-14 21:13:50,333 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:13:50,333 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@976f4a7
2018-07-14 21:13:50,334 [pool-370-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:13:50,335 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local327896914_0131_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:13:50,337 [localfetcher#108] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#108 about to shuffle
output of map attempt_local327896914_0131_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:13:50,337 [localfetcher#108] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local327896914_0131_m_000000_0
2018-07-14 21:13:50,337 [localfetcher#108] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:13:50,338 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:13:50,339 [pool-370-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:13:50,339 [pool-370-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:13:50,340 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:13:50,340 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:13:50,341 [pool-370-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:13:50,341 [pool-370-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:13:50,341 [pool-370-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:13:50,341 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:13:50,341 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:13:50,342 [pool-370-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:13:50,347 [pool-370-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:13:50,347 [pool-370-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:13:50,348 [pool-370-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:13:50,349 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Task -
```

```
Task:attempt_local327896914_0131_r_000000_0 is done. And is in the process of committing
2018-07-14 21:13:50,350 [pool-370-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:13:50,350 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local327896914_0131_r_000000_0 is allowed to commit now
2018-07-14 21:13:50,351 [pool-370-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local327896914_0131_r_000000_0' to
file:/tmp/temp-1902294536/tmp1735483955/_temporary/0/task_local327896914_0131_r_000000
2018-07-14 21:13:50,351 [pool-370-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:13:50,351 [pool-370-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local327896914_0131_r_000000_0' done.
2018-07-14 21:13:50,351 [pool-370-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local327896914_0131_r_000000_0
2018-07-14 21:13:50,351 [Thread-3698] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:13:50,464 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local327896914_0131
2018-07-14 21:13:50,464 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:13:50,464 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:13:50,465 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:13:50,466 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,467 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,467 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,470 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:13:50,471 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:13:50,471 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:13:50,471 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:13:50,472 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:13:50,472 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:13:50,473 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:13:50,486 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:13:50,489 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
```

```
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,490 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:13:50,490 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:13:50,496 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:13:50,497 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:13:50,497 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:13:50,497 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:13:50,497 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:13:50,504 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:13:50,504 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:13:50,504 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:13:50,526 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:13:50,539 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local318240264_0132
2018-07-14 21:13:50,609 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:13:50,610 [Thread-3727] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:13:50,613 [Thread-3727] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:13:50,613 [Thread-3727] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:13:50,613 [Thread-3727] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:13:50,613 [Thread-3727] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:13:50,613 [Thread-3727] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:13:50,613 [Thread-3727] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:13:50,615 [Thread-3727] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:13:50,615 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local318240264_0132_m_000000_0
2018-07-14 21:13:50,620 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:13:50,620 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:13:50,622 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:13:50,622 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:13:50,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:13:50,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:13:50,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:13:50,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:13:50,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:13:50,640 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:13:50,642 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:13:50,642 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:13:50,646 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7],ans6[72,7]
2018-07-14 21:13:50,647 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:13:50,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:13:50,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:13:50,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:13:50,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:13:50,648 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:13:50,650 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local318240264_0132_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:13:50,653 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:13:50,653 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local318240264_0132_m_000000_0' done.
2018-07-14 21:13:50,653 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local318240264_0132_m_000000_0
2018-07-14 21:13:50,653 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local318240264_0132_m_000001_0
2018-07-14 21:13:50,654 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:13:50,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:13:50,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1902294536/tmp1735483955/part-r-00000:0+62
2018-07-14 21:13:50,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:13:50,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:13:50,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:13:50,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:13:50,816 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:13:50,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:13:50,817 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:13:50,817 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:13:50,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7],ans6[72,7]
2018-07-14 21:13:50,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:13:50,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:13:50,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:13:50,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:13:50,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:13:50,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:13:50,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local318240264_0132_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:13:50,821 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:13:50,821 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local318240264_0132_m_000001_0' done.
2018-07-14 21:13:50,821 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local318240264_0132_m_000001_0
2018-07-14 21:13:50,821 [Thread-3727] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:13:50,821 [Thread-3727] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
```

```
2018-07-14 21:13:50,822 [pool-373-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local318240264_0132_r_000000_0
2018-07-14 21:13:50,827 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:13:50,827 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4b03b82b
2018-07-14 21:13:50,828 [pool-373-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:13:50,828 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local318240264_0132_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:13:50,829 [localfetcher#109] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#109 about to shuffle
output of map attempt_local318240264_0132_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:13:50,829 [localfetcher#109] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local318240264_0132_m_000000_0
2018-07-14 21:13:50,829 [localfetcher#109] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 21:13:50,835 [localfetcher#109] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#109 about to shuffle
output of map attempt_local318240264_0132_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:13:50,835 [localfetcher#109] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local318240264_0132_m_000001_0
2018-07-14 21:13:50,835 [localfetcher#109] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 21:13:50,837 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:13:50,838 [pool-373-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:13:50,838 [pool-373-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:13:50,838 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:13:50,838 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:13:50,839 [pool-373-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:13:50,839 [pool-373-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:13:50,839 [pool-373-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:13:50,839 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:13:50,839 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:13:50,839 [pool-373-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:13:50,847 [pool-373-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
```

```
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:13:50,847 [pool-373-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:13:50,850 [pool-373-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R: ans5[63,7],ans6[72,7]
2018-07-14 21:13:50,851 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local318240264_0132_r_000000_0 is done. And is in the process of committing
2018-07-14 21:13:50,856 [pool-373-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 21:13:50,856 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local318240264_0132_r_000000_0 is allowed to commit now
2018-07-14 21:13:50,860 [pool-373-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local318240264_0132_r_000000_0' to
file:/tmp/temp-1902294536/tmp2126601155/_temporary/0/task_local318240264_0132_r_000000
2018-07-14 21:13:50,861 [pool-373-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:13:50,861 [pool-373-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local318240264_0132_r_000000_0' done.
2018-07-14 21:13:50,861 [pool-373-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local318240264_0132_r_000000_0
2018-07-14 21:13:50,861 [Thread-3727] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:13:50,989 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local318240264_0132
2018-07-14 21:13:50,989 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,details
2018-07-14 21:13:50,989 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[63,7],details[8,10],details[-1,-1],ans5[63,7] C:  R:
ans5[63,7],ans6[72,7]
2018-07-14 21:13:50,991 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,991 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,992 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,994 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:13:50,995 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:13:50     2018-07-14 21:13:50
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local318240264_0132 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans5,ans6,details       HASH_JOIN
file:/tmp/temp-1902294536/tmp2126601155,
job_local327896914_0131 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1902294536/tmp2126601155"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local327896914_0131 ->      job_local318240264_0132,
job_local318240264_0132


2018-07-14 21:13:50,995 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,996 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,996 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,997 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,998 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,998 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:13:50,999 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:13:50,999 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:13:51,000 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:13:51,000 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:13:51,006 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:13:51,006 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,300)
(102,Shahrukh,10000,2,500)
(103,Akshay,11000,3,)
(104,Anubhav,5000,4,300)
(105,Pawan,2500,5,100)
(106,Aamir,25000,1,)
(107,Salman,17500,2,)
(108,Ranbir,14000,3,)
(109,Katrina,1000,4,)
(110,Priyanka,2000,5,400)
(111,Tushar,500,1,)
(112,Ajay,5000,2,)
(113,Jubeen,1000,1,)
(114,Madhuri,2000,2,200)
grunt> ans7 = filter ans6 BY ex_sum.sum is not null;
2018-07-14 21:14:29,334 [main] WARN  org.apache.pig.newplan.BaseOperatorPlan -
Encountered Warning IMPLICIT_CAST_TO_LONG 4 time(s).
```

```
grunt> dump ans7;
2018-07-14 21:14:33,026 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:14:33,041 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:33,041 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:14:33,041 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:14:33,041 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:14:33,044 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:14:33,046 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:14:33,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 21:14:33,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 21:14:33,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 21:14:33,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:14:33,056 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:33,057 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:14:33,057 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,058 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:14:33,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:14:33,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:14:33,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:14:33,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:14:33,058 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:14:33,059 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 21:14:33,059 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
```

```
2018-07-14 21:14:33,059 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:14:33,059 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583073059-0
2018-07-14 21:14:33,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:14:33,070 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,079 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:14:33,080 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:14:33,081 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:14:33,081 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:14:33,081 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:14:33,102 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:14:33,115 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local510458790_0133
2018-07-14 21:14:33,173 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:14:33,174 [Thread-3757] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:14:33,177 [Thread-3757] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:14:33,177 [Thread-3757] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:14:33,178 [Thread-3757] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:33,178 [Thread-3757] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:14:33,178 [Thread-3757] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:14:33,178 [Thread-3757] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:14:33,181 [Thread-3757] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:14:33,181 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local510458790_0133_m_000000_0
2018-07-14 21:14:33,186 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:14:33,187 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
```

```
   Locations:

------------------------

2018-07-14 21:14:33,189 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:14:33,189 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:14:33,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:14:33,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:14:33,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:14:33,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:14:33,229 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:14:33,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:14:33,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:14:33,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:14:33,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:14:33,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:14:33,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:14:33,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:14:33,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:14:33,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:14:33,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:14:33,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local510458790_0133_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:14:33,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:14:33,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local510458790_0133_m_000000_0' done.
2018-07-14 21:14:33,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local510458790_0133_m_000000_0
2018-07-14 21:14:33,248 [Thread-3757] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:14:33,249 [Thread-3757] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:14:33,249 [pool-376-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local510458790_0133_r_000000_0
```

```
2018-07-14 21:14:33,257 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:14:33,258 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6f0f8224
2018-07-14 21:14:33,258 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:14:33,259 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local510458790_0133_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:14:33,260 [localfetcher#110] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#110 about to shuffle
output of map attempt_local510458790_0133_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:14:33,260 [localfetcher#110] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local510458790_0133_m_000000_0
2018-07-14 21:14:33,260 [localfetcher#110] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:14:33,260 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:14:33,261 [pool-376-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:14:33,261 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:14:33,262 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:14:33,262 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:14:33,263 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:14:33,264 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:14:33,264 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:14:33,264 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:14:33,264 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:14:33,264 [pool-376-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:14:33,265 [pool-376-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:14:33,265 [pool-376-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:14:33,270 [pool-376-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C: ex_sum[38,8],ex_g[37,6] R:
ex_sum[38,8]
2018-07-14 21:14:33,281 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local510458790_0133_r_000000_0 is done. And is in the process of committing
2018-07-14 21:14:33,282 [pool-376-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 21:14:33,282 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local510458790_0133_r_000000_0 is allowed to commit now
2018-07-14 21:14:33,282 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local510458790_0133_r_000000_0' to
file:/tmp/temp-1902294536/tmp-957597461/_temporary/0/task_local510458790_0133_r_000000
2018-07-14 21:14:33,283 [pool-376-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local510458790_0133_r_000000_0' to
file:/tmp/temp-1902294536/tmp-506516733/_temporary/0/task_local510458790_0133_r_000000
2018-07-14 21:14:33,284 [pool-376-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 21:14:33,284 [pool-376-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local510458790_0133_r_000000_0' done.
2018-07-14 21:14:33,284 [pool-376-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local510458790_0133_r_000000_0
2018-07-14 21:14:33,284 [Thread-3757] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:14:33,477 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local510458790_0133
2018-07-14 21:14:33,477 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:14:33,477 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[35,11],expenses[-1,-1],ex_sum[38,8],ex_g[37,6] C:
ex_sum[38,8],ex_g[37,6] R: ex_sum[38,8]
2018-07-14 21:14:33,480 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:14:33,485 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,487 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,488 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,490 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:14:33,491 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:14:33,491 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:14:33,491 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:14:33,492 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:14:33,492 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:14:33,494 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:14:33,506 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:14:33,507 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
```

```
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,508 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:33,509 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:14:33,514 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:14:33,521 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:14:33,521 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:14:33,521 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:14:33,522 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:14:33,522 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:14:33,522 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:14:33,523 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:14:33,544 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:14:33,558 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1400240313_0134
2018-07-14 21:14:33,615 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:14:33,615 [Thread-3790] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:14:33,620 [Thread-3790] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:14:33,620 [Thread-3790] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:14:33,620 [Thread-3790] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:33,620 [Thread-3790] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:14:33,620 [Thread-3790] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:14:33,620 [Thread-3790] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:14:33,622 [Thread-3790] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:14:33,622 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1400240313_0134_m_000000_0
2018-07-14 21:14:33,626 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:14:33,627 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:14:33,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:14:33,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:14:33,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:14:33,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:14:33,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:14:33,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:14:33,644 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:14:33,645 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:14:33,646 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:14:33,646 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:14:33,649 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],ans7[73,7],details[-1,-1],ans5[63,7],ans5[63,7] C:  R:
ans5[63,7],ans6[72,7]
2018-07-14 21:14:33,649 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:14:33,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:14:33,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:14:33,659 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:14:33,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1400240313_0134_m_000001_0
2018-07-14 21:14:33,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:14:33,662 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:14:33,663 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
```

```
split being processed file:/tmp/temp-1902294536/tmp-506516733/part-r-00000:0+62
2018-07-14 21:14:33,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:14:33,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:14:33,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:14:33,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:14:33,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:14:33,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:14:33,708 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:14:33,708 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:14:33,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[8,10],ans7[73,7],details[-1,-1],ans5[63,7],ans5[63,7] C:  R:
ans5[63,7],ans6[72,7]
2018-07-14 21:14:33,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:14:33,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:14:33,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:14:33,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:14:33,710 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:14:33,711 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:14:33,714 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1400240313_0134_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:14:33,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:14:33,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1400240313_0134_m_000001_0' done.
2018-07-14 21:14:33,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1400240313_0134_m_000001_0
2018-07-14 21:14:33,717 [Thread-3790] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:14:33,718 [Thread-3790] WARN  org.apache.hadoop.mapred.LocalJobRunner -
job_local1400240313_0134
java.lang.Exception: org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Exception while executing (Name: ans5: Local Rearrange[tuple]{int}(false) - scope-3915
Operator Key: scope-3915): org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Exception while executing (Name: details: New For Each(false,false,false,false)[bag] -
scope-3908 Operator Key: scope-3908):
org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more than one
row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN" then "FOREACH ...
GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.hadoop.mapred.LocalJobRunner$Job.runTasks(LocalJobRunner.java:462)
        at org.apache.hadoop.mapred.LocalJobRunner$Job.run(LocalJobRunner.java:522)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception while
```

```
executing (Name: ans5: Local Rearrange[tuple]{int}(false) - scope-3915 Operator Key:
scope-3915): org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception
while executing (Name: details: New For Each(false,false,false,false)[bag] - scope-3908
Operator Key: scope-3908): org.apache.pig.backend.executionengine.ExecException: ERROR 0:
Scalar has more than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause:
"JOIN" then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:314)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POLoca
lRearrange.getNextTuple(POLocalRearrange.java:287)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POUnio
n.getNextTuple(POUnion.java:167)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.runPipe
line(PigGenericMapBase.java:280)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:275)
        at
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapBase.map(Pig
GenericMapBase.java:65)
        at org.apache.hadoop.mapreduce.Mapper.run(Mapper.java:145)
        at org.apache.hadoop.mapred.MapTask.runNewMapper(MapTask.java:784)
        at org.apache.hadoop.mapred.MapTask.run(MapTask.java:341)
        at
org.apache.hadoop.mapred.LocalJobRunner$Job$MapTaskRunnable.run(LocalJobRunner.java:24
3)
        at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
        at java.util.concurrent.FutureTask.run(FutureTask.java:266)
        at
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
        at
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
        at java.lang.Thread.run(Thread.java:748)
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Exception while
executing (Name: details: New For Each(false,false,false,false)[bag] - scope-3908 Operator
Key: scope-3908): org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar
has more than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN"
then "FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:314)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POForE
ach.getNextTuple(POForEach.java:256)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:305)
        ... 14 more
Caused by: org.apache.pig.backend.executionengine.ExecException: ERROR 0: Scalar has more
than one row in the output. 1st : (101,300), 2nd :(102,500) (common cause: "JOIN" then
"FOREACH ... GENERATE foo.bar" should be "foo::bar" )
        at org.apache.pig.impl.builtin.ReadScalars.exec(ReadScalars.java:122)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNext(POUserFunc.java:326)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POUser
Func.getNextLong(POUserFunc.java:410)
        at
```

```
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.getNext(P
hysicalOperator.java:349)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.POIsNu
ll.getNextBoolean(POIsNull.java:78)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.expressionOperators.PONot.
getNextBoolean(PONot.java:71)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.relationalOperators.POFilt
er.getNextTuple(POFilter.java:143)
        at
org.apache.pig.backend.hadoop.executionengine.physicalLayer.PhysicalOperator.processIn
put(PhysicalOperator.java:305)
        ... 16 more
2018-07-14 21:14:33,919 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1400240313_0134
2018-07-14 21:14:33,919 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,ans7,details
2018-07-14 21:14:33,919 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[8,10],ans7[73,7],details[-1,-1],ans5[63,7],ans5[63,7] C:
R: ans5[63,7],ans6[72,7]
2018-07-14 21:14:33,928 [main] WARN
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Ooops!
Some job has failed! Specify -stop_on_failure if you want Pig to stop immediately on failure.
2018-07-14 21:14:33,928 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - job
job_local1400240313_0134 has failed! Stop running all dependent jobs
2018-07-14 21:14:33,928 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:14:33,932 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,933 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,934 [main] ERROR
org.apache.pig.tools.pigstats.mapreduce.MRPigStatsUtil - 1 map reduce job(s) failed!
2018-07-14 21:14:33,938 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:14:33     2018-07-14 21:14:33
HASH_JOIN,GROUP_BY,FILTER

Some jobs have failed! Stop running all dependent jobs

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local510458790_0133 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,MULTI_QUERY,COMBINER

Failed Jobs:
JobId   Alias   Feature Message Outputs
job_local1400240313_0134        ans5,ans6,ans7,details  HASH_JOIN       Message: Job
failed!    file:/tmp/temp-1902294536/tmp-1942164302,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
```

```
Failed to read data from "file:///home/acadgild/employee_details.txt"

Output(s):
Failed to produce result in "file:/tmp/temp-1902294536/tmp-1942164302"

Counters:
Total records written : 0
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local510458790_0133 ->      job_local1400240313_0134,
job_local1400240313_0134


2018-07-14 21:14:33,940 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,941 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,942 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:14:33,943 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Some
jobs have failed! Stop running all dependent jobs
2018-07-14 21:14:33,944 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1066: Unable
to open iterator for alias ans7
Details at logfile: /home/acadgild/pig_1531574068461.log
grunt> 2018-07-14 21:14:39,625 [communication thread] INFO
org.apache.hadoop.mapred.LocalJobRunner - map > sort
^CYou have new mail in /var/spool/mail/acadgild
[acadgild.mmisra ~]$ pig -x local
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log
4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/sl
f4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/07/14 21:14:52 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/07/14 21:14:52 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
2018-07-14 21:14:52,284 [main] INFO  org.apache.pig.Main - Apache Pig version 0.16.0
(r1746530) compiled Jun 01 2016, 23:10:49
2018-07-14 21:14:52,284 [main] INFO  org.apache.pig.Main - Logging error messages to:
/home/acadgild/pig_1531583092282.log
2018-07-14 21:14:52,312 [main] INFO  org.apache.pig.impl.util.Utils - Default bootup file
/home/acadgild/.pigbootup not found
2018-07-14 21:14:52,557 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2018-07-14 21:14:52,558 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:14:52,560 [main] INFO
org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop
file system at: file:///
2018-07-14 21:14:52,657 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:52,678 [main] INFO  org.apache.pig.PigServer - Pig Script ID for the
session: PIG-default-cc84101b-e7ef-4c85-a080-52d89c56d95f
2018-07-14 21:14:52,678 [main] WARN  org.apache.pig.PigServer - ATS is disabled since
yarn.timeline-service.enabled set to false
```

```
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
>> (EmpId:Int,Expense:Int);
2018-07-14 21:14:58,971 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:14:58,972 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> ex_g= group expenses by EmpId;
grunt> ex_m = group  ex_sum all;
2018-07-14 21:15:23,003 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 4, column 14> Undefined alias: ex_sum
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
>> (EmpId:Int,Expense:Int);
2018-07-14 21:16:03,122 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:16:03,122 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> ex_sum= foreach ex_g GENERATE group as EmpId,SUM(expenses.Expense) as sum;
grunt> ex_m = group  ex_sum all;
grunt> ex_max = foreach ex_m generate group,MAX(ex_sum.sum) as max;
grunt> ans = filter ex_sum BY sum==ex_max.max;
2018-07-14 21:16:41,191 [main] WARN  org.apache.hadoop.util.NativeCodeLoader - Unable to
load native-hadoop library for your platform... using builtin-java classes where applicable
grunt> ans1 = foreach ans generate $0,$1,ans.sum;
2018-07-14 21:16:57,575 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1108:
<line 10, column 19> Duplicate schema alias: sum
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> ans = filter ex_sum BY sum==ex_max.max;
grunt> dump ans;
2018-07-14 21:17:32,636 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY,FILTER
2018-07-14 21:17:32,673 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:32,674 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:17:32,744 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:17:32,818 [main] INFO  org.apache.pig.impl.util.SpillableMemoryManager -
Selected heap (PS Old Gen) of size 699400192 to monitor. collectionUsageThreshold =
489580128, usageThreshold = 489580128
2018-07-14 21:17:32,879 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:17:32,897 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:17:32,904 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:17:32,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 21:17:32,937 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 21:17:32,938 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
```

```
2018-07-14 21:17:32,938 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 21:17:32,963 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:32,963 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:17:32,977 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
session.id is deprecated. Instead, use dfs.metrics.session-id
2018-07-14 21:17:32,980 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Initializing JVM Metrics with processName=JobTracker, sessionId=
2018-07-14 21:17:33,010 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:17:33,020 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:17:33,021 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:17:33,021 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.output.compress is deprecated. Instead, use
mapreduce.output.fileoutputformat.compress
2018-07-14 21:17:33,024 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:17:33,026 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:17:33,043 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:17:33,043 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:17:33,044 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2018-07-14 21:17:33,060 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:17:33,070 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:17:33,070 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:17:33,070 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583253069-0
2018-07-14 21:17:33,186 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:17:33,187 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker.http.address is deprecated. Instead, use
mapreduce.jobtracker.http.address
2018-07-14 21:17:33,204 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:33,226 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.task.id is deprecated. Instead,
use mapreduce.task.attempt.id
2018-07-14 21:17:33,563 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:17:33,579 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
```

```
PigTextInputFormat
2018-07-14 21:17:33,587 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:17:33,587 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:17:33,607 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:17:33,702 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:17:33,836 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local266369597_0001
2018-07-14 21:17:34,007 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:17:34,011 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local266369597_0001
2018-07-14 21:17:34,011 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:17:34,011 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:17:34,016 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 0%
complete
2018-07-14 21:17:34,016 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local266369597_0001]
2018-07-14 21:17:34,019 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:17:34,069 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:17:34,069 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:17:34,069 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:34,069 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:17:34,069 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:17:34,073 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:17:34,147 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:17:34,149 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local266369597_0001_m_000000_0
2018-07-14 21:17:34,220 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:17:34,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
```

```
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:17:34,245 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:17:34,251 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:17:34,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:17:34,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:17:34,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:17:34,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:17:34,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:17:34,293 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:17:34,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:17:34,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:17:34,337 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:17:34,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:17:34,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:17:34,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:17:34,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:17:34,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:17:34,394 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigCombiner$Combine -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:17:34,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:17:34,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local266369597_0001_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:17:34,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:17:34,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local266369597_0001_m_000000_0' done.
2018-07-14 21:17:34,425 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local266369597_0001_m_000000_0
```

```
2018-07-14 21:17:34,425 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:17:34,428 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:17:34,433 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local266369597_0001_r_000000_0
2018-07-14 21:17:34,457 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:17:34,463 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@30db55de
2018-07-14 21:17:34,483 [pool-3-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:17:34,502 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local266369597_0001_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:17:34,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 16%
complete
2018-07-14 21:17:34,516 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local266369597_0001]
2018-07-14 21:17:34,570 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#1 about to shuffle
output of map attempt_local266369597_0001_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:17:34,576 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local266369597_0001_m_000000_0
2018-07-14 21:17:34,579 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:17:34,583 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:17:34,586 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:17:34,586 [pool-3-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:17:34,594 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 21:17:34,594 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:17:34,596 [pool-3-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:17:34,597 [pool-3-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:17:34,598 [pool-3-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:17:34,598 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 21:17:34,598 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:17:34,599 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:17:34,619 [pool-3-thread-1] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.skip.on is deprecated. Instead,
use mapreduce.job.skiprecords
```

```
2018-07-14 21:17:34,620 [pool-3-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:17:34,621 [pool-3-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:17:34,637 [pool-3-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:17:34,641 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local266369597_0001_r_000000_0 is done. And is in the process of committing
2018-07-14 21:17:34,644 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:17:34,645 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local266369597_0001_r_000000_0 is allowed to commit now
2018-07-14 21:17:34,649 [pool-3-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local266369597_0001_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1755616165/_temporary/0/task_local266369597_0001_r_00000
0
2018-07-14 21:17:34,654 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:17:34,654 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local266369597_0001_r_000000_0' done.
2018-07-14 21:17:34,654 [pool-3-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local266369597_0001_r_000000_0
2018-07-14 21:17:34,654 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:17:34,818 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 21:17:34,832 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:34,850 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:34,851 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2018-07-14 21:17:34,853 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:34,912 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:17:34,912 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:17:34,913 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:17:34,913 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:17:34,916 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:17:34,952 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:17:34,960 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:34,968 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 21:17:34,968 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:17:34,977 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:17:34,995 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:17:34,995 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:17:34,995 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:17:35,029 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:17:35,074 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1489223576_0002
2018-07-14 21:17:35,192 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:17:35,196 [Thread-53] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:17:35,209 [Thread-53] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:17:35,209 [Thread-53] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:17:35,209 [Thread-53] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:35,209 [Thread-53] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:17:35,209 [Thread-53] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:17:35,212 [Thread-53] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:17:35,217 [Thread-53] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:17:35,217 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1489223576_0002_m_000000_0
2018-07-14 21:17:35,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:17:35,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:17:35,240 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1755616165/part-r-00000:0+62
2018-07-14 21:17:35,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:17:35,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:17:35,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:17:35,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:17:35,260 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:17:35,261 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:17:35,267 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:17:35,268 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:17:35,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[8,9],ex_m[7,7]
C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:17:35,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:17:35,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:17:35,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:17:35,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 21:17:35,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:17:35,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:17:35,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1489223576_0002_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:17:35,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:17:35,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1489223576_0002_m_000000_0' done.
2018-07-14 21:17:35,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1489223576_0002_m_000000_0
2018-07-14 21:17:35,307 [Thread-53] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:17:35,308 [Thread-53] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:17:35,311 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1489223576_0002_r_000000_0
2018-07-14 21:17:35,329 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:17:35,329 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5f4f67f2
2018-07-14 21:17:35,332 [pool-6-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:17:35,337 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1489223576_0002_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:17:35,341 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#2 about to shuffle
```

```
output of map attempt_local1489223576_0002_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 21:17:35,342 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1489223576_0002_m_000000_0
2018-07-14 21:17:35,343 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 21:17:35,343 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:17:35,345 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:17:35,345 [pool-6-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:17:35,346 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 21:17:35,348 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 21:17:35,349 [pool-6-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 21:17:35,350 [pool-6-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 21:17:35,350 [pool-6-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:17:35,350 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 21:17:35,350 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 21:17:35,352 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:17:35,365 [pool-6-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:17:35,365 [pool-6-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:17:35,370 [pool-6-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[8,9],ex_m[7,7]
C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:17:35,372 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1489223576_0002_r_000000_0 is done. And is in the process of committing
2018-07-14 21:17:35,381 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:17:35,381 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1489223576_0002_r_000000_0 is allowed to commit now
2018-07-14 21:17:35,387 [pool-6-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1489223576_0002_r_000000_0' to
file:/tmp/temp-1646866453/tmp-109765438/_temporary/0/task_local1489223576_0002_r_00000
0
2018-07-14 21:17:35,389 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:17:35,389 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1489223576_0002_r_000000_0' done.
2018-07-14 21:17:35,390 [pool-6-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1489223576_0002_r_000000_0
2018-07-14 21:17:35,390 [Thread-53] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:17:35,455 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1489223576_0002
2018-07-14 21:17:35,455 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 21:17:35,455 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[8,9],ex_m[7,7] C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:17:35,456 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 21:17:35,456 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1489223576_0002]
2018-07-14 21:17:35,601 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:35,601 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:35,602 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:35,612 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:17:35,613 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:17:35,614 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:17:35,623 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:17:35,628 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:35,631 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:35,631 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:17:35,647 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:17:35,677 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:17:35,677 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:17:35,677 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:17:35,708 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:17:35,733 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1189754995_0003
2018-07-14 21:17:35,857 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:17:35,857 [Thread-82] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:17:35,864 [Thread-82] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:17:35,864 [Thread-82] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:17:35,865 [Thread-82] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:35,865 [Thread-82] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:17:35,865 [Thread-82] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:17:35,871 [Thread-82] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:17:35,871 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1189754995_0003_m_000000_0
2018-07-14 21:17:35,878 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:17:35,880 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:17:35,881 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1755616165/part-r-00000:0+62
2018-07-14 21:17:35,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:17:35,892 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:17:35,899 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: ans[10,6] C:  R:
2018-07-14 21:17:35,901 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:17:35,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:17:35,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:17:35,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:17:35,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1189754995_0003_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:17:35,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:17:35,924 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1189754995_0003_m_000000_0 is allowed
to commit now
2018-07-14 21:17:35,930 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1189754995_0003_m_000000_0' to
file:/tmp/temp-1646866453/tmp1195326191/_temporary/0/task_local1189754995_0003_m_00000
0
2018-07-14 21:17:35,933 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:17:35,934 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1189754995_0003_m_000000_0' done.
2018-07-14 21:17:35,934 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1189754995_0003_m_000000_0
2018-07-14 21:17:35,934 [Thread-82] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:17:36,063 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1189754995_0003
2018-07-14 21:17:36,063 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans
2018-07-14 21:17:36,063 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[10,6] C:  R:
2018-07-14 21:17:36,067 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,068 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,069 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,079 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:17:36,092 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:17:33     2018-07-14 21:17:36
GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1189754995_0003        1       0       n/a     n/a     n/a     n/a     0
0       0       0       ans     MAP_ONLY
file:/tmp/temp-1646866453/tmp1195326191,
job_local1489223576_0002        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local266369597_0001 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1646866453/tmp1195326191"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local266369597_0001 ->      job_local1489223576_0002,job_local1189754995_0003,
job_local1489223576_0002        ->      job_local1189754995_0003,
```

```
job_local1189754995_0003


2018-07-14 21:17:36,093 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,098 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,100 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,113 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,114 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,116 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,125 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,131 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,132 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:17:36,134 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:17:36,140 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:17:36,140 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:17:36,141 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:17:36,155 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:17:36,155 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,500)
grunt> ans1 = JOIN details By EmpId,ans By EmpId;
2018-07-14 21:18:47,406 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 11, column 12> Undefined alias: details
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> describe ans;
ans: {EmpId: int,sum: long}
grunt> ans1 = JOIN details By EmpId,ans By ans.EmpId;
2018-07-14 21:19:02,343 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 11, column 12> Undefined alias: details
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> ans1 = JOIN details By EmpId,ans By ans.EmpId;
2018-07-14 21:19:25,747 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1200: Pig
script failed to parse:
<line 11, column 12> Undefined alias: details
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> dump details;
2018-07-14 21:19:29,169 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1003: Unable
to find an operator for alias details
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> dump details;
2018-07-14 21:19:34,833 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1003: Unable
to find an operator for alias details
Details at logfile: /home/acadgild/pig_1531583092282.log
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int);
```

```
2018-07-14 21:19:45,793 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:45,794 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt> ans1 = JOIN details By EmpId,ans By EmpId;
grunt> dump ans1;
2018-07-14 21:19:56,844 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:19:56,867 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:56,867 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:19:56,868 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:19:56,868 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:19:56,875 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:19:56,879 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:19:56,879 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:19:56,886 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 21:19:56,886 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 21:19:56,886 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 21:19:56,886 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 21:19:56,886 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 21:19:56,904 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:56,905 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:19:56,906 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:56,909 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:19:56,911 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:19:56,912 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:19:56,913 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:19:56,913 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:19:56,913 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:19:56,915 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:19:56,916 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:19:56,916 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:19:56,916 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583396916-0
2018-07-14 21:19:56,933 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:19:56,936 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:56,946 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:19:56,948 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:19:56,950 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:19:56,950 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:19:56,951 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:19:56,980 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:19:56,999 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1361409340_0004
2018-07-14 21:19:57,111 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:19:57,112 [Thread-107] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:19:57,119 [Thread-107] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:19:57,119 [Thread-107] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:19:57,119 [Thread-107] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:57,119 [Thread-107] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:19:57,119 [Thread-107] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:19:57,120 [Thread-107] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:19:57,127 [Thread-107] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for map tasks
2018-07-14 21:19:57,127 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1361409340_0004_m_000000_0
2018-07-14 21:19:57,137 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:57,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:19:57,147 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:19:57,148 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:19:57,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:19:57,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:19:57,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:19:57,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:19:57,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:19:57,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:19:57,184 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:57,184 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:19:57,192 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:19:57,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:19:57,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:19:57,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:19:57,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:19:57,201 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:19:57,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:19:57,215 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1361409340_0004_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:19:57,217 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
```

```
2018-07-14 21:19:57,217 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1361409340_0004_m_000000_0' done.
2018-07-14 21:19:57,217 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1361409340_0004_m_000000_0
2018-07-14 21:19:57,218 [Thread-107] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:19:57,221 [Thread-107] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:19:57,224 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1361409340_0004_r_000000_0
2018-07-14 21:19:57,242 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:57,242 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@479038a8
2018-07-14 21:19:57,247 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:19:57,250 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1361409340_0004_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:19:57,253 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#3 about to shuffle
output of map attempt_local1361409340_0004_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:19:57,254 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1361409340_0004_m_000000_0
2018-07-14 21:19:57,255 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:19:57,255 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:19:57,256 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:19:57,256 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:19:57,257 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:19:57,259 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:19:57,261 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:19:57,261 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:19:57,261 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:19:57,261 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:19:57,262 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:19:57,262 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:19:57,272 [pool-11-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:57,272 [pool-11-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
```

```
- SchemaTupleBackend has already been initialized
2018-07-14 21:19:57,281 [pool-11-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:19:57,287 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1361409340_0004_r_000000_0 is done. And is in the process of committing
2018-07-14 21:19:57,296 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:19:57,296 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1361409340_0004_r_000000_0 is allowed to commit now
2018-07-14 21:19:57,298 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1361409340_0004_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1739573052/_temporary/0/task_local1361409340_0004_r_0000
00
2018-07-14 21:19:57,299 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:19:57,299 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1361409340_0004_r_000000_0' done.
2018-07-14 21:19:57,299 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1361409340_0004_r_000000_0
2018-07-14 21:19:57,299 [Thread-107] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:19:57,412 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1361409340_0004
2018-07-14 21:19:57,413 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:19:57,413 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:19:57,413 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 21:19:57,414 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,415 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,415 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,419 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:19:57,419 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:19:57,419 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:19:57,419 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:19:57,420 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:19:57,431 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:19:57,432 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
```

```
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,437 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:57,437 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:19:57,443 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:19:57,452 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:19:57,452 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:19:57,452 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:19:57,479 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:19:57,503 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1279394388_0005
2018-07-14 21:19:57,615 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:19:57,615 [Thread-136] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:19:57,620 [Thread-136] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:19:57,620 [Thread-136] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:19:57,620 [Thread-136] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:57,620 [Thread-136] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:19:57,620 [Thread-136] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:19:57,620 [Thread-136] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:19:57,623 [Thread-136] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:19:57,624 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1279394388_0005_m_000000_0
2018-07-14 21:19:57,630 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:57,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:19:57,633 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1739573052/part-r-00000:0+62
2018-07-14 21:19:57,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:19:57,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:19:57,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:19:57,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:19:57,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:19:57,652 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:19:57,653 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:57,654 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:19:57,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[8,9],ex_m[7,7]
C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:19:57,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:19:57,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:19:57,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:19:57,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 21:19:57,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:19:57,675 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:19:57,676 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1279394388_0005_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:19:57,680 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:19:57,680 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1279394388_0005_m_000000_0' done.
2018-07-14 21:19:57,680 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1279394388_0005_m_000000_0
2018-07-14 21:19:57,681 [Thread-136] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:19:57,681 [Thread-136] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:19:57,685 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1279394388_0005_r_000000_0
2018-07-14 21:19:57,699 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:57,699 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@54f13041
2018-07-14 21:19:57,699 [pool-14-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:19:57,702 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
```

```
attempt_local1279394388_0005_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:19:57,704 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#4 about to shuffle
output of map attempt_local1279394388_0005_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 21:19:57,704 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1279394388_0005_m_000000_0
2018-07-14 21:19:57,705 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 21:19:57,705 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:19:57,707 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:19:57,707 [pool-14-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:19:57,708 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:19:57,708 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 21:19:57,710 [pool-14-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 21:19:57,710 [pool-14-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 21:19:57,710 [pool-14-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:19:57,710 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:19:57,712 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 21:19:57,712 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:19:57,721 [pool-14-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:57,721 [pool-14-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:19:57,726 [pool-14-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[8,9],ex_m[7,7]
C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:19:57,730 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1279394388_0005_r_000000_0 is done. And is in the process of committing
2018-07-14 21:19:57,732 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:19:57,732 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1279394388_0005_r_000000_0 is allowed to commit now
2018-07-14 21:19:57,734 [pool-14-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1279394388_0005_r_000000_0' to
file:/tmp/temp-1646866453/tmp-492972114/_temporary/0/task_local1279394388_0005_r_00000
0
2018-07-14 21:19:57,736 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:19:57,736 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1279394388_0005_r_000000_0' done.
2018-07-14 21:19:57,736 [pool-14-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- Finishing task: attempt_local1279394388_0005_r_000000_0
2018-07-14 21:19:57,736 [Thread-136] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:19:57,917 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1279394388_0005
2018-07-14 21:19:57,917 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 21:19:57,917 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[8,9],ex_m[7,7] C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:19:57,925 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 21:19:57,939 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,947 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,952 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:57,973 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:19:57,976 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:19:57,976 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:19:57,976 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:19:57,984 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:19:57,984 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:19:57,988 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:19:58,003 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:19:58,004 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,019 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:58,019 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:19:58,031 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:19:58,033 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:19:58,035 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:19:58,035 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
2018-07-14 21:19:58,036 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:19:58,049 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:19:58,049 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:19:58,050 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:19:58,084 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:19:58,108 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1998890417_0006
2018-07-14 21:19:58,219 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:19:58,220 [Thread-165] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:19:58,225 [Thread-165] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:19:58,225 [Thread-165] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:19:58,226 [Thread-165] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:58,228 [Thread-165] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:19:58,228 [Thread-165] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:19:58,228 [Thread-165] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:19:58,232 [Thread-165] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:19:58,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1998890417_0006_m_000000_0
2018-07-14 21:19:58,238 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:58,239 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:19:58,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:19:58,245 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:19:58,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:19:58,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:19:58,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:19:58,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:19:58,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:19:58,262 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:19:58,270 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:58,270 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:19:58,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans1[12,7],ans[10,6],ans1[12,7] C:  R:
2018-07-14 21:19:58,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:19:58,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:19:58,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:19:58,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:19:58,280 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:19:58,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:19:58,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1998890417_0006_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:19:58,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:19:58,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1998890417_0006_m_000000_0' done.
2018-07-14 21:19:58,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1998890417_0006_m_000000_0
2018-07-14 21:19:58,288 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1998890417_0006_m_000001_0
2018-07-14 21:19:58,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:58,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:19:58,294 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1739573052/part-r-00000:0+62
2018-07-14 21:19:58,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:19:58,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:19:58,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:19:58,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:19:58,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:19:58,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:19:58,314 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:58,314 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:19:58,319 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans1[12,7],ans[10,6],ans1[12,7] C:  R:
2018-07-14 21:19:58,320 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:19:58,329 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:19:58,329 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:19:58,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:19:58,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:19:58,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:19:58,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 12; bufvoid = 104857600
2018-07-14 21:19:58,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 21:19:58,332 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:19:58,334 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1998890417_0006_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:19:58,335 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:19:58,335 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1998890417_0006_m_000001_0' done.
2018-07-14 21:19:58,335 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1998890417_0006_m_000001_0
2018-07-14 21:19:58,335 [Thread-165] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:19:58,335 [Thread-165] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:19:58,339 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1998890417_0006_r_000000_0
2018-07-14 21:19:58,349 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:19:58,349 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1ee4e53f
2018-07-14 21:19:58,351 [pool-17-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
```

```
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:19:58,354 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1998890417_0006_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:19:58,358 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#5 about to shuffle
output of map attempt_local1998890417_0006_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:19:58,359 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1998890417_0006_m_000000_0
2018-07-14 21:19:58,359 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 21:19:58,362 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#5 about to shuffle
output of map attempt_local1998890417_0006_m_000001_0 decomp: 16 len: 20 to MEMORY
2018-07-14 21:19:58,362 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 16 bytes from map-output
for attempt_local1998890417_0006_m_000001_0
2018-07-14 21:19:58,362 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 16, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->368
2018-07-14 21:19:58,366 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:19:58,366 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:19:58,366 [pool-17-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:19:58,367 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:19:58,367 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 352 bytes
2018-07-14 21:19:58,369 [pool-17-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 368 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:19:58,369 [pool-17-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 370 bytes from
disk
2018-07-14 21:19:58,369 [pool-17-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:19:58,369 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:19:58,369 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 358 bytes
2018-07-14 21:19:58,369 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:19:58,382 [pool-17-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:19:58,383 [pool-17-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:19:58,384 [pool-17-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans1[12,7],ans[10,6],ans1[12,7] C:  R:
2018-07-14 21:19:58,393 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1998890417_0006_r_000000_0 is done. And is in the process of committing
2018-07-14 21:19:58,397 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- 2 / 2 copied.
2018-07-14 21:19:58,397 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1998890417_0006_r_000000_0 is allowed to commit now
2018-07-14 21:19:58,399 [pool-17-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1998890417_0006_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1761290128/_temporary/0/task_local1998890417_0006_r_0000
00
2018-07-14 21:19:58,402 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:19:58,402 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1998890417_0006_r_000000_0' done.
2018-07-14 21:19:58,402 [pool-17-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1998890417_0006_r_000000_0
2018-07-14 21:19:58,402 [Thread-165] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:19:58,504 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1998890417_0006
2018-07-14 21:19:58,504 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,ans1,details
2018-07-14 21:19:58,505 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[11,10],details[-1,-1],ans1[12,7],ans[10,6],ans1[12,7] C:
R:
2018-07-14 21:19:58,515 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,519 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,523 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,557 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:19:58,561 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:19:56     2018-07-14 21:19:58
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1279394388_0005        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local1361409340_0004        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1998890417_0006        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans,ans1,details        HASH_JOIN
file:/tmp/temp-1646866453/tmp-1761290128,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1646866453/tmp-1761290128"
```

```
Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1361409340_0004        ->
job_local1279394388_0005,job_local1998890417_0006,
job_local1279394388_0005        ->      job_local1998890417_0006,
job_local1998890417_0006


2018-07-14 21:19:58,564 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,566 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,568 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,571 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,572 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,573 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,580 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,581 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,584 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:19:58,587 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:19:58,588 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:19:58,588 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:19:58,588 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:19:58,597 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:19:58,597 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,10000,2,102,500)
grunt> ans2 = foreach ans1 generate $0,$1,ans.sum;
grunt> dump ans2;
2018-07-14 21:20:43,732 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:20:43,765 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:43,766 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:20:43,766 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:20:43,767 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
```

```
2018-07-14 21:20:43,780 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 21:20:43,796 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:20:43,800 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:20:43,801 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:20:43,802 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 21:20:43,802 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 5
2018-07-14 21:20:43,804 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 21:20:43,804 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 21:20:43,804 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 21:20:43,804 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 21:20:43,804 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 4
2018-07-14 21:20:43,816 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:43,817 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:20:43,819 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:43,821 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:20:43,822 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:20:43,822 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:20:43,822 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:20:43,822 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:20:43,823 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:20:43,826 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:20:43,827 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
```

```
2018-07-14 21:20:43,827 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:20:43,827 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583443826-0
2018-07-14 21:20:43,837 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:20:43,839 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:43,850 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:20:43,853 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:20:43,854 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:43,854 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:43,854 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:20:43,882 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:20:43,904 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local40265366_0007
2018-07-14 21:20:44,031 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:20:44,031 [Thread-197] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:20:44,041 [Thread-197] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:20:44,041 [Thread-197] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:20:44,042 [Thread-197] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:44,042 [Thread-197] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:20:44,042 [Thread-197] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:44,042 [Thread-197] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:20:44,046 [Thread-197] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:20:44,046 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local40265366_0007_m_000000_0
2018-07-14 21:20:44,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:44,057 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
```

```
   Locations:

-----------------------

2018-07-14 21:20:44,062 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:20:44,063 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:20:44,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:20:44,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:20:44,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:20:44,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:20:44,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:20:44,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:20:44,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:44,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:20:44,088 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:20:44,093 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:20:44,093 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:20:44,093 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:20:44,093 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:20:44,093 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:20:44,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:20:44,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local40265366_0007_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:20:44,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:20:44,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local40265366_0007_m_000000_0' done.
2018-07-14 21:20:44,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local40265366_0007_m_000000_0
2018-07-14 21:20:44,112 [Thread-197] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:20:44,113 [Thread-197] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:20:44,117 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local40265366_0007_r_000000_0
2018-07-14 21:20:44,122 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
```

```
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:44,122 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2283d11e
2018-07-14 21:20:44,124 [pool-20-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:20:44,126 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local40265366_0007_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:20:44,129 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#6 about to shuffle
output of map attempt_local40265366_0007_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:20:44,129 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local40265366_0007_m_000000_0
2018-07-14 21:20:44,130 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:20:44,131 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:20:44,132 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:20:44,132 [pool-20-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:20:44,133 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:20:44,133 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:20:44,135 [pool-20-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:20:44,135 [pool-20-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:20:44,135 [pool-20-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:20:44,135 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:20:44,135 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:20:44,136 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:20:44,145 [pool-20-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:44,145 [pool-20-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:20:44,149 [pool-20-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:20:44,155 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local40265366_0007_r_000000_0 is done. And is in the process of committing
2018-07-14 21:20:44,159 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:20:44,159 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local40265366_0007_r_000000_0 is allowed to commit now
```

```
2018-07-14 21:20:44,161 [pool-20-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local40265366_0007_r_000000_0' to
file:/tmp/temp-1646866453/tmp1859624829/_temporary/0/task_local40265366_0007_r_000000
2018-07-14 21:20:44,162 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:20:44,162 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local40265366_0007_r_000000_0' done.
2018-07-14 21:20:44,162 [pool-20-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local40265366_0007_r_000000_0
2018-07-14 21:20:44,162 [Thread-197] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:20:44,335 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local40265366_0007
2018-07-14 21:20:44,335 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:20:44,335 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:20:44,335 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 25%
complete
2018-07-14 21:20:44,336 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,337 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,337 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,345 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:20:44,346 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:20:44,346 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:20:44,346 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:20:44,347 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:20:44,358 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:20:44,360 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,362 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:44,362 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:44,372 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:20:44,389 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:44,389 [JobControl] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:44,389 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:20:44,417 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:20:44,444 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2089596636_0008
2018-07-14 21:20:44,551 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:20:44,552 [Thread-226] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:20:44,556 [Thread-226] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:20:44,556 [Thread-226] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:20:44,556 [Thread-226] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:44,556 [Thread-226] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:20:44,556 [Thread-226] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:44,556 [Thread-226] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:20:44,559 [Thread-226] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:20:44,560 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2089596636_0008_m_000000_0
2018-07-14 21:20:44,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:44,567 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:20:44,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp1859624829/part-r-00000:0+62
2018-07-14 21:20:44,592 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:20:44,593 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:20:44,593 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:20:44,593 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:20:44,593 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:20:44,593 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:20:44,595 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:44,596 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:20:44,600 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[8,9],ex_m[7,7]
C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:20:44,602 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:20:44,602 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:20:44,602 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:20:44,602 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 21:20:44,602 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:20:44,611 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:20:44,612 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2089596636_0008_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:20:44,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:20:44,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2089596636_0008_m_000000_0' done.
2018-07-14 21:20:44,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2089596636_0008_m_000000_0
2018-07-14 21:20:44,615 [Thread-226] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:20:44,615 [Thread-226] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:20:44,619 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local2089596636_0008_r_000000_0
2018-07-14 21:20:44,627 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:44,627 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@153d3a43
2018-07-14 21:20:44,628 [pool-23-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:20:44,629 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2089596636_0008_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:20:44,630 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#7 about to shuffle
output of map attempt_local2089596636_0008_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 21:20:44,631 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local2089596636_0008_m_000000_0
2018-07-14 21:20:44,631 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 21:20:44,632 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 21:20:44,633 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:20:44,633 [pool-23-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:20:44,634 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:20:44,634 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 21:20:44,635 [pool-23-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 21:20:44,636 [pool-23-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 21:20:44,636 [pool-23-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:20:44,636 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:20:44,636 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 21:20:44,636 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:20:44,648 [pool-23-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:44,648 [pool-23-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:20:44,652 [pool-23-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[8,9],ex_m[7,7]
C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:20:44,653 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2089596636_0008_r_000000_0 is done. And is in the process of committing
2018-07-14 21:20:44,654 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:20:44,654 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local2089596636_0008_r_000000_0 is allowed to commit now
2018-07-14 21:20:44,656 [pool-23-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2089596636_0008_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1021157819/_temporary/0/task_local2089596636_0008_r_0000
00
2018-07-14 21:20:44,657 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:20:44,657 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2089596636_0008_r_000000_0' done.
2018-07-14 21:20:44,657 [pool-23-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local2089596636_0008_r_000000_0
2018-07-14 21:20:44,657 [Thread-226] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:20:44,854 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2089596636_0008
2018-07-14 21:20:44,855 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 21:20:44,855 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[8,9],ex_m[7,7] C: ex_max[8,9],ex_m[7,7] R: ex_max[8,9]
2018-07-14 21:20:44,867 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:20:44,877 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,884 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,888 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,906 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:20:44,907 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:20:44,910 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 21:20:44,921 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:20:44,924 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:44,928 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:44,928 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:44,941 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:20:44,950 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:44,950 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:44,950 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:20:44,974 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:20:44,993 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1762214042_0009
2018-07-14 21:20:45,102 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:20:45,103 [Thread-255] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:20:45,108 [Thread-255] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:20:45,108 [Thread-255] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:20:45,108 [Thread-255] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:45,108 [Thread-255] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:45,109 [Thread-255] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
```

```
2018-07-14 21:20:45,113 [Thread-255] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:20:45,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1762214042_0009_m_000000_0
2018-07-14 21:20:45,121 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:45,122 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:20:45,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp1859624829/part-r-00000:0+62
2018-07-14 21:20:45,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:45,125 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,128 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: ans[10,6] C:  R:
2018-07-14 21:20:45,148 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,157 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:45,157 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:45,160 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:20:45,161 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1762214042_0009_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:20:45,163 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:20:45,163 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1762214042_0009_m_000000_0 is allowed
to commit now
2018-07-14 21:20:45,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1762214042_0009_m_000000_0' to
file:/tmp/temp-1646866453/tmp2021174302/_temporary/0/task_local1762214042_0009_m_00000
0
2018-07-14 21:20:45,166 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1762214042_0009_m_000000_0' to
file:/tmp/temp-1646866453/tmp1396149795/_temporary/0/task_local1762214042_0009_m_00000
0
2018-07-14 21:20:45,167 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:20:45,167 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1762214042_0009_m_000000_0' done.
2018-07-14 21:20:45,167 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1762214042_0009_m_000000_0
2018-07-14 21:20:45,167 [Thread-255] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
```

```
task executor complete.
2018-07-14 21:20:45,303 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1762214042_0009
2018-07-14 21:20:45,304 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans
2018-07-14 21:20:45,304 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[10,6] C:  R:
2018-07-14 21:20:45,308 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 75%
complete
2018-07-14 21:20:45,316 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,323 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,329 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,348 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:20:45,349 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:20:45,350 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:20:45,350 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:20:45,351 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=284
2018-07-14 21:20:45,351 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:20:45,354 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:20:45,368 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:20:45,369 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,372 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:45,372 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:45,380 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:20:45,389 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:45,389 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:45,389 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
```

```
2018-07-14 21:20:45,392 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:20:45,393 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:45,394 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:45,394 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:20:45,418 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:20:45,433 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1022045474_0010
2018-07-14 21:20:45,507 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:20:45,507 [Thread-284] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:20:45,513 [Thread-284] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:20:45,513 [Thread-284] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:20:45,513 [Thread-284] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:45,513 [Thread-284] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:20:45,513 [Thread-284] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:20:45,513 [Thread-284] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:20:45,516 [Thread-284] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:20:45,517 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1022045474_0010_m_000000_0
2018-07-14 21:20:45,522 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:45,523 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:20:45,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:20:45,526 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:20:45,550 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:20:45,550 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 21:20:45,550 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:20:45,550 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:20:45,550 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:20:45,551 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:20:45,554 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:45,554 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,557 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans1[12,7],ans[-1,-1],ans1[12,7] C:  R: ans2[13,7]
2018-07-14 21:20:45,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:20:45,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:20:45,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:20:45,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 21:20:45,561 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:20:45,563 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:20:45,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1022045474_0010_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:20:45,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:20:45,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1022045474_0010_m_000000_0' done.
2018-07-14 21:20:45,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1022045474_0010_m_000000_0
2018-07-14 21:20:45,566 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1022045474_0010_m_000001_0
2018-07-14 21:20:45,568 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:45,569 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 9
Input split[0]:
   Length = 9
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:20:45,571 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp1396149795/part-m-00000:0+9
2018-07-14 21:20:45,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:20:45,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 21:20:45,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:20:45,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:20:45,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:20:45,590 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:20:45,591 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:45,592 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,595 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans1[12,7],ans[-1,-1],ans1[12,7] C:  R: ans2[13,7]
2018-07-14 21:20:45,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:20:45,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:20:45,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:20:45,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 9; bufvoid = 104857600
2018-07-14 21:20:45,596 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 21:20:45,597 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:20:45,598 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1022045474_0010_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:20:45,599 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:20:45,599 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1022045474_0010_m_000001_0' done.
2018-07-14 21:20:45,599 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1022045474_0010_m_000001_0
2018-07-14 21:20:45,600 [Thread-284] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:20:45,605 [Thread-284] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:20:45,606 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1022045474_0010_r_000000_0
2018-07-14 21:20:45,617 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:20:45,617 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@313827bb
2018-07-14 21:20:45,618 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:20:45,620 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1022045474_0010_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:20:45,622 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#8 about to shuffle
output of map attempt_local1022045474_0010_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 21:20:45,622 [localfetcher#8] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local1022045474_0010_m_000000_0
2018-07-14 21:20:45,622 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->286
2018-07-14 21:20:45,623 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#8 about to shuffle
output of map attempt_local1022045474_0010_m_000001_0 decomp: 13 len: 17 to MEMORY
2018-07-14 21:20:45,624 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 13 bytes from map-output
for attempt_local1022045474_0010_m_000001_0
2018-07-14 21:20:45,624 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 13, inMemoryMapOutputs.size() -> 2, commitMemory -> 286, usedMemory
->299
2018-07-14 21:20:45,624 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:20:45,626 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:20:45,627 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:20:45,629 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:20:45,629 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 283 bytes
2018-07-14 21:20:45,630 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 299 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:20:45,630 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 301 bytes from
disk
2018-07-14 21:20:45,630 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:20:45,630 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:20:45,630 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 289 bytes
2018-07-14 21:20:45,631 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:20:45,643 [pool-28-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:20:45,643 [pool-28-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,648 [pool-28-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans1[12,7],ans[-1,-1],ans1[12,7] C:  R: ans2[13,7]
2018-07-14 21:20:45,649 [pool-28-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,659 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:45,659 [pool-28-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:20:45,664 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1022045474_0010_r_000000_0 is done. And is in the process of committing
2018-07-14 21:20:45,666 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:20:45,666 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local1022045474_0010_r_000000_0 is allowed to commit now
2018-07-14 21:20:45,668 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1022045474_0010_r_000000_0' to
file:/tmp/temp-1646866453/tmp195361533/_temporary/0/task_local1022045474_0010_r_000000
2018-07-14 21:20:45,669 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:20:45,669 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1022045474_0010_r_000000_0' done.
2018-07-14 21:20:45,669 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1022045474_0010_r_000000_0
2018-07-14 21:20:45,669 [Thread-284] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:20:45,811 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1022045474_0010
2018-07-14 21:20:45,811 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,ans1,ans2,details
2018-07-14 21:20:45,811 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[11,10],details[-1,-1],ans1[12,7],ans[-1,-1],ans1[12,7] C:
R: ans2[13,7]
2018-07-14 21:20:45,823 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,831 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,832 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,860 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:20:45,870 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:20:43     2018-07-14 21:20:45
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1022045474_0010        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans,ans1,ans2,details   HASH_JOIN
file:/tmp/temp-1646866453/tmp195361533,
job_local1762214042_0009        1       0       n/a     n/a     n/a     n/a     0
0       0       0       ans     MULTI_QUERY,MAP_ONLY
job_local2089596636_0008        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local40265366_0007  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp-1646866453/tmp195361533"

Counters:
```

```
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local40265366_0007  ->      job_local2089596636_0008,job_local1762214042_0009,
job_local2089596636_0008        ->      job_local1762214042_0009,
job_local1762214042_0009        ->      job_local1022045474_0010,
job_local1022045474_0010


2018-07-14 21:20:45,883 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,885 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,887 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,895 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,896 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,898 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,901 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,903 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,904 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,906 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,908 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,909 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:20:45,914 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:20:45,914 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:20:45,915 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:20:45,915 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:20:45,932 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:20:45,932 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,500)
grunt> ans3 = JOIN details BY EmpId, ex_sum BY EmpId;
grunt> ans4 = foreach ans3 GENERATE $0,$1;
grunt> dump ans4;
2018-07-14 21:21:48,811 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:21:48,835 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:21:48,835 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:21:48,836 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
```

```
2018-07-14 21:21:48,836 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:21:48,837 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 21:21:48,837 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for expenses: $1
2018-07-14 21:21:48,843 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:21:48,844 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 21:21:48,844 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:21:48,844 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:21:48,854 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:21:48,855 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:21:48,857 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:48,858 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:21:48,858 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:21:48,858 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:21:48,858 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:21:48,859 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:21:48,859 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:21:48,860 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:21:48,860 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:21:48,860 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:21:48,860 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583508860-0
2018-07-14 21:21:48,867 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:21:48,868 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 21:21:48,877 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:21:48,880 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:21:48,880 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:21:48,880 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:21:48,881 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:21:48,901 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:21:48,917 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1335006788_0011
2018-07-14 21:21:49,015 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:21:49,016 [Thread-316] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:21:49,019 [Thread-316] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:21:49,020 [Thread-316] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:21:49,020 [Thread-316] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:21:49,020 [Thread-316] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:21:49,020 [Thread-316] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:21:49,020 [Thread-316] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:21:49,023 [Thread-316] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:21:49,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1335006788_0011_m_000000_0
2018-07-14 21:21:49,027 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:21:49,028 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:21:49,031 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:21:49,031 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:21:49,047 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:21:49,047 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:21:49,047 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:21:49,047 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:21:49,047 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:21:49,048 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:21:49,049 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:21:49,049 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:21:49,050 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_g[3,6] C:  R: ex_sum[-1,-1]
2018-07-14 21:21:49,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:21:49,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:21:49,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:21:49,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 81; bufvoid = 104857600
2018-07-14 21:21:49,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:21:49,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:21:49,057 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1335006788_0011_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:21:49,058 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:21:49,059 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1335006788_0011_m_000000_0' done.
2018-07-14 21:21:49,059 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1335006788_0011_m_000000_0
2018-07-14 21:21:49,065 [Thread-316] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:21:49,065 [Thread-316] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:21:49,069 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1335006788_0011_r_000000_0
2018-07-14 21:21:49,076 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:21:49,076 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@181ca3d6
2018-07-14 21:21:49,077 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:21:49,078 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1335006788_0011_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
```

```
2018-07-14 21:21:49,080 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#9 about to shuffle
output of map attempt_local1335006788_0011_m_000000_0 decomp: 101 len: 105 to MEMORY
2018-07-14 21:21:49,080 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 101 bytes from
map-output for attempt_local1335006788_0011_m_000000_0
2018-07-14 21:21:49,080 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 101, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->101
2018-07-14 21:21:49,081 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:21:49,082 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:21:49,082 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:21:49,083 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:21:49,083 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 21:21:49,086 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 101 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:21:49,086 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 105 bytes from
disk
2018-07-14 21:21:49,086 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:21:49,086 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:21:49,086 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 21:21:49,086 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:21:49,098 [pool-31-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:21:49,098 [pool-31-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:21:49,100 [pool-31-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_g[3,6] C:  R: ex_sum[-1,-1]
2018-07-14 21:21:49,101 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1335006788_0011_r_000000_0 is done. And is in the process of committing
2018-07-14 21:21:49,104 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:21:49,104 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1335006788_0011_r_000000_0 is allowed to commit now
2018-07-14 21:21:49,107 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1335006788_0011_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1547030538/_temporary/0/task_local1335006788_0011_r_0000
00
2018-07-14 21:21:49,110 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:21:49,110 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1335006788_0011_r_000000_0' done.
2018-07-14 21:21:49,110 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1335006788_0011_r_000000_0
2018-07-14 21:21:49,110 [Thread-316] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
reduce task executor complete.
2018-07-14 21:21:49,320 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1335006788_0011
2018-07-14 21:21:49,320 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:21:49,320 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_g[3,6] C:  R: ex_sum[-1,-1]
2018-07-14 21:21:49,324 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:21:49,330 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,334 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,338 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,364 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:21:49,366 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:21:49,366 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:21:49,366 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:21:49,369 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=317
2018-07-14 21:21:49,369 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:21:49,370 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:21:49,383 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:21:49,385 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,388 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:21:49,388 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:21:49,396 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:21:49,398 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:21:49,398 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:21:49,398 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:21:49,398 [JobControl] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:21:49,407 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:21:49,407 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:21:49,407 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:21:49,431 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:21:49,449 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1161848926_0012
2018-07-14 21:21:49,524 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:21:49,525 [Thread-345] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:21:49,530 [Thread-345] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:21:49,530 [Thread-345] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:21:49,530 [Thread-345] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:21:49,530 [Thread-345] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:21:49,530 [Thread-345] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:21:49,530 [Thread-345] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:21:49,532 [Thread-345] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:21:49,532 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1161848926_0012_m_000000_0
2018-07-14 21:21:49,536 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:21:49,537 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:21:49,538 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:21:49,539 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:21:49,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:21:49,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 21:21:49,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:21:49,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:21:49,555 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:21:49,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:21:49,558 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:21:49,558 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:21:49,560 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[14,7],details[11,10],details[-1,-1],ans3[14,7] C:  R: ans4[15,7]
2018-07-14 21:21:49,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:21:49,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:21:49,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:21:49,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 21:21:49,565 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:21:49,567 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:21:49,567 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1161848926_0012_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:21:49,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:21:49,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1161848926_0012_m_000000_0' done.
2018-07-14 21:21:49,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1161848926_0012_m_000000_0
2018-07-14 21:21:49,572 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1161848926_0012_m_000001_0
2018-07-14 21:21:49,575 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:21:49,576 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 42
Input split[0]:
   Length = 42
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:21:49,577 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1547030538/part-r-00000:0+42
2018-07-14 21:21:49,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:21:49,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 21:21:49,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:21:49,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:21:49,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:21:49,694 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:21:49,695 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:21:49,696 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:21:49,697 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[14,7],details[11,10],details[-1,-1],ans3[14,7] C:  R: ans4[15,7]
2018-07-14 21:21:49,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:21:49,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:21:49,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:21:49,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 63; bufvoid = 104857600
2018-07-14 21:21:49,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:21:49,699 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:21:49,700 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1161848926_0012_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:21:49,701 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:21:49,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1161848926_0012_m_000001_0' done.
2018-07-14 21:21:49,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1161848926_0012_m_000001_0
2018-07-14 21:21:49,702 [Thread-345] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:21:49,702 [Thread-345] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:21:49,702 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1161848926_0012_r_000000_0
2018-07-14 21:21:49,712 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:21:49,712 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1ea9abc2
2018-07-14 21:21:49,718 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:21:49,718 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1161848926_0012_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:21:49,724 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#10 about to shuffle
output of map attempt_local1161848926_0012_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 21:21:49,724 [localfetcher#10] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local1161848926_0012_m_000000_0
2018-07-14 21:21:49,724 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->286
2018-07-14 21:21:49,726 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#10 about to shuffle
output of map attempt_local1161848926_0012_m_000001_0 decomp: 79 len: 83 to MEMORY
2018-07-14 21:21:49,726 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 79 bytes from map-output
for attempt_local1161848926_0012_m_000001_0
2018-07-14 21:21:49,726 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 79, inMemoryMapOutputs.size() -> 2, commitMemory -> 286, usedMemory
->365
2018-07-14 21:21:49,726 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:21:49,727 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:21:49,727 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:21:49,728 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:21:49,728 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 349 bytes
2018-07-14 21:21:49,729 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 365 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:21:49,729 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 367 bytes from
disk
2018-07-14 21:21:49,729 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:21:49,729 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:21:49,730 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 355 bytes
2018-07-14 21:21:49,730 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:21:49,741 [pool-34-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:21:49,742 [pool-34-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:21:49,744 [pool-34-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[14,7],details[11,10],details[-1,-1],ans3[14,7] C:  R: ans4[15,7]
2018-07-14 21:21:49,746 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1161848926_0012_r_000000_0 is done. And is in the process of committing
2018-07-14 21:21:49,748 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:21:49,748 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1161848926_0012_r_000000_0 is allowed to commit now
2018-07-14 21:21:49,749 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1161848926_0012_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1848871095/_temporary/0/task_local1161848926_0012_r_0000
00
2018-07-14 21:21:49,750 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- reduce > reduce
2018-07-14 21:21:49,750 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1161848926_0012_r_000000_0' done.
2018-07-14 21:21:49,750 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1161848926_0012_r_000000_0
2018-07-14 21:21:49,750 [Thread-345] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:21:49,881 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1161848926_0012
2018-07-14 21:21:49,881 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,ans4,details
2018-07-14 21:21:49,881 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans3[14,7],details[11,10],details[-1,-1],ans3[14,7] C:  R:
ans4[15,7]
2018-07-14 21:21:49,893 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,896 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,898 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,908 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:21:49,909 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:21:48     2018-07-14 21:21:49
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1161848926_0012        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans3,ans4,details       HASH_JOIN
file:/tmp/temp-1646866453/tmp-1848871095,
job_local1335006788_0011        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 6 records in: "file:/tmp/temp-1646866453/tmp-1848871095"

Counters:
Total records written : 6
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1335006788_0011        ->      job_local1161848926_0012,
job_local1161848926_0012
```

```
2018-07-14 21:21:49,909 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,910 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,911 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,915 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,916 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,917 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:21:49,921 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:21:49,921 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:21:49,922 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:21:49,922 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:21:49,934 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:21:49,934 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh)
(102,Shahrukh)
(104,Anubhav)
(105,Pawan)
(110,Priyanka)
(114,Madhuri)
grunt> ans5 = JOIN details BY EmpId LEFT OUTER, ex_sum BY EmpId;
grunt> dump ans5;
2018-07-14 21:22:51,487 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:22:51,509 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:22:51,510 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:22:51,510 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:22:51,510 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:22:51,519 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:22:51,520 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:22:51,522 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:22:51,522 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:22:51,531 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 21:22:51,531 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:22:51,532 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:51,533 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:22:51,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:22:51,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:22:51,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:22:51,534 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:22:51,535 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:22:51,537 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:22:51,537 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:22:51,537 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:22:51,537 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583571537-0
2018-07-14 21:22:51,545 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:22:51,546 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:51,557 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:22:51,560 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:22:51,560 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:22:51,560 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:22:51,563 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:22:51,587 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:22:51,601 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local693860674_0013
2018-07-14 21:22:51,679 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:22:51,680 [Thread-375] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:22:51,683 [Thread-375] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:22:51,683 [Thread-375] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:22:51,683 [Thread-375] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:22:51,683 [Thread-375] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:22:51,683 [Thread-375] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:22:51,683 [Thread-375] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:22:51,687 [Thread-375] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:22:51,687 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local693860674_0013_m_000000_0
2018-07-14 21:22:51,694 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:22:51,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:22:51,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:22:51,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:22:51,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:22:51,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:22:51,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:22:51,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:22:51,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:22:51,742 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:22:51,743 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:22:51,744 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:22:51,747 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:22:51,748 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
```

```
2018-07-14 21:22:51,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:22:51,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:22:51,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:22:51,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:22:51,758 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:22:51,760 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local693860674_0013_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:22:51,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:22:51,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local693860674_0013_m_000000_0' done.
2018-07-14 21:22:51,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local693860674_0013_m_000000_0
2018-07-14 21:22:51,761 [Thread-375] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:22:51,762 [Thread-375] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:22:51,769 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local693860674_0013_r_000000_0
2018-07-14 21:22:51,774 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:22:51,774 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@550553d9
2018-07-14 21:22:51,775 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:22:51,776 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local693860674_0013_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:22:51,778 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#11 about to shuffle
output of map attempt_local693860674_0013_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:22:51,778 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local693860674_0013_m_000000_0
2018-07-14 21:22:51,778 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:22:51,779 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:22:51,781 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:22:51,781 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:22:51,784 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:22:51,784 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:22:51,785 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
```

```
2018-07-14 21:22:51,786 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:22:51,786 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:22:51,786 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:22:51,787 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:22:51,787 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:22:51,795 [pool-37-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:22:51,795 [pool-37-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:22:51,797 [pool-37-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:22:51,799 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local693860674_0013_r_000000_0 is done. And is in the process of committing
2018-07-14 21:22:51,800 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:22:51,801 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local693860674_0013_r_000000_0 is allowed to commit now
2018-07-14 21:22:51,802 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local693860674_0013_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1476603377/_temporary/0/task_local693860674_0013_r_00000
0
2018-07-14 21:22:51,807 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:22:51,807 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local693860674_0013_r_000000_0' done.
2018-07-14 21:22:51,807 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local693860674_0013_r_000000_0
2018-07-14 21:22:51,807 [Thread-375] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:22:51,983 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local693860674_0013
2018-07-14 21:22:51,983 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:22:51,983 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:22:51,984 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:22:51,984 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:51,985 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:51,985 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:51,990 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
```

```
2018-07-14 21:22:51,991 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:22:51,991 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:22:51,991 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:22:51,996 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:22:51,996 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:22:51,996 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:22:52,006 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:22:52,007 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,010 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:22:52,010 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:22:52,017 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:22:52,026 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:22:52,026 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:22:52,026 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:22:52,028 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:22:52,028 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:22:52,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:22:52,028 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:22:52,057 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:22:52,076 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1022016777_0014
2018-07-14 21:22:52,175 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:22:52,176 [Thread-404] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:22:52,180 [Thread-404] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:22:52,180 [Thread-404] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:22:52,180 [Thread-404] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:22:52,180 [Thread-404] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:22:52,180 [Thread-404] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:22:52,180 [Thread-404] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:22:52,184 [Thread-404] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:22:52,184 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1022016777_0014_m_000000_0
2018-07-14 21:22:52,192 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:22:52,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:22:52,199 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:22:52,199 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:22:52,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:22:52,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:22:52,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:22:52,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:22:52,222 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:22:52,223 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:22:52,227 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:22:52,228 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:22:52,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[16,7],details[11,10],details[-1,-1],ans5[16,7] C:  R: ans5[16,7]
2018-07-14 21:22:52,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:22:52,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
```

```
2018-07-14 21:22:52,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:22:52,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:22:52,242 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:22:52,243 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:22:52,245 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1022016777_0014_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:22:52,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:22:52,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1022016777_0014_m_000000_0' done.
2018-07-14 21:22:52,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1022016777_0014_m_000000_0
2018-07-14 21:22:52,247 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1022016777_0014_m_000001_0
2018-07-14 21:22:52,248 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:22:52,250 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:22:52,251 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1476603377/part-r-00000:0+62
2018-07-14 21:22:52,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:22:52,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:22:52,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:22:52,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:22:52,408 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:22:52,409 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:22:52,411 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:22:52,411 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:22:52,414 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[16,7],details[11,10],details[-1,-1],ans5[16,7] C:  R: ans5[16,7]
2018-07-14 21:22:52,415 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:22:52,415 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
```

```
2018-07-14 21:22:52,415 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:22:52,415 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:22:52,415 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:22:52,417 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:22:52,419 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1022016777_0014_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:22:52,421 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:22:52,421 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1022016777_0014_m_000001_0' done.
2018-07-14 21:22:52,421 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1022016777_0014_m_000001_0
2018-07-14 21:22:52,421 [Thread-404] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:22:52,421 [Thread-404] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:22:52,423 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1022016777_0014_r_000000_0
2018-07-14 21:22:52,430 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:22:52,430 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@376d75bc
2018-07-14 21:22:52,431 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:22:52,433 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1022016777_0014_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:22:52,434 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#12 about to shuffle
output of map attempt_local1022016777_0014_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:22:52,434 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1022016777_0014_m_000000_0
2018-07-14 21:22:52,435 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 21:22:52,436 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#12 about to shuffle
output of map attempt_local1022016777_0014_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:22:52,436 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1022016777_0014_m_000001_0
2018-07-14 21:22:52,436 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 21:22:52,437 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:22:52,437 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:22:52,437 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
```

```
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:22:52,438 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:22:52,438 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:22:52,442 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:22:52,443 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:22:52,443 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:22:52,443 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:22:52,443 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:22:52,443 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:22:52,452 [pool-40-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:22:52,452 [pool-40-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:22:52,458 [pool-40-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[16,7],details[11,10],details[-1,-1],ans5[16,7] C:  R: ans5[16,7]
2018-07-14 21:22:52,467 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1022016777_0014_r_000000_0 is done. And is in the process of committing
2018-07-14 21:22:52,471 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:22:52,471 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1022016777_0014_r_000000_0 is allowed to commit now
2018-07-14 21:22:52,475 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1022016777_0014_r_000000_0' to
file:/tmp/temp-1646866453/tmp646384095/_temporary/0/task_local1022016777_0014_r_000000
2018-07-14 21:22:52,476 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:22:52,476 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1022016777_0014_r_000000_0' done.
2018-07-14 21:22:52,476 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1022016777_0014_r_000000_0
2018-07-14 21:22:52,477 [Thread-404] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:22:52,506 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1022016777_0014
2018-07-14 21:22:52,506 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 21:22:52,507 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[16,7],details[11,10],details[-1,-1],ans5[16,7] C:  R:
ans5[16,7]
2018-07-14 21:22:52,512 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1022016777_0014]
2018-07-14 21:22:52,689 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,693 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,696 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,723 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:22:52,729 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:22:51     2018-07-14 21:22:52
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1022016777_0014        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp-1646866453/tmp646384095,
job_local693860674_0013 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1646866453/tmp646384095"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local693860674_0013 ->      job_local1022016777_0014,
job_local1022016777_0014


2018-07-14 21:22:52,731 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,733 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,735 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,742 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,744 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,745 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:22:52,748 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:22:52,749 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:22:52,750 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2018-07-14 21:22:52,750 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:22:52,760 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:22:52,760 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(103,Akshay,11000,3,,)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(106,Aamir,25000,1,,)
(107,Salman,17500,2,,)
(108,Ranbir,14000,3,,)
(109,Katrina,1000,4,,)
(110,Priyanka,2000,5,110,400)
(111,Tushar,500,1,,)
(112,Ajay,5000,2,,)
(113,Jubeen,1000,1,,)
(114,Madhuri,2000,2,114,200)
grunt> ans6 = foreach ans5 generate ..$3,$5;
grunt> dump ans6;
2018-07-14 21:23:07,074 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 21:23:07,095 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:23:07,095 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:23:07,095 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:23:07,096 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:23:07,099 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:23:07,102 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:23:07,103 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:23:07,103 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:23:07,113 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:23:07,114 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:23:07,114 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:07,115 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:23:07,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:23:07,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:23:07,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:23:07,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:23:07,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:23:07,118 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:23:07,118 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:23:07,118 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:23:07,118 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583587118-0
2018-07-14 21:23:07,126 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:23:07,127 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:07,137 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:23:07,140 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:23:07,141 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:23:07,141 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:23:07,142 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:23:07,163 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:23:07,181 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2038949676_0015
2018-07-14 21:23:07,268 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:23:07,269 [Thread-434] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:23:07,272 [Thread-434] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:23:07,273 [Thread-434] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:23:07,273 [Thread-434] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:23:07,273 [Thread-434] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:23:07,273 [Thread-434] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 21:23:07,273 [Thread-434] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:23:07,275 [Thread-434] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:23:07,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2038949676_0015_m_000000_0
2018-07-14 21:23:07,281 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:23:07,282 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:23:07,285 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:23:07,287 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:23:07,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:23:07,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:23:07,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:23:07,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:23:07,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:23:07,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:23:07,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:23:07,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:23:07,308 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:23:07,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:23:07,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:23:07,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:23:07,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:23:07,316 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:23:07,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:23:07,333 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.Task - Task:attempt_local2038949676_0015_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:23:07,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:23:07,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2038949676_0015_m_000000_0' done.
2018-07-14 21:23:07,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2038949676_0015_m_000000_0
2018-07-14 21:23:07,337 [Thread-434] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:23:07,337 [Thread-434] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:23:07,339 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local2038949676_0015_r_000000_0
2018-07-14 21:23:07,346 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:23:07,346 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@32963917
2018-07-14 21:23:07,347 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:23:07,348 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2038949676_0015_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:23:07,349 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#13 about to shuffle
output of map attempt_local2038949676_0015_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:23:07,349 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local2038949676_0015_m_000000_0
2018-07-14 21:23:07,349 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:23:07,353 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:23:07,354 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:23:07,354 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:23:07,355 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:23:07,355 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:23:07,356 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:23:07,356 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:23:07,356 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:23:07,356 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:23:07,356 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:23:07,356 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
```

```
2018-07-14 21:23:07,464 [pool-43-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:23:07,464 [pool-43-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:23:07,466 [pool-43-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:23:07,467 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2038949676_0015_r_000000_0 is done. And is in the process of committing
2018-07-14 21:23:07,469 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:23:07,469 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local2038949676_0015_r_000000_0 is allowed to commit now
2018-07-14 21:23:07,469 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2038949676_0015_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1625584916/_temporary/0/task_local2038949676_0015_r_0000
00
2018-07-14 21:23:07,471 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:23:07,471 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2038949676_0015_r_000000_0' done.
2018-07-14 21:23:07,471 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local2038949676_0015_r_000000_0
2018-07-14 21:23:07,471 [Thread-434] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:23:07,627 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2038949676_0015
2018-07-14 21:23:07,627 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:23:07,627 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:23:07,631 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:23:07,638 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:07,643 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:07,647 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:07,674 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:23:07,675 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:23:07,676 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:23:07,677 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:23:07,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
```

```
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:23:07,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:23:07,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:23:07,698 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:23:07,700 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:07,703 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:23:07,703 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:23:07,710 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:23:07,711 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:23:07,712 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:23:07,712 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:23:07,712 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:23:07,720 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:23:07,720 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:23:07,720 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:23:07,743 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:23:07,760 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1427142547_0016
2018-07-14 21:23:07,842 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:23:07,842 [Thread-463] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:23:07,845 [Thread-463] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:23:07,845 [Thread-463] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:23:07,845 [Thread-463] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:23:07,845 [Thread-463] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:23:07,845 [Thread-463] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 21:23:07,845 [Thread-463] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:23:07,850 [Thread-463] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:23:07,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1427142547_0016_m_000000_0
2018-07-14 21:23:07,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:23:07,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:23:07,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:23:07,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:23:07,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:23:07,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:23:07,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:23:07,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:23:07,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:23:07,877 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:23:07,879 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:23:07,879 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:23:07,881 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R: ans5[16,7],ans6[17,7]
2018-07-14 21:23:07,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:23:07,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:23:07,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:23:07,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:23:07,883 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:23:07,884 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:23:07,887 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1427142547_0016_m_000000_0 is done. And
is in the process of committing
```

```
2018-07-14 21:23:07,888 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:23:07,888 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1427142547_0016_m_000000_0' done.
2018-07-14 21:23:07,888 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1427142547_0016_m_000000_0
2018-07-14 21:23:07,888 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1427142547_0016_m_000001_0
2018-07-14 21:23:07,890 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:23:07,890 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:23:07,891 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-1625584916/part-r-00000:0+62
2018-07-14 21:23:07,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:23:07,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:23:07,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:23:07,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:23:07,908 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:23:07,909 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:23:07,910 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:23:07,910 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:23:07,912 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R: ans5[16,7],ans6[17,7]
2018-07-14 21:23:07,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:23:07,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:23:07,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:23:07,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:23:07,913 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:23:07,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:23:07,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1427142547_0016_m_000001_0 is done. And
is in the process of committing
```

```
2018-07-14 21:23:07,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:23:07,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1427142547_0016_m_000001_0' done.
2018-07-14 21:23:07,916 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1427142547_0016_m_000001_0
2018-07-14 21:23:07,916 [Thread-463] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:23:07,917 [Thread-463] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:23:07,919 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1427142547_0016_r_000000_0
2018-07-14 21:23:07,928 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:23:07,928 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@60d88df0
2018-07-14 21:23:07,929 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:23:07,930 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1427142547_0016_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:23:07,931 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#14 about to shuffle
output of map attempt_local1427142547_0016_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:23:07,931 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1427142547_0016_m_000000_0
2018-07-14 21:23:07,931 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 21:23:07,932 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#14 about to shuffle
output of map attempt_local1427142547_0016_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:23:07,932 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1427142547_0016_m_000001_0
2018-07-14 21:23:07,932 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 21:23:07,932 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:23:07,933 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:23:07,933 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:23:07,934 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:23:07,934 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:23:07,937 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:23:07,937 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:23:07,937 [pool-46-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:23:07,937 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:23:07,938 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:23:07,938 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:23:07,946 [pool-46-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:23:07,946 [pool-46-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:23:07,948 [pool-46-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R: ans5[16,7],ans6[17,7]
2018-07-14 21:23:07,950 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1427142547_0016_r_000000_0 is done. And is in the process of committing
2018-07-14 21:23:07,951 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:23:07,952 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1427142547_0016_r_000000_0 is allowed to commit now
2018-07-14 21:23:07,954 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1427142547_0016_r_000000_0' to
file:/tmp/temp-1646866453/tmp-1109556901/_temporary/0/task_local1427142547_0016_r_0000
00
2018-07-14 21:23:07,954 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:23:07,954 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1427142547_0016_r_000000_0' done.
2018-07-14 21:23:07,954 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1427142547_0016_r_000000_0
2018-07-14 21:23:07,954 [Thread-463] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:23:08,144 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1427142547_0016
2018-07-14 21:23:08,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,details
2018-07-14 21:23:08,145 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans6[17,7]
2018-07-14 21:23:08,160 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,163 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,166 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,179 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:23:08,181 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:23:07     2018-07-14 21:23:08
HASH_JOIN,GROUP_BY

Success!
```

```
Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1427142547_0016        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,ans6,details       HASH_JOIN
file:/tmp/temp-1646866453/tmp-1109556901,
job_local2038949676_0015        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp-1646866453/tmp-1109556901"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local2038949676_0015        ->      job_local1427142547_0016,
job_local1427142547_0016


2018-07-14 21:23:08,182 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,183 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,183 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,191 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,195 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,195 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:23:08,198 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:23:08,198 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:23:08,198 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:23:08,198 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:23:08,207 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:23:08,207 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,300)
(102,Shahrukh,10000,2,500)
(103,Akshay,11000,3,)
(104,Anubhav,5000,4,300)
(105,Pawan,2500,5,100)
(106,Aamir,25000,1,)
(107,Salman,17500,2,)
```

```
(108,Ranbir,14000,3,)
(109,Katrina,1000,4,)
(110,Priyanka,2000,5,400)
(111,Tushar,500,1,)
(112,Ajay,5000,2,)
(113,Jubeen,1000,1,)
(114,Madhuri,2000,2,200)
grunt> ans7 = filter ans6 BY sum is null;
grunt> dump ans7;
2018-07-14 21:24:05,459 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:24:05,487 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:05,487 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:24:05,487 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 21:24:05,488 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:24:05,497 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:24:05,500 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:24:05,501 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:24:05,501 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:24:05,514 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:05,514 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:24:05,515 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:05,516 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:24:05,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:24:05,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:24:05,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:24:05,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:24:05,517 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:24:05,519 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
```

```
2018-07-14 21:24:05,519 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:24:05,519 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:24:05,519 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583645519-0
2018-07-14 21:24:05,529 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:24:05,530 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:05,538 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:24:05,540 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:24:05,541 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:05,541 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:24:05,541 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:24:05,564 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:24:05,588 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1457312101_0017
2018-07-14 21:24:05,680 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:24:05,681 [Thread-493] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:24:05,684 [Thread-493] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:24:05,685 [Thread-493] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:24:05,685 [Thread-493] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:05,685 [Thread-493] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:24:05,685 [Thread-493] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:24:05,685 [Thread-493] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:24:05,687 [Thread-493] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:24:05,687 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1457312101_0017_m_000000_0
2018-07-14 21:24:05,690 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:05,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
```

```
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:24:05,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:24:05,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:24:05,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:24:05,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:24:05,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:24:05,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:24:05,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:24:05,713 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:24:05,714 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:05,714 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:24:05,719 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:24:05,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:24:05,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:24:05,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:24:05,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:24:05,722 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:24:05,731 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:24:05,732 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1457312101_0017_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:24:05,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:24:05,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1457312101_0017_m_000000_0' done.
2018-07-14 21:24:05,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1457312101_0017_m_000000_0
2018-07-14 21:24:05,734 [Thread-493] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:24:05,734 [Thread-493] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:24:05,738 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- Starting task: attempt_local1457312101_0017_r_000000_0
2018-07-14 21:24:05,743 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:05,743 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@f9ff63
2018-07-14 21:24:05,744 [pool-49-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:24:05,746 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1457312101_0017_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:24:05,747 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#15 about to shuffle
output of map attempt_local1457312101_0017_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:24:05,747 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1457312101_0017_m_000000_0
2018-07-14 21:24:05,747 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:24:05,748 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:24:05,749 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:24:05,749 [pool-49-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:24:05,750 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:24:05,750 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:24:05,751 [pool-49-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:24:05,751 [pool-49-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:24:05,751 [pool-49-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:24:05,751 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:24:05,751 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:24:05,752 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:24:05,761 [pool-49-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:05,762 [pool-49-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:24:05,764 [pool-49-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:24:05,765 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1457312101_0017_r_000000_0 is done. And is in the process of committing
2018-07-14 21:24:05,766 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
```

```
2018-07-14 21:24:05,766 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1457312101_0017_r_000000_0 is allowed to commit now
2018-07-14 21:24:05,767 [pool-49-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1457312101_0017_r_000000_0' to
file:/tmp/temp-1646866453/tmp994969899/_temporary/0/task_local1457312101_0017_r_000000
2018-07-14 21:24:05,768 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:24:05,768 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1457312101_0017_r_000000_0' done.
2018-07-14 21:24:05,768 [pool-49-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1457312101_0017_r_000000_0
2018-07-14 21:24:05,768 [Thread-493] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:24:05,882 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1457312101_0017
2018-07-14 21:24:05,882 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:24:05,882 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:24:05,885 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:24:05,891 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:05,897 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:05,902 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:05,913 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:24:05,916 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:24:05,917 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:24:05,917 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:24:05,924 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:24:05,924 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:24:05,931 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:24:05,954 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:24:05,955 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:05,956 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 21:24:05,956 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:24:05,964 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:24:05,965 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:24:05,966 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:05,966 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:24:05,966 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:24:05,972 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:05,973 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:24:05,973 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:24:05,993 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:24:06,008 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1640295325_0018
2018-07-14 21:24:06,087 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:24:06,087 [Thread-522] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:24:06,093 [Thread-522] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:24:06,093 [Thread-522] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:24:06,093 [Thread-522] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:06,093 [Thread-522] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:24:06,093 [Thread-522] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:24:06,094 [Thread-522] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:24:06,096 [Thread-522] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:24:06,096 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1640295325_0018_m_000000_0
2018-07-14 21:24:06,102 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:06,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
```

```
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:24:06,105 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:24:06,105 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:24:06,132 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:24:06,132 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:24:06,132 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:24:06,132 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:24:06,132 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:24:06,134 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:24:06,135 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:06,135 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:24:06,137 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans7[18,7],ans6[17,7]
2018-07-14 21:24:06,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:24:06,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:24:06,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:24:06,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 21:24:06,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:24:06,140 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:24:06,141 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1640295325_0018_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:24:06,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:24:06,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1640295325_0018_m_000000_0' done.
2018-07-14 21:24:06,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1640295325_0018_m_000000_0
2018-07-14 21:24:06,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1640295325_0018_m_000001_0
2018-07-14 21:24:06,145 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:06,146 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
```

```
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:24:06,148 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp994969899/part-r-00000:0+62
2018-07-14 21:24:06,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:24:06,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:24:06,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:24:06,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:24:06,164 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:24:06,165 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:24:06,165 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:06,166 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:24:06,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans7[18,7],ans6[17,7]
2018-07-14 21:24:06,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:24:06,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:24:06,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:24:06,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:24:06,169 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:24:06,170 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:24:06,171 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1640295325_0018_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:24:06,176 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:24:06,176 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1640295325_0018_m_000001_0' done.
2018-07-14 21:24:06,176 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1640295325_0018_m_000001_0
2018-07-14 21:24:06,176 [Thread-522] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:24:06,176 [Thread-522] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:24:06,180 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1640295325_0018_r_000000_0
```

```
2018-07-14 21:24:06,184 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:06,185 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4a6eeaed
2018-07-14 21:24:06,186 [pool-52-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:24:06,187 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1640295325_0018_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:24:06,188 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#16 about to shuffle
output of map attempt_local1640295325_0018_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:24:06,188 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1640295325_0018_m_000001_0
2018-07-14 21:24:06,188 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 21:24:06,189 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#16 about to shuffle
output of map attempt_local1640295325_0018_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 21:24:06,189 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1640295325_0018_m_000000_0
2018-07-14 21:24:06,189 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 21:24:06,190 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:24:06,191 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:24:06,191 [pool-52-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:24:06,191 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:24:06,191 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 21:24:06,195 [pool-52-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:24:06,195 [pool-52-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 21:24:06,195 [pool-52-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:24:06,196 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:24:06,196 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 21:24:06,196 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:24:06,204 [pool-52-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:06,204 [pool-52-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
```

```
2018-07-14 21:24:06,209 [pool-52-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans7[18,7],ans6[17,7]
2018-07-14 21:24:06,212 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1640295325_0018_r_000000_0 is done. And is in the process of committing
2018-07-14 21:24:06,213 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:24:06,213 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1640295325_0018_r_000000_0 is allowed to commit now
2018-07-14 21:24:06,215 [pool-52-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1640295325_0018_r_000000_0' to
file:/tmp/temp-1646866453/tmp-890354001/_temporary/0/task_local1640295325_0018_r_00000
0
2018-07-14 21:24:06,216 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:24:06,216 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1640295325_0018_r_000000_0' done.
2018-07-14 21:24:06,216 [pool-52-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1640295325_0018_r_000000_0
2018-07-14 21:24:06,216 [Thread-522] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:24:06,391 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1640295325_0018
2018-07-14 21:24:06,391 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,ans7,details
2018-07-14 21:24:06,391 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans7[18,7],ans6[17,7]
2018-07-14 21:24:06,409 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,413 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,415 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,430 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:24:06,434 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:24:05     2018-07-14 21:24:06
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1457312101_0017        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1640295325_0018        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,ans6,ans7,details  HASH_JOIN
file:/tmp/temp-1646866453/tmp-890354001,

Input(s):
```

```
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 8 records in: "file:/tmp/temp-1646866453/tmp-890354001"

Counters:
Total records written : 8
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1457312101_0017        ->      job_local1640295325_0018,
job_local1640295325_0018


2018-07-14 21:24:06,436 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,437 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,438 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,441 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,442 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,448 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:06,451 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:24:06,451 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:06,451 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:24:06,452 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:24:06,459 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:06,459 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(103,Akshay,11000,3,)
(106,Aamir,25000,1,)
(107,Salman,17500,2,)
(108,Ranbir,14000,3,)
(109,Katrina,1000,4,)
(111,Tushar,500,1,)
(112,Ajay,5000,2,)
(113,Jubeen,1000,1,)
grunt> ans7 = filter ans6 BY sum is null;
grunt> ans8 = foreach ans7 generate EmpId,Name;
grunt> dump ans8;
2018-07-14 21:24:52,223 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 21:24:52,241 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:52,242 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:24:52,242 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
```

```
2018-07-14 21:24:52,242 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 21:24:52,245 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 21:24:52,247 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 21:24:52,248 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 21:24:52,250 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 21:24:52,250 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 21:24:52,264 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:52,265 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:24:52,265 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:52,266 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:24:52,266 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:24:52,266 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:24:52,266 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:24:52,267 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 21:24:52,267 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:24:52,268 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:24:52,268 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 21:24:52,268 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 21:24:52,269 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531583692268-0
2018-07-14 21:24:52,278 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:24:52,279 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:52,287 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
```

```
not be found. See Job or Job#setJar(String).
2018-07-14 21:24:52,289 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:24:52,290 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:52,290 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:24:52,291 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:24:52,314 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 21:24:52,328 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local530441329_0019
2018-07-14 21:24:52,402 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:24:52,402 [Thread-552] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:24:52,405 [Thread-552] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:24:52,405 [Thread-552] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:24:52,405 [Thread-552] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:52,405 [Thread-552] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:24:52,405 [Thread-552] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:24:52,405 [Thread-552] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:24:52,415 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local530441329_0019_m_000000_0
2018-07-14 21:24:52,416 [Thread-552] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:24:52,418 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:52,419 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:24:52,421 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:24:52,421 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 21:24:52,435 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:24:52,435 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:24:52,435 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:24:52,435 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:24:52,435 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:24:52,436 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:24:52,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:52,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 21:24:52,442 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:24:52,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:24:52,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:24:52,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:24:52,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 21:24:52,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 21:24:52,456 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:24:52,461 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local530441329_0019_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:24:52,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:24:52,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local530441329_0019_m_000000_0' done.
2018-07-14 21:24:52,462 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local530441329_0019_m_000000_0
2018-07-14 21:24:52,462 [Thread-552] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:24:52,463 [Thread-552] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:24:52,465 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local530441329_0019_r_000000_0
2018-07-14 21:24:52,470 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:52,470 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3f4c8875
2018-07-14 21:24:52,471 [pool-55-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:24:52,472 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local530441329_0019_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:24:52,473 [localfetcher#17] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#17 about to shuffle
output of map attempt_local530441329_0019_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 21:24:52,473 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local530441329_0019_m_000000_0
2018-07-14 21:24:52,473 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 21:24:52,474 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:24:52,474 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:24:52,474 [pool-55-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:24:52,476 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:24:52,476 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:24:52,476 [pool-55-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:24:52,477 [pool-55-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 21:24:52,477 [pool-55-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:24:52,477 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:24:52,477 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 21:24:52,477 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:24:52,486 [pool-55-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:52,486 [pool-55-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:24:52,490 [pool-55-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C: ex_sum[6,8],ex_g[3,6] R:
ex_sum[6,8]
2018-07-14 21:24:52,493 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local530441329_0019_r_000000_0 is done. And is in the process of committing
2018-07-14 21:24:52,494 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 21:24:52,494 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local530441329_0019_r_000000_0 is allowed to commit now
2018-07-14 21:24:52,494 [pool-55-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local530441329_0019_r_000000_0' to
file:/tmp/temp-1646866453/tmp-2032353218/_temporary/0/task_local530441329_0019_r_00000
0
2018-07-14 21:24:52,495 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:24:52,495 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local530441329_0019_r_000000_0' done.
2018-07-14 21:24:52,496 [pool-55-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local530441329_0019_r_000000_0
2018-07-14 21:24:52,496 [Thread-552] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
reduce task executor complete.
2018-07-14 21:24:52,705 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local530441329_0019
2018-07-14 21:24:52,706 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 21:24:52,707 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[1,11],expenses[-1,-1],ex_sum[6,8],ex_g[3,6] C:
ex_sum[6,8],ex_g[3,6] R: ex_sum[6,8]
2018-07-14 21:24:52,710 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 21:24:52,716 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:52,720 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:52,724 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:52,737 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 21:24:52,738 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 21:24:52,739 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 21:24:52,739 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 21:24:52,742 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 21:24:52,742 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 21:24:52,744 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 21:24:52,760 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 21:24:52,765 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:52,767 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:52,768 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:24:52,774 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 21:24:52,777 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 21:24:52,777 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:52,777 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
```

```
2018-07-14 21:24:52,777 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:24:52,785 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:52,785 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 21:24:52,785 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 21:24:52,805 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 21:24:52,821 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local39874636_0020
2018-07-14 21:24:52,897 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 21:24:52,898 [Thread-581] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 21:24:52,902 [Thread-581] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 21:24:52,902 [Thread-581] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 21:24:52,903 [Thread-581] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:52,903 [Thread-581] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 21:24:52,903 [Thread-581] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 21:24:52,903 [Thread-581] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 21:24:52,905 [Thread-581] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 21:24:52,905 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local39874636_0020_m_000000_0
2018-07-14 21:24:52,911 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:52,912 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:24:52,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 21:24:52,914 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 21:24:52,935 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 21:24:52,935 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:24:52,935 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:24:52,935 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:24:52,935 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:24:52,935 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:24:52,937 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:52,937 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:24:52,941 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans5[-1,-1],ans7[19,7],ans8[-1,-1]
2018-07-14 21:24:52,947 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:24:52,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:24:52,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:24:52,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 21:24:52,948 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 21:24:52,949 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:24:52,950 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local39874636_0020_m_000000_0 is done. And
is in the process of committing
2018-07-14 21:24:52,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:24:52,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local39874636_0020_m_000000_0' done.
2018-07-14 21:24:52,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local39874636_0020_m_000000_0
2018-07-14 21:24:52,951 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local39874636_0020_m_000001_0
2018-07-14 21:24:52,955 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:52,956 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 21:24:52,959 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp-1646866453/tmp-2032353218/part-r-00000:0+62
2018-07-14 21:24:53,070 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 21:24:53,071 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 21:24:53,071 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 21:24:53,071 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 21:24:53,071 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 21:24:53,071 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 21:24:53,072 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:53,072 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 21:24:53,075 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans5[-1,-1],ans7[19,7],ans8[-1,-1]
2018-07-14 21:24:53,077 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 21:24:53,077 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 21:24:53,077 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 21:24:53,077 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 21:24:53,077 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 21:24:53,078 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 21:24:53,080 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local39874636_0020_m_000001_0 is done. And
is in the process of committing
2018-07-14 21:24:53,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 21:24:53,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local39874636_0020_m_000001_0' done.
2018-07-14 21:24:53,082 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local39874636_0020_m_000001_0
2018-07-14 21:24:53,082 [Thread-581] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 21:24:53,083 [Thread-581] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 21:24:53,084 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local39874636_0020_r_000000_0
2018-07-14 21:24:53,092 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 21:24:53,092 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@7018697c
2018-07-14 21:24:53,093 [pool-58-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 21:24:53,094 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local39874636_0020_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 21:24:53,095 [localfetcher#18] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#18 about to shuffle
output of map attempt_local39874636_0020_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 21:24:53,096 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local39874636_0020_m_000001_0
2018-07-14 21:24:53,097 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 21:24:53,098 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#18 about to shuffle
output of map attempt_local39874636_0020_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 21:24:53,098 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local39874636_0020_m_000000_0
2018-07-14 21:24:53,098 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->385
2018-07-14 21:24:53,098 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 21:24:53,099 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:24:53,099 [pool-58-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 21:24:53,100 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 21:24:53,100 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 369 bytes
2018-07-14 21:24:53,102 [pool-58-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 385 bytes
to disk to satisfy reduce memory limit
2018-07-14 21:24:53,102 [pool-58-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 387 bytes from
disk
2018-07-14 21:24:53,102 [pool-58-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 21:24:53,102 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 21:24:53,103 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 375 bytes
2018-07-14 21:24:53,103 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:24:53,111 [pool-58-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 21:24:53,111 [pool-58-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 21:24:53,118 [pool-58-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans5[-1,-1],ans7[19,7],ans8[-1,-1]
2018-07-14 21:24:53,128 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local39874636_0020_r_000000_0 is done. And is in the process of committing
2018-07-14 21:24:53,129 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 21:24:53,129 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local39874636_0020_r_000000_0 is allowed to commit now
2018-07-14 21:24:53,134 [pool-58-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
```

```
'attempt_local39874636_0020_r_000000_0' to
file:/tmp/temp-1646866453/tmp105100876/_temporary/0/task_local39874636_0020_r_000000
2018-07-14 21:24:53,135 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 21:24:53,135 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local39874636_0020_r_000000_0' done.
2018-07-14 21:24:53,135 [pool-58-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local39874636_0020_r_000000_0
2018-07-14 21:24:53,135 [Thread-581] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 21:24:53,254 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local39874636_0020
2018-07-14 21:24:53,254 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans7,ans8,details
2018-07-14 21:24:53,254 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[11,10],details[-1,-1],ans5[16,7],ans5[16,7] C:  R:
ans5[16,7],ans5[-1,-1],ans7[19,7],ans8[-1,-1]
2018-07-14 21:24:53,266 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,269 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,272 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,295 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 21:24:53,300 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 21:24:52     2018-07-14 21:24:53
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local39874636_0020  2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans5,ans7,ans8,details  HASH_JOIN
file:/tmp/temp-1646866453/tmp105100876,
job_local530441329_0019 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 8 records in: "file:/tmp/temp-1646866453/tmp105100876"

Counters:
Total records written : 8
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
```

```
job_local530441329_0019 ->      job_local39874636_0020,
job_local39874636_0020


2018-07-14 21:24:53,310 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,312 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,314 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,318 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,319 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,321 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 21:24:53,325 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 21:24:53,325 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:24:53,326 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:24:53,327 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 21:24:53,336 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 21:24:53,336 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(103,Akshay)
(106,Aamir)
(107,Salman)
(108,Ranbir)
(109,Katrina)
(111,Tushar)
(112,Ajay)
(113,Jubeen)
grunt> ans8 = foreach ans7 generate EmpId,Name;^CYou have new mail in
/var/spool/mail/acadgild
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$
[acadgild.mmisra ~]$ pig -x local
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log
4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in
[jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/sl
f4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/07/14 21:35:59 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
18/07/14 21:35:59 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
2018-07-14 21:35:59,425 [main] INFO  org.apache.pig.Main - Apache Pig version 0.16.0
(r1746530) compiled Jun 01 2016, 23:10:49
2018-07-14 21:35:59,425 [main] INFO  org.apache.pig.Main - Logging error messages to:
/home/acadgild/pig_1531584359423.log
2018-07-14 21:35:59,451 [main] INFO  org.apache.pig.impl.util.Utils - Default bootup file
/home/acadgild/.pigbootup not found
2018-07-14 21:35:59,648 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2018-07-14 21:35:59,649 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 21:35:59,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop
file system at: file:///
2018-07-14 21:35:59,740 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:35:59,762 [main] INFO  org.apache.pig.PigServer - Pig Script ID for the
session: PIG-default-4efd996d-c7e4-42b3-b850-cc391024d60b
2018-07-14 21:35:59,762 [main] WARN  org.apache.pig.PigServer - ATS is disabled since
yarn.timeline-service.enabled set to false
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> clear


grunt>
grunt>
grunt>
grunt>
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int);
2018-07-14 21:36:10,912 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:36:10,913 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
(EmpId:Int,Expense:Int);
2018-07-14 21:36:11,081 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 21:36:11,081 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> sorted =ORDER details BY Salary DESC;
grunt> y = LIMIT sorted 5;
grunt>
grunt> dump y;
2018-07-14 22:03:16,325 [main] WARN  org.apache.hadoop.util.NativeCodeLoader - Unable to
load native-hadoop library for your platform... using builtin-java classes where applicable
2018-07-14 22:03:16,386 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: ORDER_BY,LIMIT
2018-07-14 22:03:16,434 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:16,435 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:03:16,497 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:03:16,589 [main] INFO  org.apache.pig.impl.util.SpillableMemoryManager -
Selected heap (PS Old Gen) of size 699400192 to monitor. collectionUsageThreshold =
489580128, usageThreshold = 489580128
2018-07-14 22:03:16,658 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:03:16,719 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.SecondaryKeyOptimizerMR -
Using Secondary Key Optimization for MapReduce node scope-20
```

```
2018-07-14 22:03:16,734 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 4
2018-07-14 22:03:16,734 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 4
2018-07-14 22:03:16,771 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:16,771 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:03:16,788 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
session.id is deprecated. Instead, use dfs.metrics.session-id
2018-07-14 22:03:16,789 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Initializing JVM Metrics with processName=JobTracker, sessionId=
2018-07-14 22:03:16,837 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:03:16,846 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:03:16,846 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:03:16,847 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.output.compress is deprecated. Instead, use
mapreduce.output.fileoutputformat.compress
2018-07-14 22:03:16,870 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:03:16,891 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:03:16,891 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:03:16,891 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531585996891-0
2018-07-14 22:03:16,941 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:03:16,942 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.tracker.http.address is deprecated. Instead, use
mapreduce.jobtracker.http.address
2018-07-14 22:03:16,967 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:17,000 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.task.id is deprecated. Instead,
use mapreduce.task.attempt.id
2018-07-14 22:03:17,477 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:03:17,514 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:03:17,524 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:03:17,525 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:03:17,553 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:03:17,685 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:03:17,854 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
```

```
Submitting tokens for job: job_local1837380255_0001
2018-07-14 22:03:18,092 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:03:18,094 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1837380255_0001
2018-07-14 22:03:18,094 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases details
2018-07-14 22:03:18,094 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[1,10],details[-1,-1] C:  R:
2018-07-14 22:03:18,118 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:03:18,126 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 0%
complete
2018-07-14 22:03:18,126 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1837380255_0001]
2018-07-14 22:03:18,155 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:03:18,155 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:03:18,155 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:18,155 [Thread-22] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:18,162 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:03:18,245 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:03:18,246 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1837380255_0001_m_000000_0
2018-07-14 22:03:18,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:18,344 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:03:18,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:03:18,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:03:18,404 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:03:18,407 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
```

```
generate code.
2018-07-14 22:03:18,435 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: details[1,10],details[-1,-1]
C:  R:
2018-07-14 22:03:18,452 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:03:18,453 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1837380255_0001_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:03:18,472 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:03:18,472 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1837380255_0001_m_000000_0 is allowed
to commit now
2018-07-14 22:03:18,475 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1837380255_0001_m_000000_0' to
file:/tmp/temp1426823468/tmp-1153408833/_temporary/0/task_local1837380255_0001_m_00000
0
2018-07-14 22:03:18,479 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:03:18,479 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1837380255_0001_m_000000_0' done.
2018-07-14 22:03:18,479 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1837380255_0001_m_000000_0
2018-07-14 22:03:18,480 [Thread-22] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:03:18,629 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 25%
complete
2018-07-14 22:03:18,632 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:18,637 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:18,638 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2018-07-14 22:03:18,638 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2018-07-14 22:03:18,639 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:18,709 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:03:18,710 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:03:18,712 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:03:18,714 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:03:18,717 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=280
2018-07-14 22:03:18,719 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:03:18,723 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Setting up single store job
2018-07-14 22:03:18,770 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:03:18,786 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:18,797 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:18,797 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:18,814 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:03:18,840 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:03:18,840 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:03:18,841 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:03:18,888 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:03:18,931 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1778678846_0002
2018-07-14 22:03:19,105 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:03:19,110 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:03:19,119 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:03:19,119 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:03:19,119 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:19,119 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:03:19,119 [Thread-47] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:19,119 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:03:19,125 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:03:19,125 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1778678846_0002_m_000000_0
2018-07-14 22:03:19,136 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:19,138 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 280
Input split[0]:
   Length = 280
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
```

```
   Locations:

-----------------------

2018-07-14 22:03:19,144 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1153408833/part-m-00000:0+280
2018-07-14 22:03:19,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:03:19,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:03:19,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:03:19,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:03:19,182 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:03:19,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:03:19,197 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:03:19,197 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:03:19,204 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:19,254 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:03:19,254 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:03:19,254 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:03:19,254 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 252; bufvoid = 104857600
2018-07-14 22:03:19,254 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:03:19,263 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:03:19,267 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1778678846_0002_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:03:19,270 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:03:19,270 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1778678846_0002_m_000000_0' done.
2018-07-14 22:03:19,270 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1778678846_0002_m_000000_0
2018-07-14 22:03:19,270 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:03:19,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1778678846_0002
2018-07-14 22:03:19,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases sorted
2018-07-14 22:03:19,271 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: sorted[3,8] C:  R:
2018-07-14 22:03:19,273 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
Waiting for reduce tasks
2018-07-14 22:03:19,278 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 37%
complete
2018-07-14 22:03:19,278 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1778678846_0002]
2018-07-14 22:03:19,278 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1778678846_0002_r_000000_0
2018-07-14 22:03:19,298 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:19,306 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5ea18c47
2018-07-14 22:03:19,331 [pool-5-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:03:19,349 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1778678846_0002_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:03:19,416 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#1 about to shuffle
output of map attempt_local1778678846_0002_m_000000_0 decomp: 282 len: 286 to MEMORY
2018-07-14 22:03:19,420 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 282 bytes from
map-output for attempt_local1778678846_0002_m_000000_0
2018-07-14 22:03:19,428 [localfetcher#1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 282, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->282
2018-07-14 22:03:19,431 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:03:19,432 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:19,432 [pool-5-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:03:19,442 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 22:03:19,442 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 268 bytes
2018-07-14 22:03:19,444 [pool-5-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 282 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:03:19,444 [pool-5-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 286 bytes from
disk
2018-07-14 22:03:19,445 [pool-5-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:03:19,445 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 22:03:19,446 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 268 bytes
2018-07-14 22:03:19,446 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:19,469 [pool-5-thread-1] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.skip.on is deprecated. Instead,
use mapreduce.job.skiprecords
2018-07-14 22:03:19,481 [pool-5-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 22:03:19,481 [pool-5-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:03:19,505 [pool-5-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:19,521 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1778678846_0002_r_000000_0 is done. And is in the process of committing
2018-07-14 22:03:19,527 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:19,528 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1778678846_0002_r_000000_0 is allowed to commit now
2018-07-14 22:03:19,536 [pool-5-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1778678846_0002_r_000000_0' to
file:/tmp/temp1426823468/tmp-277124160/_temporary/0/task_local1778678846_0002_r_000000
2018-07-14 22:03:19,539 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:03:19,541 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1778678846_0002_r_000000_0' done.
2018-07-14 22:03:19,541 [pool-5-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1778678846_0002_r_000000_0
2018-07-14 22:03:19,543 [Thread-47] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:03:19,709 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:03:19,712 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:19,713 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:19,714 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:19,729 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:03:19,729 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:03:19,729 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:03:19,729 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:03:19,737 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:03:19,771 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:03:19,773 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:19,779 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:19,779 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:19,792 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:03:19,814 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
```

```
2018-07-14 22:03:19,814 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:03:19,814 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:03:19,857 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:03:19,890 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1881632685_0003
2018-07-14 22:03:20,052 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:03:20,054 [Thread-76] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:03:20,065 [Thread-76] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:03:20,065 [Thread-76] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:03:20,065 [Thread-76] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:20,065 [Thread-76] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:03:20,065 [Thread-76] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:20,066 [Thread-76] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:03:20,072 [Thread-76] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:03:20,074 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1881632685_0003_m_000000_0
2018-07-14 22:03:20,090 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:20,097 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 280
Input split[0]:
   Length = 280
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:03:20,101 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1153408833/part-m-00000:0+280
2018-07-14 22:03:20,126 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:03:20,126 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:03:20,126 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:03:20,127 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:03:20,127 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
```

```
2018-07-14 22:03:20,139 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:03:20,141 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:03:20,144 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:03:20,150 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:20,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:03:20,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:03:20,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:03:20,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 308; bufvoid = 104857600
2018-07-14 22:03:20,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:03:20,189 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigCombiner$Combine -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:20,208 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:03:20,210 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1881632685_0003_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:03:20,214 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:03:20,214 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1881632685_0003_m_000000_0' done.
2018-07-14 22:03:20,215 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1881632685_0003_m_000000_0
2018-07-14 22:03:20,216 [Thread-76] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:03:20,217 [Thread-76] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:03:20,221 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1881632685_0003_r_000000_0
2018-07-14 22:03:20,237 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:20,237 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@42574ab0
2018-07-14 22:03:20,243 [pool-8-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:03:20,246 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1881632685_0003_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:03:20,251 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#2 about to shuffle
output of map attempt_local1881632685_0003_m_000000_0 decomp: 120 len: 124 to MEMORY
2018-07-14 22:03:20,256 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 120 bytes from
map-output for attempt_local1881632685_0003_m_000000_0
2018-07-14 22:03:20,256 [localfetcher#2] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
```

```
map-output of size: 120, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->120
2018-07-14 22:03:20,257 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:03:20,258 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:20,258 [pool-8-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:03:20,260 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 22:03:20,260 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 112 bytes
2018-07-14 22:03:20,263 [pool-8-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 120 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:03:20,263 [pool-8-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 124 bytes from
disk
2018-07-14 22:03:20,263 [pool-8-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:03:20,263 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Merger - Merging
1 sorted segments
2018-07-14 22:03:20,264 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down to
the last merge-pass, with 1 segments left of total size: 112 bytes
2018-07-14 22:03:20,264 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:20,272 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1881632685_0003
2018-07-14 22:03:20,272 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases sorted
2018-07-14 22:03:20,272 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: sorted[3,8] C:  R:
2018-07-14 22:03:20,277 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 62%
complete
2018-07-14 22:03:20,278 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Running jobs are [job_local1881632685_0003]
2018-07-14 22:03:20,291 [pool-8-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:03:20,291 [pool-8-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:03:20,297 [pool-8-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:20,301 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1881632685_0003_r_000000_0 is done. And is in the process of committing
2018-07-14 22:03:20,306 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:20,309 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1881632685_0003_r_000000_0 is allowed to commit now
2018-07-14 22:03:20,311 [pool-8-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1881632685_0003_r_000000_0' to
file:/tmp/temp1426823468/tmp1886958563/_temporary/0/task_local1881632685_0003_r_000000
2018-07-14 22:03:20,317 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
```

```
2018-07-14 22:03:20,318 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1881632685_0003_r_000000_0' done.
2018-07-14 22:03:20,320 [pool-8-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1881632685_0003_r_000000_0
2018-07-14 22:03:20,321 [Thread-76] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:03:20,457 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 75%
complete
2018-07-14 22:03:20,460 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:20,461 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:20,462 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:20,478 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:03:20,479 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:03:20,479 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:03:20,479 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:03:20,486 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:03:20,508 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:03:20,514 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:20,521 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:20,522 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:20,532 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:03:20,559 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:03:20,559 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:03:20,559 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:03:20,595 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:03:20,626 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local82626024_0004
2018-07-14 22:03:20,783 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:03:20,784 [Thread-105] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:03:20,791 [Thread-105] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
```

```
2018-07-14 22:03:20,791 [Thread-105] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:03:20,791 [Thread-105] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:20,791 [Thread-105] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:03:20,792 [Thread-105] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:03:20,792 [Thread-105] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:03:20,798 [Thread-105] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:03:20,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local82626024_0004_m_000000_0
2018-07-14 22:03:20,808 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:20,810 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 98
Input split[0]:
   Length = 98
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:03:20,814 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp1886958563/part-r-00000:0+98
2018-07-14 22:03:20,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:03:20,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:03:20,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:03:20,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:03:20,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:03:20,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:03:20,844 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:03:20,844 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:03:20,846 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:20,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:03:20,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:03:20,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
```

```
2018-07-14 22:03:20,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 108; bufvoid = 104857600
2018-07-14 22:03:20,847 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214380(104857520); length = 17/6553600
2018-07-14 22:03:20,849 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:03:20,851 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local82626024_0004_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:03:20,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:03:20,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local82626024_0004_m_000000_0' done.
2018-07-14 22:03:20,854 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local82626024_0004_m_000000_0
2018-07-14 22:03:20,854 [Thread-105] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:03:20,854 [Thread-105] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:03:20,859 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local82626024_0004_r_000000_0
2018-07-14 22:03:20,879 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:03:20,879 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@563b4a26
2018-07-14 22:03:20,880 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:03:20,882 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local82626024_0004_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:03:20,888 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#3 about to shuffle
output of map attempt_local82626024_0004_m_000000_0 decomp: 120 len: 124 to MEMORY
2018-07-14 22:03:20,889 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 120 bytes from
map-output for attempt_local82626024_0004_m_000000_0
2018-07-14 22:03:20,889 [localfetcher#3] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 120, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->120
2018-07-14 22:03:20,890 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:03:20,891 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:20,892 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:03:20,893 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:03:20,894 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 112 bytes
2018-07-14 22:03:20,896 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 120 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:03:20,896 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 124 bytes from
disk
2018-07-14 22:03:20,896 [pool-11-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:03:20,896 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:03:20,896 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 112 bytes
2018-07-14 22:03:20,897 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:20,929 [pool-11-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:03:20,930 [pool-11-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:03:20,933 [pool-11-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:03:20,935 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local82626024_0004_r_000000_0 is done. And is in the process of committing
2018-07-14 22:03:20,939 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:03:20,939 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local82626024_0004_r_000000_0 is allowed to commit now
2018-07-14 22:03:20,942 [pool-11-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local82626024_0004_r_000000_0' to
file:/tmp/temp1426823468/tmp1857435674/_temporary/0/task_local82626024_0004_r_000000
2018-07-14 22:03:20,943 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:03:20,943 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local82626024_0004_r_000000_0' done.
2018-07-14 22:03:20,943 [pool-11-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local82626024_0004_r_000000_0
2018-07-14 22:03:20,944 [Thread-105] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:03:21,010 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local82626024_0004
2018-07-14 22:03:21,010 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases sorted
2018-07-14 22:03:21,010 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: sorted[3,8] C:  R:
2018-07-14 22:03:21,014 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,015 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,016 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,039 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:03:21,049 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:03:16     2018-07-14 22:03:21
ORDER_BY,LIMIT

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
```

```
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1778678846_0002        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     sorted  SAMPLER
job_local1837380255_0001        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details MAP_ONLY
job_local1881632685_0003        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     sorted  ORDER_BY,COMBINER
job_local82626024_0004  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     sorted          file:/tmp/temp1426823468/tmp1857435674,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 5 records in: "file:/tmp/temp1426823468/tmp1857435674"

Counters:
Total records written : 5
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1837380255_0001        ->      job_local1778678846_0002,
job_local1778678846_0002        ->      job_local1881632685_0003,
job_local1881632685_0003        ->      job_local82626024_0004,
job_local82626024_0004


2018-07-14 22:03:21,051 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,054 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,055 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,066 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,070 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,071 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,080 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,081 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,085 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,096 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,106 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,107 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:03:21,114 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:03:21,118 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:03:21,119 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:03:21,119 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
```

```
SchemaTupleBackend has already been initialized
2018-07-14 22:03:21,132 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:03:21,132 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(106,Aamir,25000,1)
(101,Amitabh,20000,1)
(107,Salman,17500,2)
(108,Ranbir,14000,3)
(103,Akshay,11000,3)
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> filter_sorted= FILTER sorted BY ((EmpId%2) != 0);
grunt> ans2 =LIMIT filter_sorted 3;
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> dump ans2;
2018-07-14 22:06:41,082 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: ORDER_BY,FILTER,LIMIT
2018-07-14 22:06:41,108 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:41,109 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:06:41,110 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:06:41,110 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:06:41,118 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:06:41,121 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.SecondaryKeyOptimizerMR -
Using Secondary Key Optimization for MapReduce node scope-83
2018-07-14 22:06:41,122 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 4
2018-07-14 22:06:41,122 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 4
2018-07-14 22:06:41,134 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:41,134 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:06:41,135 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:41,137 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:06:41,137 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:06:41,140 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:06:41,141 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:06:41,141 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:06:41,141 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531586201141-0
2018-07-14 22:06:41,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:06:41,153 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:41,164 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:06:41,167 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:06:41,167 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:06:41,167 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:06:41,168 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:06:41,198 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:06:41,220 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1822761447_0005
2018-07-14 22:06:41,343 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:06:41,343 [Thread-134] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:06:41,347 [Thread-134] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:06:41,347 [Thread-134] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:06:41,347 [Thread-134] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:41,347 [Thread-134] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 22:06:41,348 [Thread-134] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:06:41,355 [Thread-134] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:06:41,356 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1822761447_0005_m_000000_0
2018-07-14 22:06:41,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:41,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:06:41,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:06:41,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:06:41,380 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:41,381 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:06:41,386 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M:
details[1,10],details[-1,-1],filter_sorted[5,15] C:  R:
2018-07-14 22:06:41,395 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:06:41,395 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1822761447_0005_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:06:41,400 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:06:41,400 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1822761447_0005_m_000000_0 is allowed
to commit now
2018-07-14 22:06:41,402 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1822761447_0005_m_000000_0' to
file:/tmp/temp1426823468/tmp903791605/_temporary/0/task_local1822761447_0005_m_000000
2018-07-14 22:06:41,403 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:06:41,403 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1822761447_0005_m_000000_0' done.
2018-07-14 22:06:41,403 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1822761447_0005_m_000000_0
2018-07-14 22:06:41,403 [Thread-134] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:06:41,544 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1822761447_0005
2018-07-14 22:06:41,544 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
Processing aliases details,filter_sorted
2018-07-14 22:06:41,544 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[1,10],details[-1,-1],filter_sorted[5,15] C:  R:
2018-07-14 22:06:41,545 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 25%
complete
2018-07-14 22:06:41,546 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:41,547 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:41,547 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:41,551 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:06:41,552 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:06:41,553 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:06:41,553 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:06:41,555 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=138
2018-07-14 22:06:41,555 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:06:41,556 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:06:41,567 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:06:41,569 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:41,572 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:41,572 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:06:41,579 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:06:41,590 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:06:41,590 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:06:41,590 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:06:41,623 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:06:41,646 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local273376727_0006
2018-07-14 22:06:41,780 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
```

```
2018-07-14 22:06:41,781 [Thread-157] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:06:41,784 [Thread-157] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:06:41,785 [Thread-157] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:06:41,785 [Thread-157] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:41,785 [Thread-157] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:06:41,785 [Thread-157] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:06:41,785 [Thread-157] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:06:41,790 [Thread-157] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:06:41,790 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local273376727_0006_m_000000_0
2018-07-14 22:06:41,797 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:41,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 138
Input split[0]:
   Length = 138
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:06:41,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp903791605/part-m-00000:0+138
2018-07-14 22:06:41,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:06:41,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:06:41,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:06:41,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:06:41,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:06:41,835 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:06:41,837 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:41,837 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:06:41,839 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:06:41,841 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:06:41,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:06:41,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:06:41,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 126; bufvoid = 104857600
2018-07-14 22:06:41,842 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:06:41,843 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:06:41,845 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local273376727_0006_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:06:41,846 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:06:41,846 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local273376727_0006_m_000000_0' done.
2018-07-14 22:06:41,846 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local273376727_0006_m_000000_0
2018-07-14 22:06:41,846 [Thread-157] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:06:41,846 [Thread-157] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:06:41,849 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local273376727_0006_r_000000_0
2018-07-14 22:06:41,857 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:41,857 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6a6f1861
2018-07-14 22:06:41,858 [pool-16-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:06:41,859 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local273376727_0006_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:06:41,864 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#4 about to shuffle
output of map attempt_local273376727_0006_m_000000_0 decomp: 142 len: 146 to MEMORY
2018-07-14 22:06:41,864 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 142 bytes from
map-output for attempt_local273376727_0006_m_000000_0
2018-07-14 22:06:41,864 [localfetcher#4] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 142, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->142
2018-07-14 22:06:41,866 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:06:41,867 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:41,867 [pool-16-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:06:41,868 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:06:41,868 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 128 bytes
2018-07-14 22:06:41,870 [pool-16-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 142 bytes
```

```
to disk to satisfy reduce memory limit
2018-07-14 22:06:41,871 [pool-16-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 146 bytes from
disk
2018-07-14 22:06:41,871 [pool-16-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:06:41,871 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:06:41,871 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 128 bytes
2018-07-14 22:06:41,871 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:41,883 [pool-16-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:41,883 [pool-16-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:06:41,891 [pool-16-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:06:41,893 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local273376727_0006_r_000000_0 is done. And is in the process of committing
2018-07-14 22:06:41,896 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:41,896 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local273376727_0006_r_000000_0 is allowed to commit now
2018-07-14 22:06:41,898 [pool-16-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local273376727_0006_r_000000_0' to
file:/tmp/temp1426823468/tmp-961620584/_temporary/0/task_local273376727_0006_r_000000
2018-07-14 22:06:41,899 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:06:41,899 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local273376727_0006_r_000000_0' done.
2018-07-14 22:06:41,899 [pool-16-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local273376727_0006_r_000000_0
2018-07-14 22:06:41,899 [Thread-157] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:06:42,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local273376727_0006
2018-07-14 22:06:42,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases sorted
2018-07-14 22:06:42,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: sorted[3,8] C:  R:
2018-07-14 22:06:42,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:06:42,070 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,071 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,072 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,078 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:06:42,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
```

```
2018-07-14 22:06:42,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:06:42,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:06:42,082 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:06:42,094 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:06:42,095 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,099 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:42,099 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:06:42,109 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:06:42,121 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:06:42,121 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:06:42,121 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:06:42,150 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:06:42,172 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1671689007_0007
2018-07-14 22:06:42,287 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:06:42,287 [Thread-186] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:06:42,291 [Thread-186] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:06:42,291 [Thread-186] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:06:42,291 [Thread-186] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:42,291 [Thread-186] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:06:42,292 [Thread-186] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:06:42,292 [Thread-186] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:06:42,295 [Thread-186] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:06:42,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1671689007_0007_m_000000_0
```

```
2018-07-14 22:06:42,301 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:42,303 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 138
Input split[0]:
   Length = 138
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:06:42,306 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp903791605/part-m-00000:0+138
2018-07-14 22:06:42,330 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:06:42,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:06:42,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:06:42,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:06:42,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:06:42,337 [Service Thread] INFO
org.apache.pig.impl.util.SpillableMemoryManager - first memory handler call- Usage
threshold init = 88080384(86016K) used = 536873520(524290K) committed = 576192512(562688K)
max = 699400192(683008K), toFree = 292083456
2018-07-14 22:06:42,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:06:42,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:42,343 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:06:42,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:06:42,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:06:42,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:06:42,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:06:42,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 152; bufvoid = 104857600
2018-07-14 22:06:42,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:06:42,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:06:42,357 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1671689007_0007_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:06:42,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:06:42,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1671689007_0007_m_000000_0' done.
2018-07-14 22:06:42,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1671689007_0007_m_000000_0
```

```
2018-07-14 22:06:42,359 [Thread-186] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:06:42,359 [Thread-186] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:06:42,363 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1671689007_0007_r_000000_0
2018-07-14 22:06:42,371 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:42,372 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@39891b28
2018-07-14 22:06:42,374 [pool-19-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:06:42,376 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1671689007_0007_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:06:42,377 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#5 about to shuffle
output of map attempt_local1671689007_0007_m_000000_0 decomp: 74 len: 78 to MEMORY
2018-07-14 22:06:42,378 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 74 bytes from map-output
for attempt_local1671689007_0007_m_000000_0
2018-07-14 22:06:42,378 [localfetcher#5] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 74, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->74
2018-07-14 22:06:42,379 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:06:42,379 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:42,379 [pool-19-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:06:42,381 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:06:42,381 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 66 bytes
2018-07-14 22:06:42,382 [pool-19-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 74 bytes to
disk to satisfy reduce memory limit
2018-07-14 22:06:42,382 [pool-19-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 78 bytes from
disk
2018-07-14 22:06:42,382 [pool-19-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:06:42,382 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:06:42,383 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 66 bytes
2018-07-14 22:06:42,383 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:42,393 [pool-19-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:42,393 [pool-19-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:06:42,395 [pool-19-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:06:42,396 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Task -
```

```
Task:attempt_local1671689007_0007_r_000000_0 is done. And is in the process of committing
2018-07-14 22:06:42,400 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:42,400 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1671689007_0007_r_000000_0 is allowed to commit now
2018-07-14 22:06:42,401 [pool-19-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1671689007_0007_r_000000_0' to
file:/tmp/temp1426823468/tmp-1575608734/_temporary/0/task_local1671689007_0007_r_00000
0
2018-07-14 22:06:42,404 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:06:42,404 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1671689007_0007_r_000000_0' done.
2018-07-14 22:06:42,404 [pool-19-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1671689007_0007_r_000000_0
2018-07-14 22:06:42,404 [Thread-186] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:06:42,588 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1671689007_0007
2018-07-14 22:06:42,589 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases sorted
2018-07-14 22:06:42,589 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: sorted[3,8] C:  R:
2018-07-14 22:06:42,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 75%
complete
2018-07-14 22:06:42,607 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,611 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,614 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,631 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:06:42,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:06:42,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:06:42,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:06:42,635 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:06:42,649 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:06:42,651 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:42,657 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:42,657 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:06:42,664 [JobControl] WARN
```

```
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:06:42,672 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:06:42,672 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:06:42,672 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:06:42,701 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:06:42,719 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local299712536_0008
2018-07-14 22:06:42,821 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:06:42,821 [Thread-215] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:06:42,825 [Thread-215] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:06:42,825 [Thread-215] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:06:42,825 [Thread-215] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:42,825 [Thread-215] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:06:42,825 [Thread-215] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:06:42,825 [Thread-215] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:06:42,829 [Thread-215] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:06:42,829 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local299712536_0008_m_000000_0
2018-07-14 22:06:42,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:42,836 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 60
Input split[0]:
   Length = 60
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:06:42,838 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1575608734/part-r-00000:0+60
2018-07-14 22:06:42,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:06:42,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:06:42,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
```

```
2018-07-14 22:06:42,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:06:42,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:06:42,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:06:42,858 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:42,858 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:06:42,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:06:42,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:06:42,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:06:42,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:06:42,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 66; bufvoid = 104857600
2018-07-14 22:06:42,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214388(104857552); length = 9/6553600
2018-07-14 22:06:42,863 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:06:42,865 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local299712536_0008_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:06:42,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:06:42,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local299712536_0008_m_000000_0' done.
2018-07-14 22:06:42,866 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local299712536_0008_m_000000_0
2018-07-14 22:06:42,866 [Thread-215] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:06:42,866 [Thread-215] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:06:42,870 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local299712536_0008_r_000000_0
2018-07-14 22:06:42,877 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:06:42,877 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@752e4104
2018-07-14 22:06:42,878 [pool-22-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:06:42,879 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local299712536_0008_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:06:42,881 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#6 about to shuffle
output of map attempt_local299712536_0008_m_000000_0 decomp: 74 len: 78 to MEMORY
2018-07-14 22:06:42,881 [localfetcher#6] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 74 bytes from map-output
for attempt_local299712536_0008_m_000000_0
2018-07-14 22:06:42,881 [localfetcher#6] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 74, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->74
2018-07-14 22:06:42,882 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:06:42,883 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:42,883 [pool-22-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:06:42,884 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:06:42,884 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 66 bytes
2018-07-14 22:06:42,885 [pool-22-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 74 bytes to
disk to satisfy reduce memory limit
2018-07-14 22:06:42,885 [pool-22-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 78 bytes from
disk
2018-07-14 22:06:42,885 [pool-22-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:06:42,885 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:06:42,886 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 66 bytes
2018-07-14 22:06:42,886 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:42,896 [pool-22-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:06:42,896 [pool-22-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:06:42,900 [pool-22-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: sorted[3,8] C:  R:
2018-07-14 22:06:42,901 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local299712536_0008_r_000000_0 is done. And is in the process of committing
2018-07-14 22:06:42,904 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:06:42,904 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local299712536_0008_r_000000_0 is allowed to commit now
2018-07-14 22:06:42,906 [pool-22-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local299712536_0008_r_000000_0' to
file:/tmp/temp1426823468/tmp1345235158/_temporary/0/task_local299712536_0008_r_000000
2018-07-14 22:06:42,907 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:06:42,907 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local299712536_0008_r_000000_0' done.
2018-07-14 22:06:42,907 [pool-22-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local299712536_0008_r_000000_0
2018-07-14 22:06:42,907 [Thread-215] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:06:43,025 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local299712536_0008
2018-07-14 22:06:43,025 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases sorted
2018-07-14 22:06:43,025 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
detailed locations: M: sorted[3,8] C:  R:
2018-07-14 22:06:43,046 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,052 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,060 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,074 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:06:43,083 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:06:41     2018-07-14 22:06:43
ORDER_BY,FILTER,LIMIT

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1671689007_0007        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     sorted  ORDER_BY,COMBINER
job_local1822761447_0005        1       0       n/a     n/a     n/a     n/a     0
0       0       0       details,filter_sorted   MAP_ONLY
job_local273376727_0006 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     sorted  SAMPLER
job_local299712536_0008 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     sorted          file:/tmp/temp1426823468/tmp1345235158,

Input(s):
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 3 records in: "file:/tmp/temp1426823468/tmp1345235158"

Counters:
Total records written : 3
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1822761447_0005        ->      job_local273376727_0006,
job_local273376727_0006 ->      job_local1671689007_0007,
job_local1671689007_0007        ->      job_local299712536_0008,
job_local299712536_0008


2018-07-14 22:06:43,087 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,088 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,093 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,096 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,097 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,098 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,101 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,101 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,102 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,106 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,107 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,111 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:06:43,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:06:43,116 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:06:43,117 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:06:43,117 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:06:43,135 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:06:43,135 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1)
(107,Salman,17500,2)
(103,Akshay,11000,3)
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
>> (EmpId:Int,Expense:Int);
2018-07-14 22:09:01,928 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:09:01,928 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> expenses = load 'employee_expenses.txt' using PigStorage(' ') as
>> (EmpId:Int,Expense:Int);
2018-07-14 22:09:07,351 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:09:07,351 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt> ex_g = group expenses BY EmpId;
grunt>
grunt>
grunt>
grunt>
grunt> dump ex_g;
2018-07-14 22:10:56,160 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 22:10:56,193 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:10:56,193 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:10:56,193 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:10:56,193 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:10:56,215 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:10:56,216 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 22:10:56,216 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 22:10:56,229 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:10:56,230 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:10:56,230 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,231 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:10:56,231 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:10:56,232 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:10:56,232 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:10:56,232 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:10:56,232 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:10:56,233 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:10:56,234 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:10:56,234 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:10:56,234 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531586456233-0
2018-07-14 22:10:56,245 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:10:56,247 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,257 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
```

```
2018-07-14 22:10:56,264 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:10:56,264 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:10:56,264 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:10:56,265 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:10:56,295 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:10:56,316 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1847510987_0009
2018-07-14 22:10:56,426 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:10:56,427 [Thread-244] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:10:56,433 [Thread-244] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:10:56,433 [Thread-244] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:10:56,433 [Thread-244] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:10:56,433 [Thread-244] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:10:56,433 [Thread-244] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:10:56,433 [Thread-244] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:10:56,438 [Thread-244] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:10:56,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1847510987_0009_m_000000_0
2018-07-14 22:10:56,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:10:56,445 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:10:56,451 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:10:56,451 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:10:56,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:10:56,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 22:10:56,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:10:56,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:10:56,485 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:10:56,486 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:10:56,490 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:10:56,490 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:10:56,494 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_g[11,7] C:  R:
2018-07-14 22:10:56,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:10:56,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:10:56,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:10:56,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 22:10:56,496 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:10:56,499 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:10:56,501 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1847510987_0009_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:10:56,502 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:10:56,502 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1847510987_0009_m_000000_0' done.
2018-07-14 22:10:56,502 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1847510987_0009_m_000000_0
2018-07-14 22:10:56,503 [Thread-244] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:10:56,503 [Thread-244] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:10:56,506 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1847510987_0009_r_000000_0
2018-07-14 22:10:56,521 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:10:56,521 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@27c45497
2018-07-14 22:10:56,522 [pool-25-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:10:56,524 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1847510987_0009_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:10:56,526 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#7 about to shuffle
output of map attempt_local1847510987_0009_m_000000_0 decomp: 125 len: 129 to MEMORY
```

```
2018-07-14 22:10:56,526 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local1847510987_0009_m_000000_0
2018-07-14 22:10:56,526 [localfetcher#7] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 22:10:56,528 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:10:56,620 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:10:56,620 [pool-25-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:10:56,620 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:10:56,620 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 22:10:56,624 [pool-25-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:10:56,624 [pool-25-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 22:10:56,624 [pool-25-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:10:56,624 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:10:56,624 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 22:10:56,624 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:10:56,634 [pool-25-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:10:56,634 [pool-25-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:10:56,638 [pool-25-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_g[11,7] C:  R:
2018-07-14 22:10:56,640 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1847510987_0009_r_000000_0 is done. And is in the process of committing
2018-07-14 22:10:56,644 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:10:56,644 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1847510987_0009_r_000000_0 is allowed to commit now
2018-07-14 22:10:56,648 [pool-25-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1847510987_0009_r_000000_0' to
file:/tmp/temp1426823468/tmp850178937/_temporary/0/task_local1847510987_0009_r_000000
2018-07-14 22:10:56,649 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:10:56,649 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1847510987_0009_r_000000_0' done.
2018-07-14 22:10:56,649 [pool-25-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1847510987_0009_r_000000_0
2018-07-14 22:10:56,649 [Thread-244] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:10:56,746 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1847510987_0009
```

```
2018-07-14 22:10:56,746 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,expenses
2018-07-14 22:10:56,746 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_g[11,7] C:  R:
2018-07-14 22:10:56,747 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,748 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,748 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,754 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:10:56,754 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:10:56     2018-07-14 22:10:56
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1847510987_0009        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,expenses   GROUP_BY
file:/tmp/temp1426823468/tmp850178937,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp1426823468/tmp850178937"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1847510987_0009


2018-07-14 22:10:56,755 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,756 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,756 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:10:56,760 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:10:56,760 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:10:56,760 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:10:56,760 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
```

```
SchemaTupleBackend has already been initialized
2018-07-14 22:10:56,770 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:10:56,770 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt>
grunt>
grunt>
grunt>
grunt> ex_sum= foreach ex_g GENERATE group as EmpId,SUM(expenses.Expense) as sum;
grunt>
grunt> dump ex_sum;
2018-07-14 22:13:05,051 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 22:13:05,072 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:13:05,072 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:13:05,072 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:13:05,073 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:13:05,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:13:05,079 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:13:05,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 22:13:05,089 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 22:13:05,113 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:13:05,114 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:13:05,114 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,115 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:13:05,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:13:05,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:13:05,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
```

```
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:13:05,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:13:05,116 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:13:05,118 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:13:05,119 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:13:05,119 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:13:05,119 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531586585119-0
2018-07-14 22:13:05,133 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:13:05,134 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,145 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:13:05,150 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:13:05,150 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:13:05,150 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:13:05,151 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:13:05,174 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:13:05,197 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local349891731_0010
2018-07-14 22:13:05,293 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:13:05,294 [Thread-273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:13:05,298 [Thread-273] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:13:05,298 [Thread-273] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:13:05,298 [Thread-273] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:13:05,298 [Thread-273] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:13:05,298 [Thread-273] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:13:05,298 [Thread-273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
```

```
2018-07-14 22:13:05,302 [Thread-273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:13:05,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local349891731_0010_m_000000_0
2018-07-14 22:13:05,312 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:13:05,314 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:13:05,319 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:13:05,319 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:13:05,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:13:05,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:13:05,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:13:05,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:13:05,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:13:05,340 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:13:05,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:13:05,341 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:13:05,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:13:05,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:13:05,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:13:05,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:13:05,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:13:05,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:13:05,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:13:05,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local349891731_0010_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:13:05,372 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:13:05,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local349891731_0010_m_000000_0' done.
2018-07-14 22:13:05,372 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local349891731_0010_m_000000_0
2018-07-14 22:13:05,372 [Thread-273] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:13:05,372 [Thread-273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:13:05,377 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local349891731_0010_r_000000_0
2018-07-14 22:13:05,387 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:13:05,388 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@305806f
2018-07-14 22:13:05,389 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:13:05,390 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local349891731_0010_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:13:05,392 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#8 about to shuffle
output of map attempt_local349891731_0010_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:13:05,392 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local349891731_0010_m_000000_0
2018-07-14 22:13:05,392 [localfetcher#8] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:13:05,394 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:13:05,395 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:13:05,395 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:13:05,396 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:13:05,396 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:13:05,398 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:13:05,398 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:13:05,398 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:13:05,398 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:13:05,398 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:13:05,398 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:13:05,406 [pool-28-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 22:13:05,406 [pool-28-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:13:05,414 [pool-28-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:13:05,417 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local349891731_0010_r_000000_0 is done. And is in the process of committing
2018-07-14 22:13:05,419 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:13:05,419 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local349891731_0010_r_000000_0 is allowed to commit now
2018-07-14 22:13:05,423 [pool-28-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local349891731_0010_r_000000_0' to
file:/tmp/temp1426823468/tmp-1088343716/_temporary/0/task_local349891731_0010_r_000000
2018-07-14 22:13:05,425 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:13:05,425 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local349891731_0010_r_000000_0' done.
2018-07-14 22:13:05,425 [pool-28-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local349891731_0010_r_000000_0
2018-07-14 22:13:05,426 [Thread-273] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:13:05,597 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local349891731_0010
2018-07-14 22:13:05,597 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:13:05,597 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C:
ex_sum[12,8],ex_g[11,7] R: ex_sum[12,8]
2018-07-14 22:13:05,610 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,616 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,621 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,653 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:13:05,655 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:13:05     2018-07-14 22:13:05
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local349891731_0010 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp1426823468/tmp-1088343716,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
```

```
Output(s):
Successfully stored 7 records in: "file:/tmp/temp1426823468/tmp-1088343716"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local349891731_0010


2018-07-14 22:13:05,656 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,657 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,657 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:13:05,661 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:13:05,662 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:13:05,663 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:13:05,663 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:13:05,674 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:13:05,674 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt>
grunt>
grunt>
grunt>
grunt> ex_m = group  ex_sum all;
grunt>
grunt> dump ex_m;
2018-07-14 22:14:56,445 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 22:14:56,471 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:14:56,471 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:14:56,471 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:14:56,471 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
```

```
2018-07-14 22:14:56,479 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:14:56,481 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:14:56,484 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:14:56,484 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:14:56,495 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:14:56,496 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:14:56,496 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:56,497 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:14:56,497 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:14:56,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:14:56,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:14:56,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:14:56,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:14:56,499 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:14:56,499 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:14:56,499 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:14:56,499 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531586696499-0
2018-07-14 22:14:56,510 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:14:56,511 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:56,521 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:14:56,525 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:14:56,525 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:14:56,525 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:14:56,525 [JobControl] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:14:56,552 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:14:56,569 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1771646364_0011
2018-07-14 22:14:56,660 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:14:56,661 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:14:56,664 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:14:56,664 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:14:56,664 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:14:56,664 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:14:56,664 [Thread-302] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:14:56,664 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:14:56,666 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:14:56,666 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1771646364_0011_m_000000_0
2018-07-14 22:14:56,672 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:14:56,673 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:14:56,675 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:14:56,675 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:14:56,691 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:14:56,691 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:14:56,691 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:14:56,691 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:14:56,691 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:14:56,692 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
```

```
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:14:56,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:14:56,693 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:14:56,696 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:14:56,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:14:56,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:14:56,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:14:56,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:14:56,698 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:14:56,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:14:56,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1771646364_0011_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:14:56,708 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:14:56,708 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1771646364_0011_m_000000_0' done.
2018-07-14 22:14:56,708 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1771646364_0011_m_000000_0
2018-07-14 22:14:56,708 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:14:56,708 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:14:56,711 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1771646364_0011_r_000000_0
2018-07-14 22:14:56,724 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:14:56,724 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@400f4de5
2018-07-14 22:14:56,725 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:14:56,730 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1771646364_0011_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:14:56,731 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#9 about to shuffle
output of map attempt_local1771646364_0011_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:14:56,731 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1771646364_0011_m_000000_0
2018-07-14 22:14:56,731 [localfetcher#9] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:14:56,732 [EventFetcher for fetching Map Completion Events] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:14:56,732 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:14:56,732 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:14:56,733 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:14:56,733 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:14:56,736 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:14:56,736 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:14:56,736 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:14:56,736 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:14:56,736 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:14:56,736 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:14:56,747 [pool-31-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:14:56,747 [pool-31-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:14:56,754 [pool-31-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:14:56,760 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1771646364_0011_r_000000_0 is done. And is in the process of committing
2018-07-14 22:14:56,763 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:14:56,763 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1771646364_0011_r_000000_0 is allowed to commit now
2018-07-14 22:14:56,765 [pool-31-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1771646364_0011_r_000000_0' to
file:/tmp/temp1426823468/tmp-1753229361/_temporary/0/task_local1771646364_0011_r_00000
0
2018-07-14 22:14:56,766 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:14:56,766 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1771646364_0011_r_000000_0' done.
2018-07-14 22:14:56,766 [pool-31-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1771646364_0011_r_000000_0
2018-07-14 22:14:56,769 [Thread-302] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:14:56,963 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1771646364_0011
2018-07-14 22:14:56,964 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:14:56,964 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C:
ex_sum[12,8],ex_g[11,7] R: ex_sum[12,8]
2018-07-14 22:14:56,969 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:14:56,977 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:56,981 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:56,985 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:56,997 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:14:56,997 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:14:56,998 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:14:56,998 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:14:57,000 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:14:57,018 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:14:57,019 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,027 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:14:57,027 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:14:57,034 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:14:57,045 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:14:57,045 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:14:57,046 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:14:57,073 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:14:57,089 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local533036808_0012
2018-07-14 22:14:57,190 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:14:57,190 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:14:57,194 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:14:57,195 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
```

```
2018-07-14 22:14:57,195 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:14:57,195 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:14:57,195 [Thread-331] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:14:57,195 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:14:57,198 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:14:57,198 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local533036808_0012_m_000000_0
2018-07-14 22:14:57,203 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:14:57,204 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:14:57,206 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1753229361/part-r-00000:0+62
2018-07-14 22:14:57,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:14:57,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:14:57,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:14:57,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:14:57,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:14:57,225 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:14:57,226 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:14:57,226 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:14:57,227 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[13,7] C:  R:
2018-07-14 22:14:57,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:14:57,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:14:57,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:14:57,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 97; bufvoid = 104857600
2018-07-14 22:14:57,228 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
```

```
26214372(104857488); length = 25/6553600
2018-07-14 22:14:57,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:14:57,231 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local533036808_0012_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:14:57,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:14:57,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local533036808_0012_m_000000_0' done.
2018-07-14 22:14:57,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local533036808_0012_m_000000_0
2018-07-14 22:14:57,233 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:14:57,233 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:14:57,237 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local533036808_0012_r_000000_0
2018-07-14 22:14:57,245 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:14:57,246 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4218a024
2018-07-14 22:14:57,247 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:14:57,248 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local533036808_0012_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:14:57,252 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#10 about to shuffle
output of map attempt_local533036808_0012_m_000000_0 decomp: 113 len: 117 to MEMORY
2018-07-14 22:14:57,253 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 113 bytes from
map-output for attempt_local533036808_0012_m_000000_0
2018-07-14 22:14:57,253 [localfetcher#10] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 113, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->113
2018-07-14 22:14:57,253 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:14:57,254 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:14:57,254 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:14:57,255 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:14:57,255 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 22:14:57,257 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 113 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:14:57,257 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 117 bytes from
disk
2018-07-14 22:14:57,257 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:14:57,257 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
```

```
2018-07-14 22:14:57,258 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 105 bytes
2018-07-14 22:14:57,259 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:14:57,267 [pool-34-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:14:57,267 [pool-34-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:14:57,268 [pool-34-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_m[13,7] C:  R:
2018-07-14 22:14:57,268 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local533036808_0012_r_000000_0 is done. And is in the process of committing
2018-07-14 22:14:57,271 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:14:57,271 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local533036808_0012_r_000000_0 is allowed to commit now
2018-07-14 22:14:57,272 [pool-34-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local533036808_0012_r_000000_0' to
file:/tmp/temp1426823468/tmp1967489388/_temporary/0/task_local533036808_0012_r_000000
2018-07-14 22:14:57,274 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:14:57,274 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local533036808_0012_r_000000_0' done.
2018-07-14 22:14:57,274 [pool-34-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local533036808_0012_r_000000_0
2018-07-14 22:14:57,274 [Thread-331] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:14:57,393 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local533036808_0012
2018-07-14 22:14:57,393 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m
2018-07-14 22:14:57,393 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_m[13,7] C:  R:
2018-07-14 22:14:57,395 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,396 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,397 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,400 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:14:57,401 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:14:56     2018-07-14 22:14:57
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1771646364_0011        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
```

```
job_local533036808_0012 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m    GROUP_BY        file:/tmp/temp1426823468/tmp1967489388,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp1426823468/tmp1967489388"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1771646364_0011        ->      job_local533036808_0012,
job_local533036808_0012


2018-07-14 22:14:57,401 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,402 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,403 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,405 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,407 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,407 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:14:57,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:14:57,409 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:14:57,410 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:14:57,410 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:14:57,417 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:14:57,417 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,{(119,200),(114,200),(110,400),(105,100),(104,300),(102,500),(101,300)})
grunt>
grunt>
grunt>
grunt>
grunt> ex_max = foreach ex_m generate group,MAX(ex_sum.sum) as max;
grunt> dump ex_max;
2018-07-14 22:16:31,971 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 22:16:31,993 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:16:31,994 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:16:31,994 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:16:31,994 [main] INFO
```

```
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:16:32,012 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:16:32,013 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:16:32,014 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:16:32,015 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:16:32,015 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:16:32,026 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:16:32,027 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:16:32,027 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:32,029 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:16:32,029 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:16:32,029 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:16:32,029 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:16:32,030 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:16:32,030 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:16:32,031 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:16:32,031 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:16:32,031 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:16:32,031 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531586792031-0
2018-07-14 22:16:32,040 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:16:32,042 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:32,058 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
```

```
2018-07-14 22:16:32,060 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:16:32,060 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:16:32,060 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:16:32,061 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:16:32,083 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:16:32,099 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2081496176_0013
2018-07-14 22:16:32,181 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:16:32,181 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:16:32,187 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:16:32,187 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:16:32,187 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:16:32,188 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:16:32,188 [Thread-360] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:16:32,188 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:16:32,190 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:16:32,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2081496176_0013_m_000000_0
2018-07-14 22:16:32,194 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:16:32,195 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:16:32,197 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:16:32,197 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:16:32,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:16:32,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
```

```
2018-07-14 22:16:32,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:16:32,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:16:32,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:16:32,213 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:16:32,214 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:16:32,215 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:16:32,221 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:16:32,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:16:32,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:16:32,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:16:32,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:16:32,224 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:16:32,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:16:32,232 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2081496176_0013_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:16:32,233 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:16:32,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2081496176_0013_m_000000_0' done.
2018-07-14 22:16:32,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2081496176_0013_m_000000_0
2018-07-14 22:16:32,234 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:16:32,234 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:16:32,234 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local2081496176_0013_r_000000_0
2018-07-14 22:16:32,240 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:16:32,240 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@1b7f66c7
2018-07-14 22:16:32,242 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:16:32,242 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2081496176_0013_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:16:32,244 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#11 about to shuffle
```

```
output of map attempt_local2081496176_0013_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:16:32,244 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local2081496176_0013_m_000000_0
2018-07-14 22:16:32,244 [localfetcher#11] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:16:32,247 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:16:32,248 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:16:32,248 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:16:32,250 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:16:32,250 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:16:32,253 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:16:32,253 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:16:32,253 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:16:32,254 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:16:32,254 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:16:32,255 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:16:32,262 [pool-37-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:16:32,262 [pool-37-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:16:32,264 [pool-37-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:16:32,265 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2081496176_0013_r_000000_0 is done. And is in the process of committing
2018-07-14 22:16:32,267 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:16:32,267 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local2081496176_0013_r_000000_0 is allowed to commit now
2018-07-14 22:16:32,268 [pool-37-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2081496176_0013_r_000000_0' to
file:/tmp/temp1426823468/tmp-1227210921/_temporary/0/task_local2081496176_0013_r_00000
0
2018-07-14 22:16:32,270 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:16:32,270 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2081496176_0013_r_000000_0' done.
2018-07-14 22:16:32,270 [pool-37-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local2081496176_0013_r_000000_0
2018-07-14 22:16:32,271 [Thread-360] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
```

```
2018-07-14 22:16:32,486 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2081496176_0013
2018-07-14 22:16:32,487 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:16:32,487 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C:
ex_sum[12,8],ex_g[11,7] R: ex_sum[12,8]
2018-07-14 22:16:32,493 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:16:32,499 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:32,506 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:32,509 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:32,526 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:16:32,527 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:16:32,527 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:16:32,527 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:16:32,530 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:16:32,543 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:16:32,545 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:32,556 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:16:32,556 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:16:32,565 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:16:32,578 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:16:32,578 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:16:32,578 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:16:32,600 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:16:32,616 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local70206610_0014
2018-07-14 22:16:32,696 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:16:32,697 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter set in config null
2018-07-14 22:16:32,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:16:32,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:16:32,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:16:32,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:16:32,701 [Thread-389] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:16:32,701 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:16:32,705 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:16:32,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local70206610_0014_m_000000_0
2018-07-14 22:16:32,712 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:16:32,715 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:16:32,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1227210921/part-r-00000:0+62
2018-07-14 22:16:32,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:16:32,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:16:32,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:16:32,734 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:16:32,735 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:16:32,736 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:16:32,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:16:32,738 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:16:32,740 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:16:32,746 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:16:32,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:16:32,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:16:32,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 22:16:32,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:16:32,759 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:16:32,761 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local70206610_0014_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:16:32,765 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:16:32,765 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local70206610_0014_m_000000_0' done.
2018-07-14 22:16:32,765 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local70206610_0014_m_000000_0
2018-07-14 22:16:32,768 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:16:32,769 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:16:32,769 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local70206610_0014_r_000000_0
2018-07-14 22:16:32,777 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:16:32,777 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2d8018bc
2018-07-14 22:16:32,778 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:16:32,779 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local70206610_0014_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:16:32,780 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#12 about to shuffle
output of map attempt_local70206610_0014_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 22:16:32,780 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local70206610_0014_m_000000_0
2018-07-14 22:16:32,780 [localfetcher#12] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 22:16:32,780 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:16:32,781 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:16:32,781 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:16:32,782 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:16:32,782 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:16:32,783 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
```

```
disk to satisfy reduce memory limit
2018-07-14 22:16:32,783 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 22:16:32,783 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:16:32,783 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:16:32,783 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:16:32,783 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:16:32,791 [pool-40-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:16:32,792 [pool-40-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:16:32,798 [pool-40-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:16:32,799 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local70206610_0014_r_000000_0 is done. And is in the process of committing
2018-07-14 22:16:32,801 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:16:32,801 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local70206610_0014_r_000000_0 is allowed to commit now
2018-07-14 22:16:32,804 [pool-40-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local70206610_0014_r_000000_0' to
file:/tmp/temp1426823468/tmp1923252521/_temporary/0/task_local70206610_0014_r_000000
2018-07-14 22:16:32,805 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:16:32,805 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local70206610_0014_r_000000_0' done.
2018-07-14 22:16:32,805 [pool-40-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local70206610_0014_r_000000_0
2018-07-14 22:16:32,806 [Thread-389] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:16:33,004 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local70206610_0014
2018-07-14 22:16:33,005 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 22:16:33,005 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[14,9],ex_m[13,7] C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:16:33,025 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,031 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,033 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:16:33,049 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:16:32     2018-07-14 22:16:33
```

```
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local2081496176_0013        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local70206610_0014  1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
file:/tmp/temp1426823468/tmp1923252521,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp1426823468/tmp1923252521"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local2081496176_0013        ->      job_local70206610_0014,
job_local70206610_0014


2018-07-14 22:16:33,051 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,054 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,056 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,058 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,060 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,061 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:16:33,064 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:16:33,067 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:16:33,067 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:16:33,067 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:16:33,077 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:16:33,077 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(all,500)
grunt>
grunt>
grunt>
grunt> ans = filter ex_sum BY sum==ex_max.max;
```

```
grunt> dump ans;
2018-07-14 22:18:05,031 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY,FILTER
2018-07-14 22:18:05,053 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:05,054 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:18:05,054 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:18:05,054 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:18:05,067 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:18:05,074 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:18:05,074 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:18:05,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 22:18:05,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 22:18:05,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 22:18:05,078 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 22:18:05,088 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:05,088 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:18:05,089 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,090 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:18:05,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:18:05,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:18:05,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:18:05,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:18:05,091 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:18:05,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
```

```
Setting up single store job
2018-07-14 22:18:05,093 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:18:05,093 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:18:05,093 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531586885093-0
2018-07-14 22:18:05,103 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:18:05,105 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,115 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:18:05,117 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:18:05,118 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:18:05,118 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:18:05,118 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:18:05,139 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:18:05,158 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local666634513_0015
2018-07-14 22:18:05,263 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:18:05,264 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:18:05,267 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:18:05,267 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:18:05,267 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:05,267 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:18:05,267 [Thread-418] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:18:05,267 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:18:05,269 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:18:05,269 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local666634513_0015_m_000000_0
2018-07-14 22:18:05,273 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:18:05,274 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
```

```
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:18:05,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:18:05,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:18:05,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:18:05,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:18:05,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:18:05,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:18:05,291 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:18:05,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:18:05,293 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:18:05,293 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:18:05,295 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:18:05,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:18:05,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:18:05,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:18:05,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:18:05,296 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:18:05,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:18:05,306 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local666634513_0015_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:18:05,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:18:05,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local666634513_0015_m_000000_0' done.
2018-07-14 22:18:05,307 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local666634513_0015_m_000000_0
2018-07-14 22:18:05,308 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:18:05,308 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
```

```
2018-07-14 22:18:05,309 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local666634513_0015_r_000000_0
2018-07-14 22:18:05,315 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:18:05,315 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5487769
2018-07-14 22:18:05,316 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:18:05,318 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local666634513_0015_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:18:05,318 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#13 about to shuffle
output of map attempt_local666634513_0015_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:18:05,319 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local666634513_0015_m_000000_0
2018-07-14 22:18:05,319 [localfetcher#13] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:18:05,327 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:18:05,327 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:18:05,328 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:18:05,331 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:18:05,331 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:18:05,332 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:18:05,332 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:18:05,332 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:18:05,332 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:18:05,332 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:18:05,332 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:18:05,342 [pool-43-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:18:05,342 [pool-43-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:18:05,344 [pool-43-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:18:05,351 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local666634513_0015_r_000000_0 is done. And is in the process of committing
2018-07-14 22:18:05,352 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
```

```
- 1 / 1 copied.
2018-07-14 22:18:05,352 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local666634513_0015_r_000000_0 is allowed to commit now
2018-07-14 22:18:05,354 [pool-43-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local666634513_0015_r_000000_0' to
file:/tmp/temp1426823468/tmp1950520068/_temporary/0/task_local666634513_0015_r_000000
2018-07-14 22:18:05,356 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:18:05,356 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local666634513_0015_r_000000_0' done.
2018-07-14 22:18:05,356 [pool-43-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local666634513_0015_r_000000_0
2018-07-14 22:18:05,356 [Thread-418] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:18:05,466 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local666634513_0015
2018-07-14 22:18:05,466 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:18:05,466 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C:
ex_sum[12,8],ex_g[11,7] R: ex_sum[12,8]
2018-07-14 22:18:05,471 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 22:18:05,479 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,484 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,485 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,495 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:18:05,495 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:18:05,495 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:18:05,495 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:18:05,498 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:18:05,513 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:18:05,518 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,520 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:05,520 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:18:05,526 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
```

```
2018-07-14 22:18:05,535 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:18:05,535 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:18:05,535 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:18:05,561 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:18:05,577 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1235975428_0016
2018-07-14 22:18:05,664 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:18:05,665 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:18:05,668 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:18:05,668 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:18:05,668 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:05,668 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:18:05,668 [Thread-447] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:18:05,668 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:18:05,671 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:18:05,671 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1235975428_0016_m_000000_0
2018-07-14 22:18:05,674 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:18:05,676 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:18:05,678 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp1950520068/part-r-00000:0+62
2018-07-14 22:18:05,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:18:05,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:18:05,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:18:05,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
```

```
2018-07-14 22:18:05,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:18:05,702 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:18:05,703 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:18:05,703 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:18:05,705 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:18:05,706 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:18:05,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:18:05,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:18:05,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 22:18:05,707 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:18:05,717 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:18:05,718 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1235975428_0016_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:18:05,721 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:18:05,721 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1235975428_0016_m_000000_0' done.
2018-07-14 22:18:05,721 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1235975428_0016_m_000000_0
2018-07-14 22:18:05,721 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:18:05,722 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:18:05,726 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1235975428_0016_r_000000_0
2018-07-14 22:18:05,734 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:18:05,734 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3a1f0881
2018-07-14 22:18:05,734 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:18:05,735 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1235975428_0016_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:18:05,736 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#14 about to shuffle
output of map attempt_local1235975428_0016_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 22:18:05,736 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1235975428_0016_m_000000_0
2018-07-14 22:18:05,736 [localfetcher#14] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
```

```
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 22:18:05,737 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:18:05,737 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:18:05,737 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:18:05,738 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:18:05,738 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:18:05,740 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 22:18:05,741 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 22:18:05,741 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:18:05,741 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:18:05,741 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:18:05,741 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:18:05,750 [pool-46-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:18:05,750 [pool-46-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:18:05,752 [pool-46-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:18:05,752 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1235975428_0016_r_000000_0 is done. And is in the process of committing
2018-07-14 22:18:05,757 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:18:05,757 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1235975428_0016_r_000000_0 is allowed to commit now
2018-07-14 22:18:05,759 [pool-46-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1235975428_0016_r_000000_0' to
file:/tmp/temp1426823468/tmp-652627794/_temporary/0/task_local1235975428_0016_r_000000
2018-07-14 22:18:05,760 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:18:05,760 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1235975428_0016_r_000000_0' done.
2018-07-14 22:18:05,760 [pool-46-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1235975428_0016_r_000000_0
2018-07-14 22:18:05,760 [Thread-447] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:18:05,969 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1235975428_0016
2018-07-14 22:18:05,970 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 22:18:05,970 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
```

```
detailed locations: M: ex_max[14,9],ex_m[13,7] C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:18:05,979 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 22:18:05,987 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,989 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:05,991 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,001 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:18:06,001 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:18:06,004 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:18:06,012 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:18:06,014 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,017 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:06,017 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:18:06,028 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:18:06,036 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:18:06,036 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:18:06,036 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:18:06,056 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:18:06,071 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local708496389_0017
2018-07-14 22:18:06,142 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:18:06,142 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:18:06,145 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:18:06,145 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:18:06,145 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:06,145 [Thread-476] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:18:06,145 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:18:06,149 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:18:06,149 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local708496389_0017_m_000000_0
2018-07-14 22:18:06,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:18:06,159 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:18:06,160 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp1950520068/part-r-00000:0+62
2018-07-14 22:18:06,167 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:18:06,167 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:18:06,168 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: ans[15,6] C:  R:
2018-07-14 22:18:06,169 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:18:06,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:18:06,180 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:18:06,184 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:18:06,184 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local708496389_0017_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:18:06,186 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:18:06,186 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local708496389_0017_m_000000_0 is allowed to
commit now
2018-07-14 22:18:06,190 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local708496389_0017_m_000000_0' to
file:/tmp/temp1426823468/tmp1352723527/_temporary/0/task_local708496389_0017_m_000000
2018-07-14 22:18:06,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:18:06,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local708496389_0017_m_000000_0' done.
2018-07-14 22:18:06,191 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local708496389_0017_m_000000_0
2018-07-14 22:18:06,191 [Thread-476] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:18:06,345 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local708496389_0017
```

```
2018-07-14 22:18:06,346 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans
2018-07-14 22:18:06,346 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[15,6] C:  R:
2018-07-14 22:18:06,363 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,368 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,369 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,376 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:18:06,380 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:18:05     2018-07-14 22:18:06
GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1235975428_0016        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local666634513_0015 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local708496389_0017 1       0       n/a     n/a     n/a     n/a     0       0
0       0       ans     MAP_ONLY        file:/tmp/temp1426823468/tmp1352723527,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp1426823468/tmp1352723527"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local666634513_0015 ->      job_local1235975428_0016,job_local708496389_0017,
job_local1235975428_0016        ->      job_local708496389_0017,
job_local708496389_0017


2018-07-14 22:18:06,382 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,384 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,385 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,388 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,389 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
```

```
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,390 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,393 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,394 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,395 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:18:06,396 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:18:06,397 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:18:06,398 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:18:06,398 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:18:06,409 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:18:06,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,500)
grunt> ans1 = JOIN details By EmpId,ans By EmpId;
grunt> ans2 = foreach ans1 generate $0,$1,ans.sum;
grunt>
grunt>
grunt> dump ans1;
2018-07-14 22:20:08,060 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 22:20:08,078 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:08,078 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:20:08,078 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:20:08,079 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:20:08,093 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:20:08,095 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:20:08,096 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:20:08,098 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 22:20:08,098 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 3
2018-07-14 22:20:08,098 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 22:20:08,098 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
```

```
Merged 0 out of total 3 MR operators.
2018-07-14 22:20:08,098 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 3
2018-07-14 22:20:08,107 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:08,108 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:20:08,108 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:08,109 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:20:08,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:20:08,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:20:08,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:20:08,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:20:08,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:20:08,111 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:20:08,111 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:20:08,111 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:20:08,111 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531587008111-0
2018-07-14 22:20:08,122 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:20:08,123 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:08,133 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:20:08,135 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:20:08,135 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:20:08,135 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:20:08,139 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:20:08,160 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:20:08,176 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1980885730_0018
2018-07-14 22:20:08,254 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
```

```
2018-07-14 22:20:08,254 [Thread-501] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:20:08,257 [Thread-501] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:20:08,257 [Thread-501] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:20:08,257 [Thread-501] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:08,257 [Thread-501] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:20:08,257 [Thread-501] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:20:08,258 [Thread-501] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:20:08,261 [Thread-501] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:20:08,261 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1980885730_0018_m_000000_0
2018-07-14 22:20:08,265 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:08,265 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:20:08,267 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:20:08,267 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:20:08,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:20:08,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:20:08,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:20:08,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:20:08,292 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:20:08,293 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:20:08,294 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:08,294 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:20:08,296 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:20:08,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:20:08,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:20:08,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:20:08,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:20:08,299 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:20:08,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:20:08,318 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1980885730_0018_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:20:08,322 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:20:08,322 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1980885730_0018_m_000000_0' done.
2018-07-14 22:20:08,322 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1980885730_0018_m_000000_0
2018-07-14 22:20:08,322 [Thread-501] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:20:08,322 [Thread-501] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:20:08,326 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1980885730_0018_r_000000_0
2018-07-14 22:20:08,330 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:08,330 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5b78fe99
2018-07-14 22:20:08,331 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:20:08,332 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1980885730_0018_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:20:08,333 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#15 about to shuffle
output of map attempt_local1980885730_0018_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:20:08,333 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1980885730_0018_m_000000_0
2018-07-14 22:20:08,333 [localfetcher#15] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:20:08,340 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:20:08,341 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:20:08,341 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:20:08,342 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger -
```

```
Merging 1 sorted segments
2018-07-14 22:20:08,342 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:20:08,343 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:20:08,343 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:20:08,343 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:20:08,343 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:20:08,343 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:20:08,343 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:20:08,349 [pool-51-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:08,349 [pool-51-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:20:08,351 [pool-51-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:20:08,352 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1980885730_0018_r_000000_0 is done. And is in the process of committing
2018-07-14 22:20:08,353 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:20:08,353 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1980885730_0018_r_000000_0 is allowed to commit now
2018-07-14 22:20:08,356 [pool-51-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1980885730_0018_r_000000_0' to
file:/tmp/temp1426823468/tmp72908523/_temporary/0/task_local1980885730_0018_r_000000
2018-07-14 22:20:08,357 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:20:08,357 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1980885730_0018_r_000000_0' done.
2018-07-14 22:20:08,357 [pool-51-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1980885730_0018_r_000000_0
2018-07-14 22:20:08,357 [Thread-501] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:20:08,559 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1980885730_0018
2018-07-14 22:20:08,559 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:20:08,559 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C:
ex_sum[12,8],ex_g[11,7] R: ex_sum[12,8]
2018-07-14 22:20:08,563 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 33%
complete
2018-07-14 22:20:08,577 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:08,583 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 22:20:08,590 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:08,598 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:20:08,598 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:20:08,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:20:08,599 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:20:08,601 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:20:08,615 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:20:08,617 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:08,620 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:08,620 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:20:08,628 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:20:08,636 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:20:08,636 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:20:08,636 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:20:08,657 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:20:08,672 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local632328365_0019
2018-07-14 22:20:08,740 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:20:08,740 [Thread-530] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:20:08,744 [Thread-530] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:20:08,744 [Thread-530] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:20:08,744 [Thread-530] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:08,744 [Thread-530] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:20:08,744 [Thread-530] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
```

```
2018-07-14 22:20:08,744 [Thread-530] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:20:08,746 [Thread-530] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:20:08,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local632328365_0019_m_000000_0
2018-07-14 22:20:08,750 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:08,751 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:20:08,752 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp72908523/part-r-00000:0+62
2018-07-14 22:20:08,768 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:20:08,768 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:20:08,768 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:20:08,768 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:20:08,768 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:20:08,769 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:20:08,770 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:08,770 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:20:08,772 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:20:08,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:20:08,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:20:08,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:20:08,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 22:20:08,773 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:20:08,781 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:20:08,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local632328365_0019_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:20:08,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
```

```
2018-07-14 22:20:08,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local632328365_0019_m_000000_0' done.
2018-07-14 22:20:08,788 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local632328365_0019_m_000000_0
2018-07-14 22:20:08,788 [Thread-530] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:20:08,792 [Thread-530] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:20:08,792 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local632328365_0019_r_000000_0
2018-07-14 22:20:08,797 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:08,798 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6d00152e
2018-07-14 22:20:08,798 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:20:08,800 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local632328365_0019_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:20:08,801 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#16 about to shuffle
output of map attempt_local632328365_0019_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 22:20:08,801 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local632328365_0019_m_000000_0
2018-07-14 22:20:08,801 [localfetcher#16] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 22:20:08,802 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:20:08,802 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:20:08,802 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:20:08,803 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:20:08,803 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:20:08,806 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 22:20:08,806 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 22:20:08,806 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:20:08,806 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:20:08,807 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:20:08,807 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:20:08,814 [pool-54-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:08,814 [pool-54-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
```

```
- SchemaTupleBackend has already been initialized
2018-07-14 22:20:08,819 [pool-54-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:20:08,822 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local632328365_0019_r_000000_0 is done. And is in the process of committing
2018-07-14 22:20:08,828 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:20:08,828 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local632328365_0019_r_000000_0 is allowed to commit now
2018-07-14 22:20:08,829 [pool-54-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local632328365_0019_r_000000_0' to
file:/tmp/temp1426823468/tmp-871824933/_temporary/0/task_local632328365_0019_r_000000
2018-07-14 22:20:08,830 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:20:08,830 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local632328365_0019_r_000000_0' done.
2018-07-14 22:20:08,830 [pool-54-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local632328365_0019_r_000000_0
2018-07-14 22:20:08,830 [Thread-530] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:20:09,042 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local632328365_0019
2018-07-14 22:20:09,042 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 22:20:09,042 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[14,9],ex_m[13,7] C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:20:09,044 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 66%
complete
2018-07-14 22:20:09,046 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,047 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,047 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,068 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:20:09,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:20:09,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:20:09,069 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:20:09,071 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 22:20:09,071 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:20:09,074 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
```

```
2018-07-14 22:20:09,108 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:20:09,118 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,120 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:09,120 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:20:09,126 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:20:09,129 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:20:09,130 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:20:09,130 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:20:09,130 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:20:09,138 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:20:09,138 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:20:09,139 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:20:09,161 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:20:09,179 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local915432389_0020
2018-07-14 22:20:09,267 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:20:09,268 [Thread-559] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:20:09,272 [Thread-559] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:20:09,272 [Thread-559] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:20:09,272 [Thread-559] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:09,272 [Thread-559] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:20:09,272 [Thread-559] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:20:09,273 [Thread-559] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:20:09,276 [Thread-559] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:20:09,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
```

```
attempt_local915432389_0020_m_000000_0
2018-07-14 22:20:09,282 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:09,283 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:20:09,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:20:09,286 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:20:09,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:20:09,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:20:09,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:20:09,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:20:09,304 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:20:09,305 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:20:09,311 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:09,312 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:20:09,319 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[15,6],ans1[16,7],details[1,10],details[-1,-1],ans1[16,7] C:  R:
2018-07-14 22:20:09,326 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:20:09,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:20:09,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:20:09,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 22:20:09,327 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:20:09,328 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:20:09,330 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local915432389_0020_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:20:09,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:20:09,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local915432389_0020_m_000000_0' done.
2018-07-14 22:20:09,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local915432389_0020_m_000000_0
```

```
2018-07-14 22:20:09,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local915432389_0020_m_000001_0
2018-07-14 22:20:09,333 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:09,334 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:20:09,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp72908523/part-r-00000:0+62
2018-07-14 22:20:09,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:20:09,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:20:09,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:20:09,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:20:09,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:20:09,358 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:20:09,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:09,360 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:20:09,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[15,6],ans1[16,7],details[1,10],details[-1,-1],ans1[16,7] C:  R:
2018-07-14 22:20:09,364 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:20:09,377 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:20:09,377 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:20:09,379 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:20:09,379 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:20:09,379 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:20:09,379 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 12; bufvoid = 104857600
2018-07-14 22:20:09,379 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 22:20:09,380 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:20:09,382 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local915432389_0020_m_000001_0 is done. And
is in the process of committing
```

```
2018-07-14 22:20:09,384 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:20:09,384 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local915432389_0020_m_000001_0' done.
2018-07-14 22:20:09,384 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local915432389_0020_m_000001_0
2018-07-14 22:20:09,384 [Thread-559] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:20:09,384 [Thread-559] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:20:09,389 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local915432389_0020_r_000000_0
2018-07-14 22:20:09,398 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:20:09,398 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@fb4c390
2018-07-14 22:20:09,399 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:20:09,401 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local915432389_0020_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:20:09,403 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#17 about to shuffle
output of map attempt_local915432389_0020_m_000001_0 decomp: 16 len: 20 to MEMORY
2018-07-14 22:20:09,403 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 16 bytes from map-output
for attempt_local915432389_0020_m_000001_0
2018-07-14 22:20:09,403 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 16, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->16
2018-07-14 22:20:09,406 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#17 about to shuffle
output of map attempt_local915432389_0020_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 22:20:09,409 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local915432389_0020_m_000000_0
2018-07-14 22:20:09,410 [localfetcher#17] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 16, usedMemory
->368
2018-07-14 22:20:09,410 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:20:09,411 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:20:09,411 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:20:09,412 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:20:09,412 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 352 bytes
2018-07-14 22:20:09,414 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 368 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:20:09,414 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 370 bytes from
disk
2018-07-14 22:20:09,414 [pool-57-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:20:09,414 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:20:09,414 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 358 bytes
2018-07-14 22:20:09,414 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:20:09,425 [pool-57-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:20:09,425 [pool-57-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:20:09,426 [pool-57-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans[15,6],ans1[16,7],details[1,10],details[-1,-1],ans1[16,7] C:  R:
2018-07-14 22:20:09,428 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local915432389_0020_r_000000_0 is done. And is in the process of committing
2018-07-14 22:20:09,432 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:20:09,432 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local915432389_0020_r_000000_0 is allowed to commit now
2018-07-14 22:20:09,434 [pool-57-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local915432389_0020_r_000000_0' to
file:/tmp/temp1426823468/tmp-289388368/_temporary/0/task_local915432389_0020_r_000000
2018-07-14 22:20:09,435 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:20:09,435 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local915432389_0020_r_000000_0' done.
2018-07-14 22:20:09,435 [pool-57-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local915432389_0020_r_000000_0
2018-07-14 22:20:09,435 [Thread-559] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:20:09,569 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local915432389_0020
2018-07-14 22:20:09,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,ans1,details
2018-07-14 22:20:09,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[15,6],ans1[16,7],details[1,10],details[-1,-1],ans1[16,7] C:
R:
2018-07-14 22:20:09,587 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,595 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,601 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,630 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:20:09,636 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:20:08     2018-07-14 22:20:09
HASH_JOIN,GROUP_BY,FILTER

Success!
```

```
Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1980885730_0018        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local632328365_0019 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local915432389_0020 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans,ans1,details        HASH_JOIN
file:/tmp/temp1426823468/tmp-289388368,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp1426823468/tmp-289388368"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1980885730_0018        ->
job_local632328365_0019,job_local915432389_0020,
job_local632328365_0019 ->      job_local915432389_0020,
job_local915432389_0020


2018-07-14 22:20:09,639 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,641 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,643 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,646 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,647 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,648 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,650 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,650 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,651 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:20:09,652 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:20:09,653 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:20:09,653 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:20:09,653 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:20:09,661 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:20:09,661 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,10000,2,102,500)
grunt>
grunt>
grunt>
grunt>
grunt> dump ans2;
2018-07-14 22:21:53,639 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 22:21:53,655 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:53,655 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:21:53,655 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:21:53,655 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:21:53,662 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 22:21:53,669 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:21:53,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:21:53,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 5
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 map-only splittees.
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 1 out of total 3 MR operators.
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 map-reduce splittees.
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer -
Merged 0 out of total 3 MR operators.
2018-07-14 22:21:53,671 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 4
2018-07-14 22:21:53,681 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:53,681 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:21:53,681 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:53,682 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
```

```
the job
2018-07-14 22:21:53,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:21:53,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:21:53,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:21:53,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:21:53,683 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:21:53,684 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:21:53,684 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:21:53,684 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:21:53,684 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531587113684-0
2018-07-14 22:21:53,704 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:21:53,705 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:53,716 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:21:53,717 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:21:53,717 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:53,717 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:53,718 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:21:53,739 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:21:53,755 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1806058153_0021
2018-07-14 22:21:53,828 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:21:53,828 [Thread-591] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:21:53,831 [Thread-591] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:21:53,832 [Thread-591] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:21:53,832 [Thread-591] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 22:21:53,832 [Thread-591] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:21:53,832 [Thread-591] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:53,832 [Thread-591] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:21:53,834 [Thread-591] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:21:53,834 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1806058153_0021_m_000000_0
2018-07-14 22:21:53,838 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:53,838 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:21:53,840 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:21:53,841 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:21:53,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:21:53,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:21:53,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:21:53,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:21:53,855 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:21:53,856 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:21:53,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:53,857 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:21:53,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:21:53,861 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:21:53,861 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:21:53,861 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:21:53,861 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
```

```
2018-07-14 22:21:53,861 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:21:53,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:21:53,869 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1806058153_0021_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:21:53,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:21:53,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1806058153_0021_m_000000_0' done.
2018-07-14 22:21:53,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1806058153_0021_m_000000_0
2018-07-14 22:21:53,871 [Thread-591] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:21:53,871 [Thread-591] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:21:53,872 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1806058153_0021_r_000000_0
2018-07-14 22:21:53,875 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:53,875 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@74b5d96d
2018-07-14 22:21:53,876 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:21:53,877 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1806058153_0021_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:21:53,885 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#18 about to shuffle
output of map attempt_local1806058153_0021_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:21:53,885 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1806058153_0021_m_000000_0
2018-07-14 22:21:53,885 [localfetcher#18] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:21:53,886 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:21:53,887 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:21:53,887 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:21:53,888 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:21:53,888 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:21:53,889 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:21:53,889 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:21:53,889 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
```

```
2018-07-14 22:21:53,889 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:21:53,889 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:21:53,889 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:21:53,897 [pool-60-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:53,897 [pool-60-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:21:53,899 [pool-60-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C: ex_sum[12,8],ex_g[11,7] R:
ex_sum[12,8]
2018-07-14 22:21:53,901 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1806058153_0021_r_000000_0 is done. And is in the process of committing
2018-07-14 22:21:53,902 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:21:53,902 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1806058153_0021_r_000000_0 is allowed to commit now
2018-07-14 22:21:53,903 [pool-60-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1806058153_0021_r_000000_0' to
file:/tmp/temp1426823468/tmp1750155620/_temporary/0/task_local1806058153_0021_r_000000
2018-07-14 22:21:53,904 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:21:53,904 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1806058153_0021_r_000000_0' done.
2018-07-14 22:21:53,904 [pool-60-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1806058153_0021_r_000000_0
2018-07-14 22:21:53,905 [Thread-591] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:21:54,031 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1806058153_0021
2018-07-14 22:21:54,031 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:21:54,031 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[12,8],ex_g[11,7] C:
ex_sum[12,8],ex_g[11,7] R: ex_sum[12,8]
2018-07-14 22:21:54,034 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 25%
complete
2018-07-14 22:21:54,040 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,045 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,050 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,071 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:21:54,072 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:21:54,073 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:21:54,073 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:21:54,075 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:21:54,090 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:21:54,092 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,094 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:54,094 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:54,101 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:21:54,114 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:54,114 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:54,114 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:21:54,135 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:21:54,150 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1422169772_0022
2018-07-14 22:21:54,222 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:21:54,222 [Thread-620] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:21:54,227 [Thread-620] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:21:54,227 [Thread-620] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:21:54,228 [Thread-620] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:54,228 [Thread-620] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:21:54,228 [Thread-620] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:54,228 [Thread-620] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:21:54,230 [Thread-620] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:21:54,230 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1422169772_0022_m_000000_0
2018-07-14 22:21:54,234 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:54,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
```

```
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:21:54,235 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp1750155620/part-r-00000:0+62
2018-07-14 22:21:54,261 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:21:54,261 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:21:54,261 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:21:54,263 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:21:54,263 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:21:54,264 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:21:54,264 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:54,264 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:21:54,266 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:21:54,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:21:54,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:21:54,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:21:54,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 90; bufvoid = 104857600
2018-07-14 22:21:54,268 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:21:54,273 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:21:54,275 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1422169772_0022_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:21:54,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:21:54,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1422169772_0022_m_000000_0' done.
2018-07-14 22:21:54,276 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1422169772_0022_m_000000_0
2018-07-14 22:21:54,276 [Thread-620] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:21:54,276 [Thread-620] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:21:54,276 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1422169772_0022_r_000000_0
2018-07-14 22:21:54,282 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
```

```
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:54,282 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@43c4fd47
2018-07-14 22:21:54,283 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:21:54,284 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1422169772_0022_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:21:54,285 [localfetcher#19] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#19 about to shuffle
output of map attempt_local1422169772_0022_m_000000_0 decomp: 17 len: 21 to MEMORY
2018-07-14 22:21:54,285 [localfetcher#19] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 17 bytes from map-output
for attempt_local1422169772_0022_m_000000_0
2018-07-14 22:21:54,285 [localfetcher#19] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 17, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->17
2018-07-14 22:21:54,285 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:21:54,286 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:21:54,286 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:21:54,290 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:21:54,290 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:21:54,290 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 17 bytes to
disk to satisfy reduce memory limit
2018-07-14 22:21:54,290 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 21 bytes from
disk
2018-07-14 22:21:54,290 [pool-63-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:21:54,290 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:21:54,291 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 9 bytes
2018-07-14 22:21:54,291 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:21:54,299 [pool-63-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:54,300 [pool-63-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:21:54,304 [pool-63-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M: ex_max[14,9],ex_m[13,7]
C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:21:54,305 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1422169772_0022_r_000000_0 is done. And is in the process of committing
2018-07-14 22:21:54,306 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:21:54,306 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1422169772_0022_r_000000_0 is allowed to commit now
2018-07-14 22:21:54,307 [pool-63-thread-1] INFO
```

```
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1422169772_0022_r_000000_0' to
file:/tmp/temp1426823468/tmp-757600371/_temporary/0/task_local1422169772_0022_r_000000
2018-07-14 22:21:54,307 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:21:54,307 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1422169772_0022_r_000000_0' done.
2018-07-14 22:21:54,307 [pool-63-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1422169772_0022_r_000000_0
2018-07-14 22:21:54,307 [Thread-620] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:21:54,426 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1422169772_0022
2018-07-14 22:21:54,426 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_m,ex_max
2018-07-14 22:21:54,426 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ex_max[14,9],ex_m[13,7] C: ex_max[14,9],ex_m[13,7] R: ex_max[14,9]
2018-07-14 22:21:54,432 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:21:54,451 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,457 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,459 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,465 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:21:54,465 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:21:54,467 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up multi store job
2018-07-14 22:21:54,486 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:21:54,487 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,490 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:54,490 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:54,498 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:21:54,506 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:54,506 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:54,506 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:21:54,529 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:21:54,544 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
```

```
Submitting tokens for job: job_local1867272350_0023
2018-07-14 22:21:54,610 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:21:54,611 [Thread-649] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:21:54,616 [Thread-649] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:21:54,616 [Thread-649] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:21:54,616 [Thread-649] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:54,617 [Thread-649] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:54,617 [Thread-649] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:21:54,620 [Thread-649] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:21:54,621 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1867272350_0023_m_000000_0
2018-07-14 22:21:54,627 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:54,629 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:21:54,630 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp1750155620/part-r-00000:0+62
2018-07-14 22:21:54,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:54,631 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:21:54,635 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapOnly$Map - Aliases
being processed per job phase (AliasName[line,offset]): M: ans[15,6] C:  R:
2018-07-14 22:21:54,650 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:21:54,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:54,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:54,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:21:54,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1867272350_0023_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:21:54,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
```

```
2018-07-14 22:21:54,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task attempt_local1867272350_0023_m_000000_0 is allowed
to commit now
2018-07-14 22:21:54,660 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1867272350_0023_m_000000_0' to
file:/tmp/temp1426823468/tmp-1851513699/_temporary/0/task_local1867272350_0023_m_00000
0
2018-07-14 22:21:54,661 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1867272350_0023_m_000000_0' to
file:/tmp/temp1426823468/tmp-1786772177/_temporary/0/task_local1867272350_0023_m_00000
0
2018-07-14 22:21:54,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:21:54,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1867272350_0023_m_000000_0' done.
2018-07-14 22:21:54,665 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1867272350_0023_m_000000_0
2018-07-14 22:21:54,665 [Thread-649] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:21:54,812 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1867272350_0023
2018-07-14 22:21:54,813 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans
2018-07-14 22:21:54,813 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans[15,6] C:  R:
2018-07-14 22:21:54,814 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 75%
complete
2018-07-14 22:21:54,815 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,816 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,816 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,819 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:21:54,820 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:21:54,820 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:21:54,820 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:21:54,821 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=284
2018-07-14 22:21:54,823 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:21:54,825 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:21:54,832 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:21:54,834 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:54,836 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:54,836 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:54,842 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:21:54,844 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:21:54,844 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:54,844 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:54,845 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:21:54,853 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:54,853 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:54,853 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:21:54,874 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:21:54,889 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local571845289_0024
2018-07-14 22:21:54,958 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:21:54,959 [Thread-678] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:21:54,963 [Thread-678] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:21:54,963 [Thread-678] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:21:54,963 [Thread-678] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:54,963 [Thread-678] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:21:54,963 [Thread-678] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:21:54,963 [Thread-678] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:21:54,966 [Thread-678] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:21:54,966 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local571845289_0024_m_000000_0
```

```
2018-07-14 22:21:54,970 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:54,971 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:21:54,974 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:21:54,974 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:21:54,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:21:54,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:21:54,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:21:54,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:21:54,999 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:21:55,000 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:21:55,005 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:55,005 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:21:55,008 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[1,10],details[-1,-1],ans1[16,7],ans[-1,-1],ans1[16,7] C:  R: ans2[17,7]
2018-07-14 22:21:55,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:21:55,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:21:55,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:21:55,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 22:21:55,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:21:55,010 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:21:55,015 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local571845289_0024_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:21:55,017 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:21:55,017 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local571845289_0024_m_000000_0' done.
2018-07-14 22:21:55,017 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local571845289_0024_m_000000_0
2018-07-14 22:21:55,017 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local571845289_0024_m_000001_0
2018-07-14 22:21:55,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:55,025 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 9
Input split[0]:
   Length = 9
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:21:55,026 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1786772177/part-m-00000:0+9
2018-07-14 22:21:55,045 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:21:55,045 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:21:55,045 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:21:55,045 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:21:55,045 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:21:55,047 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:21:55,048 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:55,049 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:21:55,050 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[1,10],details[-1,-1],ans1[16,7],ans[-1,-1],ans1[16,7] C:  R: ans2[17,7]
2018-07-14 22:21:55,052 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:21:55,052 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:21:55,052 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:21:55,052 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 9; bufvoid = 104857600
2018-07-14 22:21:55,052 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214396(104857584); length = 1/6553600
2018-07-14 22:21:55,053 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:21:55,054 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local571845289_0024_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:21:55,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:21:55,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local571845289_0024_m_000001_0' done.
2018-07-14 22:21:55,055 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local571845289_0024_m_000001_0
2018-07-14 22:21:55,056 [Thread-678] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
```

```
task executor complete.
2018-07-14 22:21:55,056 [Thread-678] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:21:55,056 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local571845289_0024_r_000000_0
2018-07-14 22:21:55,061 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:21:55,061 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@6643a1ae
2018-07-14 22:21:55,063 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:21:55,064 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local571845289_0024_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:21:55,066 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#20 about to shuffle
output of map attempt_local571845289_0024_m_000001_0 decomp: 13 len: 17 to MEMORY
2018-07-14 22:21:55,066 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 13 bytes from map-output
for attempt_local571845289_0024_m_000001_0
2018-07-14 22:21:55,066 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 13, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->13
2018-07-14 22:21:55,067 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#20 about to shuffle
output of map attempt_local571845289_0024_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 22:21:55,068 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local571845289_0024_m_000000_0
2018-07-14 22:21:55,068 [localfetcher#20] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 2, commitMemory -> 13, usedMemory
->299
2018-07-14 22:21:55,069 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:21:55,070 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:21:55,071 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:21:55,071 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:21:55,071 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 283 bytes
2018-07-14 22:21:55,074 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 299 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:21:55,074 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 301 bytes from
disk
2018-07-14 22:21:55,074 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:21:55,074 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:21:55,074 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 289 bytes
2018-07-14 22:21:55,075 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
```

```
2018-07-14 22:21:55,082 [pool-68-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:21:55,082 [pool-68-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:21:55,088 [pool-68-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[1,10],details[-1,-1],ans1[16,7],ans[-1,-1],ans1[16,7] C:  R: ans2[17,7]
2018-07-14 22:21:55,088 [pool-68-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:21:55,096 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:55,096 [pool-68-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:21:55,100 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local571845289_0024_r_000000_0 is done. And is in the process of committing
2018-07-14 22:21:55,101 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:21:55,101 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local571845289_0024_r_000000_0 is allowed to commit now
2018-07-14 22:21:55,102 [pool-68-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local571845289_0024_r_000000_0' to
file:/tmp/temp1426823468/tmp1462878024/_temporary/0/task_local571845289_0024_r_000000
2018-07-14 22:21:55,104 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:21:55,105 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local571845289_0024_r_000000_0' done.
2018-07-14 22:21:55,105 [pool-68-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local571845289_0024_r_000000_0
2018-07-14 22:21:55,105 [Thread-678] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:21:55,264 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local571845289_0024
2018-07-14 22:21:55,264 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans,ans1,ans2,details
2018-07-14 22:21:55,264 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[1,10],details[-1,-1],ans1[16,7],ans[-1,-1],ans1[16,7] C:
R: ans2[17,7]
2018-07-14 22:21:55,278 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,281 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,283 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,291 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:21:55,300 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:21:53     2018-07-14 22:21:55
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
```

```
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1422169772_0022        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_m,ex_max     GROUP_BY,COMBINER
job_local1806058153_0021        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1867272350_0023        1       0       n/a     n/a     n/a     n/a     0
0       0       0       ans     MULTI_QUERY,MAP_ONLY
job_local571845289_0024 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans,ans1,ans2,details   HASH_JOIN
file:/tmp/temp1426823468/tmp1462878024,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 1 records in: "file:/tmp/temp1426823468/tmp1462878024"

Counters:
Total records written : 1
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1806058153_0021        ->
job_local1422169772_0022,job_local1867272350_0023,
job_local1422169772_0022        ->      job_local1867272350_0023,
job_local1867272350_0023        ->      job_local571845289_0024,
job_local571845289_0024


2018-07-14 22:21:55,300 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,304 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,305 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,308 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,313 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,314 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,318 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,318 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,319 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,320 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,320 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,321 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:21:55,323 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:21:55,324 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:21:55,324 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:21:55,324 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:21:55,332 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:21:55,332 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(102,Shahrukh,500)
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> details = load 'employee_details.txt' using PigStorage(',') as
(EmpId:Int,Name:CharArray,Salary:Int,Dept:Int);
2018-07-14 22:24:39,778 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:24:39,778 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
grunt>
grunt> ex_g = group expenses BY EmpId;
grunt> ans3 = JOIN details BY EmpId, ex_g BY EmpId;
2018-07-14 22:26:59,309 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1025:
<line 20, column 38> Invalid field projection. Projected field [EmpId] does not exist in
schema: group:int,expenses:bag{:tuple(EmpId:int,Expense:int)}.
Details at logfile: /home/acadgild/pig_1531584359423.log
grunt> ans3 = JOIN details BY EmpId, ex_g BY group.EmpId;
2018-07-14 22:27:24,569 [main] ERROR org.apache.pig.tools.grunt.Grunt - ERROR 1000:
<line 20, column 43> Invalid field reference. Referenced field [EmpId] does not exist in
schema: .
Details at logfile: /home/acadgild/pig_1531584359423.log
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
```

```
grunt>
grunt>
grunt> ex_g = group expenses BY EmpId;
grunt>
grunt> dump ex_g;
2018-07-14 22:28:09,839 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 22:28:09,853 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:28:09,854 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:28:09,854 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:28:09,854 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:28:09,855 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:28:09,856 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 22:28:09,856 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 22:28:09,865 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:28:09,866 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:28:09,866 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:09,867 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:28:09,867 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:28:09,868 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:28:09,868 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:28:09,868 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:28:09,868 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:28:09,869 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:28:09,869 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:28:09,869 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:28:09,869 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531587489869-0
```

```
2018-07-14 22:28:09,874 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:28:09,877 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:09,884 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:28:09,885 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:28:09,885 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:28:09,885 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:28:09,886 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:28:09,907 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:28:09,921 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1477905694_0025
2018-07-14 22:28:09,997 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:28:09,998 [Thread-710] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:28:10,001 [Thread-710] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:28:10,001 [Thread-710] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:28:10,001 [Thread-710] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:28:10,001 [Thread-710] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:28:10,001 [Thread-710] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:28:10,001 [Thread-710] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:28:10,003 [Thread-710] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:28:10,003 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1477905694_0025_m_000000_0
2018-07-14 22:28:10,006 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:28:10,007 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:28:10,009 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:28:10,009 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:28:10,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:28:10,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:28:10,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:28:10,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:28:10,023 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:28:10,024 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:28:10,025 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:28:10,026 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:28:10,027 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_g[20,7] C:  R:
2018-07-14 22:28:10,028 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:28:10,028 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:28:10,028 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:28:10,028 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 105; bufvoid = 104857600
2018-07-14 22:28:10,028 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:28:10,029 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:28:10,030 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1477905694_0025_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:28:10,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:28:10,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1477905694_0025_m_000000_0' done.
2018-07-14 22:28:10,032 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1477905694_0025_m_000000_0
2018-07-14 22:28:10,032 [Thread-710] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:28:10,033 [Thread-710] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:28:10,034 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1477905694_0025_r_000000_0
2018-07-14 22:28:10,040 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:28:10,040 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@737bf06c
2018-07-14 22:28:10,041 [pool-71-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
```

```
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:28:10,043 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1477905694_0025_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:28:10,044 [localfetcher#21] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#21 about to shuffle
output of map attempt_local1477905694_0025_m_000000_0 decomp: 125 len: 129 to MEMORY
2018-07-14 22:28:10,044 [localfetcher#21] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 125 bytes from
map-output for attempt_local1477905694_0025_m_000000_0
2018-07-14 22:28:10,045 [localfetcher#21] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 125, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->125
2018-07-14 22:28:10,046 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:28:10,047 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:28:10,047 [pool-71-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:28:10,049 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:28:10,055 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 22:28:10,057 [pool-71-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 125 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:28:10,057 [pool-71-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 129 bytes from
disk
2018-07-14 22:28:10,057 [pool-71-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:28:10,057 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:28:10,057 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 117 bytes
2018-07-14 22:28:10,057 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:28:10,069 [pool-71-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:28:10,070 [pool-71-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:28:10,071 [pool-71-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_g[20,7] C:  R:
2018-07-14 22:28:10,071 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1477905694_0025_r_000000_0 is done. And is in the process of committing
2018-07-14 22:28:10,074 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:28:10,074 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1477905694_0025_r_000000_0 is allowed to commit now
2018-07-14 22:28:10,075 [pool-71-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1477905694_0025_r_000000_0' to
file:/tmp/temp1426823468/tmp1868512290/_temporary/0/task_local1477905694_0025_r_000000
2018-07-14 22:28:10,081 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:28:10,081 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
'attempt_local1477905694_0025_r_000000_0' done.
2018-07-14 22:28:10,081 [pool-71-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1477905694_0025_r_000000_0
2018-07-14 22:28:10,082 [Thread-710] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:28:10,198 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1477905694_0025
2018-07-14 22:28:10,198 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,expenses
2018-07-14 22:28:10,198 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_g[20,7] C:  R:
2018-07-14 22:28:10,205 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:10,208 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:10,211 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:10,218 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:28:10,220 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:28:09     2018-07-14 22:28:10
GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1477905694_0025        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,expenses   GROUP_BY
file:/tmp/temp1426823468/tmp1868512290,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp1426823468/tmp1868512290"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1477905694_0025


2018-07-14 22:28:10,221 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:10,222 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:28:10,224 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 22:28:10,226 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:28:10,226 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:28:10,227 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:28:10,227 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:28:10,238 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:28:10,238 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,{(101,100),(101,200)})
(102,{(102,400),(102,100)})
(104,{(104,300)})
(105,{(105,100)})
(110,{(110,400)})
(114,{(114,200)})
(119,{(119,200)})
grunt>
grunt> ex_g = group expenses BY EmpId;
grunt>
grunt>
grunt> ex_sum= foreach ex_g GENERATE group as EmpId,SUM(expenses.Expense) as sum;
grunt>
grunt>
grunt> dump ex_sum;
2018-07-14 22:29:31,541 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: GROUP_BY
2018-07-14 22:29:31,561 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:29:31,561 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:29:31,562 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:29:31,562 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:29:31,569 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:29:31,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:29:31,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 1
2018-07-14 22:29:31,570 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 1
2018-07-14 22:29:31,582 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:29:31,582 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:29:31,583 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:31,584 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
```

```
the job
2018-07-14 22:29:31,585 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:29:31,585 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:29:31,585 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:29:31,585 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:29:31,585 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:29:31,586 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:29:31,586 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:29:31,586 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:29:31,587 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531587571586-0
2018-07-14 22:29:31,598 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:29:31,599 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:31,608 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:29:31,611 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:29:31,611 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:29:31,611 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:29:31,612 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:29:31,633 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:29:31,649 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1810796886_0026
2018-07-14 22:29:31,732 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:29:31,733 [Thread-739] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:29:31,736 [Thread-739] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:29:31,736 [Thread-739] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:29:31,736 [Thread-739] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 22:29:31,736 [Thread-739] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:29:31,736 [Thread-739] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:29:31,736 [Thread-739] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:29:31,738 [Thread-739] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:29:31,739 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1810796886_0026_m_000000_0
2018-07-14 22:29:31,744 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:29:31,746 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:29:31,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:29:31,749 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:29:31,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:29:31,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:29:31,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:29:31,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:29:31,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:29:31,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:29:31,779 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:29:31,779 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:29:31,782 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:29:31,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:29:31,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:29:31,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:29:31,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
```

```
2018-07-14 22:29:31,785 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:29:31,799 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:29:31,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1810796886_0026_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:29:31,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:29:31,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1810796886_0026_m_000000_0' done.
2018-07-14 22:29:31,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1810796886_0026_m_000000_0
2018-07-14 22:29:31,809 [Thread-739] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:29:31,809 [Thread-739] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:29:31,813 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1810796886_0026_r_000000_0
2018-07-14 22:29:31,817 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:29:31,817 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@724fc1f1
2018-07-14 22:29:31,818 [pool-74-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:29:31,818 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1810796886_0026_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:29:31,819 [localfetcher#22] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#22 about to shuffle
output of map attempt_local1810796886_0026_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:29:31,819 [localfetcher#22] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1810796886_0026_m_000000_0
2018-07-14 22:29:31,820 [localfetcher#22] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:29:31,820 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:29:31,820 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:29:31,820 [pool-74-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:29:31,821 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:29:31,821 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:29:31,822 [pool-74-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:29:31,822 [pool-74-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:29:31,822 [pool-74-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
```

```
2018-07-14 22:29:31,822 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:29:31,822 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:29:31,822 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:29:31,833 [pool-74-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:29:31,833 [pool-74-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:29:31,836 [pool-74-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:29:31,843 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1810796886_0026_r_000000_0 is done. And is in the process of committing
2018-07-14 22:29:31,844 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:29:31,844 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1810796886_0026_r_000000_0 is allowed to commit now
2018-07-14 22:29:31,845 [pool-74-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1810796886_0026_r_000000_0' to
file:/tmp/temp1426823468/tmp1036442657/_temporary/0/task_local1810796886_0026_r_000000
2018-07-14 22:29:31,846 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:29:31,846 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1810796886_0026_r_000000_0' done.
2018-07-14 22:29:31,846 [pool-74-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1810796886_0026_r_000000_0
2018-07-14 22:29:31,848 [Thread-739] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:29:32,034 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1810796886_0026
2018-07-14 22:29:32,035 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:29:32,035 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C:
ex_sum[22,8],ex_g[21,7] R: ex_sum[22,8]
2018-07-14 22:29:32,036 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:32,036 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:32,037 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:32,041 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:29:32,042 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:29:31     2018-07-14 22:29:32
GROUP_BY

Success!

Job Stats (time in seconds):
```

```
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1810796886_0026        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
file:/tmp/temp1426823468/tmp1036442657,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"

Output(s):
Successfully stored 7 records in: "file:/tmp/temp1426823468/tmp1036442657"

Counters:
Total records written : 7
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1810796886_0026


2018-07-14 22:29:32,042 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:32,043 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:32,043 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:29:32,047 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:29:32,047 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:29:32,047 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:29:32,047 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:29:32,054 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:29:32,054 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,300)
(102,500)
(104,300)
(105,100)
(110,400)
(114,200)
(119,200)
grunt>
grunt>
grunt>
grunt>
grunt> ans3 = JOIN details BY EmpId, ex_sum BY EmpId;
grunt> dump ans3;
2018-07-14 22:31:08,166 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 22:31:08,187 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:31:08,187 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
```

```
2018-07-14 22:31:08,187 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:31:08,188 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:31:08,190 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:31:08,191 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:31:08,192 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 22:31:08,193 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:31:08,193 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:31:08,201 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:31:08,201 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:31:08,202 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:08,203 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:31:08,203 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:31:08,203 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:31:08,203 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:31:08,203 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:31:08,203 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:31:08,204 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:31:08,204 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:31:08,204 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:31:08,204 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531587668204-0
2018-07-14 22:31:08,214 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:31:08,216 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 22:31:08,224 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:31:08,226 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:31:08,227 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:31:08,227 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:31:08,227 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:31:08,248 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:31:08,262 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local869315231_0027
2018-07-14 22:31:08,326 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:31:08,326 [Thread-768] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:31:08,329 [Thread-768] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:31:08,329 [Thread-768] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:31:08,329 [Thread-768] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:31:08,329 [Thread-768] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:31:08,329 [Thread-768] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:31:08,329 [Thread-768] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:31:08,331 [Thread-768] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:31:08,331 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local869315231_0027_m_000000_0
2018-07-14 22:31:08,334 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:31:08,335 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:31:08,337 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:31:08,337 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:31:08,360 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:31:08,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:31:08,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:31:08,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:31:08,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:31:08,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:31:08,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:31:08,364 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:31:08,366 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:31:08,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:31:08,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:31:08,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:31:08,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:31:08,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:31:08,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:31:08,375 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local869315231_0027_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:31:08,378 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:31:08,378 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local869315231_0027_m_000000_0' done.
2018-07-14 22:31:08,378 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local869315231_0027_m_000000_0
2018-07-14 22:31:08,378 [Thread-768] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:31:08,378 [Thread-768] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:31:08,382 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local869315231_0027_r_000000_0
2018-07-14 22:31:08,386 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:31:08,386 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4bbbc3bc
2018-07-14 22:31:08,386 [pool-77-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:31:08,387 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local869315231_0027_r_000000_0 Thread started: EventFetcher for fetching Map
```

```
Completion Events
2018-07-14 22:31:08,389 [localfetcher#23] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#23 about to shuffle
output of map attempt_local869315231_0027_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:31:08,389 [localfetcher#23] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local869315231_0027_m_000000_0
2018-07-14 22:31:08,389 [localfetcher#23] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:31:08,390 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:31:08,391 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:31:08,391 [pool-77-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:31:08,392 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:31:08,400 [pool-77-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:31:08,400 [pool-77-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:31:08,405 [pool-77-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:31:08,408 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local869315231_0027_r_000000_0 is done. And is in the process of committing
2018-07-14 22:31:08,410 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:31:08,410 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local869315231_0027_r_000000_0 is allowed to commit now
2018-07-14 22:31:08,411 [pool-77-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local869315231_0027_r_000000_0' to
file:/tmp/temp1426823468/tmp572901011/_temporary/0/task_local869315231_0027_r_000000
2018-07-14 22:31:08,412 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:31:08,412 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local869315231_0027_r_000000_0' done.
2018-07-14 22:31:08,412 [pool-77-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local869315231_0027_r_000000_0
```

```
2018-07-14 22:31:08,412 [Thread-768] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:31:08,528 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local869315231_0027
2018-07-14 22:31:08,528 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:31:08,528 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C:
ex_sum[22,8],ex_g[21,7] R: ex_sum[22,8]
2018-07-14 22:31:08,531 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:31:08,537 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:08,542 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:08,545 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:08,558 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:31:08,561 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:31:08,562 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:31:08,563 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:31:08,565 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 22:31:08,565 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:31:08,568 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:31:08,581 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:31:08,584 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:08,589 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:31:08,590 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:31:08,598 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:31:08,602 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:31:08,603 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:31:08,603 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
```

```
process : 1
2018-07-14 22:31:08,603 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:31:08,611 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:31:08,611 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:31:08,611 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:31:08,633 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:31:08,648 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1834679643_0028
2018-07-14 22:31:08,728 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:31:08,729 [Thread-797] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:31:08,731 [Thread-797] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:31:08,731 [Thread-797] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:31:08,731 [Thread-797] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:31:08,731 [Thread-797] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:31:08,732 [Thread-797] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:31:08,732 [Thread-797] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:31:08,734 [Thread-797] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:31:08,735 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1834679643_0028_m_000000_0
2018-07-14 22:31:08,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:31:08,738 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:31:08,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:31:08,741 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:31:08,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 22:31:08,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:31:08,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:31:08,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:31:08,763 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:31:08,764 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:31:08,765 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:31:08,765 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:31:08,769 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans3[23,7],ans3[23,7] C:  R:
2018-07-14 22:31:08,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:31:08,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:31:08,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:31:08,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 22:31:08,771 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:31:08,772 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:31:08,774 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1834679643_0028_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:31:08,775 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:31:08,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1834679643_0028_m_000000_0' done.
2018-07-14 22:31:08,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1834679643_0028_m_000000_0
2018-07-14 22:31:08,776 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1834679643_0028_m_000001_0
2018-07-14 22:31:08,777 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:31:08,778 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:31:08,779 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp572901011/part-r-00000:0+62
2018-07-14 22:31:08,797 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 22:31:08,797 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:31:08,797 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:31:08,797 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:31:08,797 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:31:08,798 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:31:08,799 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:31:08,799 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:31:08,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans3[23,7],ans3[23,7] C:  R:
2018-07-14 22:31:08,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:31:08,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:31:08,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:31:08,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 22:31:08,802 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:31:08,803 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:31:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1834679643_0028_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:31:08,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:31:08,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1834679643_0028_m_000001_0' done.
2018-07-14 22:31:08,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1834679643_0028_m_000001_0
2018-07-14 22:31:08,805 [Thread-797] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:31:08,806 [Thread-797] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:31:08,809 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1834679643_0028_r_000000_0
2018-07-14 22:31:08,817 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:31:08,817 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@3dc4b77f
2018-07-14 22:31:08,818 [pool-80-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:31:08,819 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1834679643_0028_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:31:08,821 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#24 about to shuffle
```

```
output of map attempt_local1834679643_0028_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 22:31:08,822 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1834679643_0028_m_000001_0
2018-07-14 22:31:08,822 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 22:31:08,825 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#24 about to shuffle
output of map attempt_local1834679643_0028_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 22:31:08,825 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1834679643_0028_m_000000_0
2018-07-14 22:31:08,825 [localfetcher#24] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 22:31:08,826 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:31:08,826 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:31:08,826 [pool-80-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:31:08,827 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:31:08,828 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 22:31:08,830 [pool-80-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:31:08,830 [pool-80-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 22:31:08,830 [pool-80-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:31:08,830 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:31:08,830 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 22:31:08,830 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:31:08,836 [pool-80-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:31:08,836 [pool-80-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:31:08,837 [pool-80-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans3[23,7],ans3[23,7] C:  R:
2018-07-14 22:31:08,838 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1834679643_0028_r_000000_0 is done. And is in the process of committing
2018-07-14 22:31:08,841 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:31:08,841 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1834679643_0028_r_000000_0 is allowed to commit now
2018-07-14 22:31:08,843 [pool-80-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1834679643_0028_r_000000_0' to
file:/tmp/temp1426823468/tmp-2094988715/_temporary/0/task_local1834679643_0028_r_00000
```

```
0
2018-07-14 22:31:08,844 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:31:08,844 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1834679643_0028_r_000000_0' done.
2018-07-14 22:31:08,844 [pool-80-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1834679643_0028_r_000000_0
2018-07-14 22:31:08,847 [Thread-797] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:31:09,033 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1834679643_0028
2018-07-14 22:31:09,033 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,details
2018-07-14 22:31:09,034 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[18,10],details[-1,-1],ans3[23,7],ans3[23,7] C:  R:
2018-07-14 22:31:09,047 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,052 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,064 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,086 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:31:09,087 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:31:08     2018-07-14 22:31:09
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1834679643_0028        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans3,details    HASH_JOIN
file:/tmp/temp1426823468/tmp-2094988715,
job_local869315231_0027 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 6 records in: "file:/tmp/temp1426823468/tmp-2094988715"

Counters:
Total records written : 6
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local869315231_0027 ->      job_local1834679643_0028,
job_local1834679643_0028
```

```
2018-07-14 22:31:09,088 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,089 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,091 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,094 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,096 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,097 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:31:09,099 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:31:09,099 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:31:09,100 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:31:09,100 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:31:09,109 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:31:09,109 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(110,Priyanka,2000,5,110,400)
(114,Madhuri,2000,2,114,200)
grunt>
grunt>
grunt> ans4 = foreach ans3 GENERATE $0,$1;
grunt>
grunt> dump ans4;
2018-07-14 22:33:57,214 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 22:33:57,236 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:33:57,236 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:33:57,236 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:33:57,236 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:33:57,237 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 22:33:57,237 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for expenses: $1
2018-07-14 22:33:57,239 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:33:57,240 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler$LastInputStrea
```

```
mingOptimizer - Rewrite: POPackage->POForEach to POPackage(JoinPackager)
2018-07-14 22:33:57,240 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:33:57,240 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:33:57,250 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:33:57,251 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:33:57,251 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:57,252 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:33:57,252 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:33:57,253 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:33:57,253 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:33:57,253 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:33:57,253 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:33:57,254 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:33:57,254 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:33:57,254 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:33:57,254 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531587837254-0
2018-07-14 22:33:57,260 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:33:57,262 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:57,271 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:33:57,273 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:33:57,274 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:33:57,274 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:33:57,274 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:33:57,312 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:33:57,328 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
```

```
Submitting tokens for job: job_local971338538_0029
2018-07-14 22:33:57,405 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:33:57,405 [Thread-827] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:33:57,408 [Thread-827] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:33:57,409 [Thread-827] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:33:57,409 [Thread-827] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:33:57,409 [Thread-827] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:33:57,409 [Thread-827] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:33:57,409 [Thread-827] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:33:57,413 [Thread-827] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:33:57,414 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local971338538_0029_m_000000_0
2018-07-14 22:33:57,417 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:33:57,418 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:33:57,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:33:57,420 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:33:57,436 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:33:57,436 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:33:57,436 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:33:57,436 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:33:57,436 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:33:57,437 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:33:57,438 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:33:57,438 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:33:57,441 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_g[21,7] C:  R: ex_sum[-1,-1]
2018-07-14 22:33:57,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:33:57,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:33:57,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:33:57,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 81; bufvoid = 104857600
2018-07-14 22:33:57,443 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:33:57,446 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:33:57,448 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local971338538_0029_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:33:57,449 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:33:57,449 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local971338538_0029_m_000000_0' done.
2018-07-14 22:33:57,449 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local971338538_0029_m_000000_0
2018-07-14 22:33:57,449 [Thread-827] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:33:57,449 [Thread-827] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:33:57,453 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local971338538_0029_r_000000_0
2018-07-14 22:33:57,458 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:33:57,458 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@56776346
2018-07-14 22:33:57,459 [pool-83-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:33:57,460 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local971338538_0029_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:33:57,462 [localfetcher#25] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#25 about to shuffle
output of map attempt_local971338538_0029_m_000000_0 decomp: 101 len: 105 to MEMORY
2018-07-14 22:33:57,462 [localfetcher#25] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 101 bytes from
map-output for attempt_local971338538_0029_m_000000_0
2018-07-14 22:33:57,462 [localfetcher#25] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 101, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->101
2018-07-14 22:33:57,463 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:33:57,464 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:33:57,464 [pool-83-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
```

```
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:33:57,465 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:33:57,465 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 22:33:57,466 [pool-83-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 101 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:33:57,466 [pool-83-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 105 bytes from
disk
2018-07-14 22:33:57,466 [pool-83-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:33:57,466 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:33:57,466 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 93 bytes
2018-07-14 22:33:57,466 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:33:57,474 [pool-83-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:33:57,475 [pool-83-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:33:57,477 [pool-83-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_g[21,7] C:  R: ex_sum[-1,-1]
2018-07-14 22:33:57,479 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local971338538_0029_r_000000_0 is done. And is in the process of committing
2018-07-14 22:33:57,481 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:33:57,481 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local971338538_0029_r_000000_0 is allowed to commit now
2018-07-14 22:33:57,482 [pool-83-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local971338538_0029_r_000000_0' to
file:/tmp/temp1426823468/tmp-1954700578/_temporary/0/task_local971338538_0029_r_000000
2018-07-14 22:33:57,483 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:33:57,483 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local971338538_0029_r_000000_0' done.
2018-07-14 22:33:57,483 [pool-83-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local971338538_0029_r_000000_0
2018-07-14 22:33:57,483 [Thread-827] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:33:57,607 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local971338538_0029
2018-07-14 22:33:57,607 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:33:57,607 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_g[21,7] C:  R: ex_sum[-1,-1]
2018-07-14 22:33:57,613 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:33:57,620 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:57,624 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
```

```
2018-07-14 22:33:57,627 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:57,649 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:33:57,650 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:33:57,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:33:57,651 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:33:57,653 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=317
2018-07-14 22:33:57,653 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:33:57,656 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:33:57,670 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:33:57,674 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:57,676 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:33:57,677 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:33:57,688 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:33:57,697 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:33:57,697 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:33:57,697 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:33:57,698 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:33:57,698 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:33:57,699 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:33:57,699 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:33:57,720 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:33:57,737 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1655978638_0030
2018-07-14 22:33:57,816 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:33:57,816 [Thread-856] INFO  org.apache.hadoop.mapred.LocalJobRunner -
```

```
OutputCommitter set in config null
2018-07-14 22:33:57,820 [Thread-856] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:33:57,820 [Thread-856] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:33:57,820 [Thread-856] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:33:57,820 [Thread-856] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:33:57,820 [Thread-856] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:33:57,820 [Thread-856] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:33:57,822 [Thread-856] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:33:57,823 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1655978638_0030_m_000000_0
2018-07-14 22:33:57,826 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:33:57,827 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:33:57,832 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:33:57,832 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:33:57,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:33:57,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:33:57,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:33:57,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:33:57,859 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:33:57,860 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:33:57,863 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:33:57,863 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:33:57,865 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
```

```
ans3[23,7],details[18,10],details[-1,-1],ans3[23,7] C:  R: ans4[24,7]
2018-07-14 22:33:57,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:33:57,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:33:57,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:33:57,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 22:33:57,868 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:33:57,870 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:33:57,872 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1655978638_0030_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:33:57,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:33:57,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1655978638_0030_m_000000_0' done.
2018-07-14 22:33:57,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1655978638_0030_m_000000_0
2018-07-14 22:33:57,873 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1655978638_0030_m_000001_0
2018-07-14 22:33:57,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:33:57,875 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 42
Input split[0]:
   Length = 42
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:33:57,876 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1954700578/part-r-00000:0+42
2018-07-14 22:33:57,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:33:57,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:33:57,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:33:57,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:33:57,896 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:33:57,897 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:33:57,898 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:33:57,907 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:33:57,911 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
```

```
ans3[23,7],details[18,10],details[-1,-1],ans3[23,7] C:  R: ans4[24,7]
2018-07-14 22:33:57,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:33:57,915 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:33:57,917 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:33:57,917 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 63; bufvoid = 104857600
2018-07-14 22:33:57,917 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:33:57,920 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:33:57,922 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1655978638_0030_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:33:57,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:33:57,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1655978638_0030_m_000001_0' done.
2018-07-14 22:33:57,923 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1655978638_0030_m_000001_0
2018-07-14 22:33:57,923 [Thread-856] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:33:57,924 [Thread-856] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:33:57,928 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1655978638_0030_r_000000_0
2018-07-14 22:33:57,933 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:33:57,933 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4077cd9c
2018-07-14 22:33:57,934 [pool-86-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:33:57,935 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1655978638_0030_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:33:57,936 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#26 about to shuffle
output of map attempt_local1655978638_0030_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 22:33:57,936 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local1655978638_0030_m_000000_0
2018-07-14 22:33:57,936 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->286
2018-07-14 22:33:57,937 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#26 about to shuffle
output of map attempt_local1655978638_0030_m_000001_0 decomp: 79 len: 83 to MEMORY
2018-07-14 22:33:57,938 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 79 bytes from map-output
for attempt_local1655978638_0030_m_000001_0
2018-07-14 22:33:57,938 [localfetcher#26] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 79, inMemoryMapOutputs.size() -> 2, commitMemory -> 286, usedMemory
->365
2018-07-14 22:33:57,941 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
```

```
Returning
2018-07-14 22:33:57,942 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:33:57,942 [pool-86-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:33:57,944 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:33:57,944 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 349 bytes
2018-07-14 22:33:57,948 [pool-86-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 365 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:33:57,948 [pool-86-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 367 bytes from
disk
2018-07-14 22:33:57,948 [pool-86-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:33:57,948 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:33:57,949 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 355 bytes
2018-07-14 22:33:57,949 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:33:57,956 [pool-86-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:33:57,956 [pool-86-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:33:57,957 [pool-86-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans3[23,7],details[18,10],details[-1,-1],ans3[23,7] C:  R: ans4[24,7]
2018-07-14 22:33:57,960 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1655978638_0030_r_000000_0 is done. And is in the process of committing
2018-07-14 22:33:57,962 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:33:57,963 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1655978638_0030_r_000000_0 is allowed to commit now
2018-07-14 22:33:57,965 [pool-86-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1655978638_0030_r_000000_0' to
file:/tmp/temp1426823468/tmp-1498375764/_temporary/0/task_local1655978638_0030_r_00000
0
2018-07-14 22:33:57,967 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:33:57,967 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1655978638_0030_r_000000_0' done.
2018-07-14 22:33:57,967 [pool-86-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1655978638_0030_r_000000_0
2018-07-14 22:33:57,967 [Thread-856] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:33:58,121 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1655978638_0030
2018-07-14 22:33:58,121 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans3,ans4,details
2018-07-14 22:33:58,121 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans3[23,7],details[18,10],details[-1,-1],ans3[23,7] C:  R:
ans4[24,7]
```

```
2018-07-14 22:33:58,127 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,132 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,134 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,146 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:33:58,148 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:33:57     2018-07-14 22:33:58
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1655978638_0030        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans3,ans4,details       HASH_JOIN
file:/tmp/temp1426823468/tmp-1498375764,
job_local971338538_0029 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 6 records in: "file:/tmp/temp1426823468/tmp-1498375764"

Counters:
Total records written : 6
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local971338538_0029 ->      job_local1655978638_0030,
job_local1655978638_0030


2018-07-14 22:33:58,149 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,149 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,150 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,153 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,154 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,154 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:33:58,157 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:33:58,157 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
```

```
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:33:58,158 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:33:58,158 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:33:58,167 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:33:58,167 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh)
(102,Shahrukh)
(104,Anubhav)
(105,Pawan)
(110,Priyanka)
(114,Madhuri)
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt>
grunt> ans5 = JOIN details BY EmpId LEFT OUTER, ex_sum BY EmpId;
grunt> dump ans5;
2018-07-14 22:39:38,774 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 22:39:38,795 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:39:38,795 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:39:38,795 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:39:38,795 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:39:38,804 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:39:38,805 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:39:38,809 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:39:38,809 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:39:38,818 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:39:38,818 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:39:38,818 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:38,820 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
```

```
the job
2018-07-14 22:39:38,820 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:39:38,820 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:39:38,820 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:39:38,821 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:39:38,821 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:39:38,822 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:39:38,822 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:39:38,822 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:39:38,822 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531588178822-0
2018-07-14 22:39:38,832 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:39:38,833 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:38,842 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:39:38,844 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:39:38,844 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:39:38,844 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:39:38,844 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:39:38,869 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:39:38,886 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local212290170_0031
2018-07-14 22:39:38,959 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:39:38,959 [Thread-886] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:39:38,962 [Thread-886] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:39:38,962 [Thread-886] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:39:38,963 [Thread-886] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
```

```
2018-07-14 22:39:38,963 [Thread-886] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:39:38,963 [Thread-886] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:39:38,963 [Thread-886] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:39:38,965 [Thread-886] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:39:38,965 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local212290170_0031_m_000000_0
2018-07-14 22:39:38,968 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:39:38,968 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:39:38,970 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:39:38,971 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:39:38,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:39:38,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:39:38,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:39:38,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:39:38,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:39:38,987 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:39:38,988 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:39:38,989 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:39:38,991 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:39:38,993 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:39:38,994 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:39:38,994 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:39:38,994 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
```

```
2018-07-14 22:39:38,994 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:39:39,002 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:39:39,003 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local212290170_0031_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:39:39,004 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:39:39,004 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local212290170_0031_m_000000_0' done.
2018-07-14 22:39:39,004 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local212290170_0031_m_000000_0
2018-07-14 22:39:39,004 [Thread-886] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:39:39,004 [Thread-886] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:39:39,005 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local212290170_0031_r_000000_0
2018-07-14 22:39:39,010 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:39:39,010 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@4e6e07e
2018-07-14 22:39:39,011 [pool-89-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:39:39,015 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local212290170_0031_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:39:39,016 [localfetcher#27] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#27 about to shuffle
output of map attempt_local212290170_0031_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:39:39,016 [localfetcher#27] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local212290170_0031_m_000000_0
2018-07-14 22:39:39,016 [localfetcher#27] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:39:39,017 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:39:39,018 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:39:39,018 [pool-89-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
```

```
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:39:39,019 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:39:39,020 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:39:39,027 [pool-89-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:39:39,027 [pool-89-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:39:39,028 [pool-89-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:39:39,029 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local212290170_0031_r_000000_0 is done. And is in the process of committing
2018-07-14 22:39:39,030 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:39:39,030 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local212290170_0031_r_000000_0 is allowed to commit now
2018-07-14 22:39:39,032 [pool-89-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local212290170_0031_r_000000_0' to
file:/tmp/temp1426823468/tmp666182937/_temporary/0/task_local212290170_0031_r_000000
2018-07-14 22:39:39,037 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:39:39,037 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local212290170_0031_r_000000_0' done.
2018-07-14 22:39:39,037 [pool-89-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local212290170_0031_r_000000_0
2018-07-14 22:39:39,037 [Thread-886] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:39:39,161 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local212290170_0031
2018-07-14 22:39:39,161 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:39:39,161 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C:
ex_sum[22,8],ex_g[21,7] R: ex_sum[22,8]
2018-07-14 22:39:39,162 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:39:39,163 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,164 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,164 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,175 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:39:39,176 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:39:39,176 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:39:39,177 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:39:39,178 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 22:39:39,178 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:39:39,179 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:39:39,195 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:39:39,197 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,200 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:39:39,200 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:39:39,209 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:39:39,217 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:39:39,217 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:39:39,217 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:39:39,218 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:39:39,218 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:39:39,218 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:39:39,218 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:39:39,237 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:39:39,251 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1004303690_0032
2018-07-14 22:39:39,326 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:39:39,327 [Thread-915] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:39:39,329 [Thread-915] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:39:39,329 [Thread-915] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:39:39,330 [Thread-915] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:39:39,330 [Thread-915] INFO
```

```
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:39:39,330 [Thread-915] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:39:39,331 [Thread-915] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:39:39,334 [Thread-915] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:39:39,334 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1004303690_0032_m_000000_0
2018-07-14 22:39:39,337 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:39:39,338 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:39:39,340 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:39:39,340 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:39:39,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:39:39,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:39:39,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:39:39,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:39:39,354 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:39:39,355 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:39:39,357 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:39:39,357 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:39:39,359 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R: ans5[25,7]
2018-07-14 22:39:39,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:39:39,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:39:39,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:39:39,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 22:39:39,360 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
```

```
2018-07-14 22:39:39,362 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:39:39,363 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1004303690_0032_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:39:39,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:39:39,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1004303690_0032_m_000000_0' done.
2018-07-14 22:39:39,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1004303690_0032_m_000000_0
2018-07-14 22:39:39,365 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1004303690_0032_m_000001_0
2018-07-14 22:39:39,368 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:39:39,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:39:39,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp666182937/part-r-00000:0+62
2018-07-14 22:39:39,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:39:39,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:39:39,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:39:39,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:39:39,510 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:39:39,511 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:39:39,512 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:39:39,512 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:39:39,514 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R: ans5[25,7]
2018-07-14 22:39:39,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:39:39,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:39:39,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:39:39,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 22:39:39,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
```

```
2018-07-14 22:39:39,516 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:39:39,517 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1004303690_0032_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:39:39,518 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:39:39,518 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1004303690_0032_m_000001_0' done.
2018-07-14 22:39:39,518 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1004303690_0032_m_000001_0
2018-07-14 22:39:39,518 [Thread-915] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:39:39,519 [Thread-915] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:39:39,519 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1004303690_0032_r_000000_0
2018-07-14 22:39:39,526 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:39:39,526 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@36e8f9f9
2018-07-14 22:39:39,527 [pool-92-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:39:39,528 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1004303690_0032_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:39:39,529 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#28 about to shuffle
output of map attempt_local1004303690_0032_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 22:39:39,529 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1004303690_0032_m_000000_0
2018-07-14 22:39:39,529 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->352
2018-07-14 22:39:39,530 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#28 about to shuffle
output of map attempt_local1004303690_0032_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 22:39:39,530 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1004303690_0032_m_000001_0
2018-07-14 22:39:39,530 [localfetcher#28] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 352, usedMemory
->451
2018-07-14 22:39:39,530 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:39:39,531 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:39:39,531 [pool-92-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:39:39,532 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:39:39,532 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 22:39:39,534 [pool-92-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
```

```
to disk to satisfy reduce memory limit
2018-07-14 22:39:39,534 [pool-92-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 22:39:39,534 [pool-92-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:39:39,534 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:39:39,534 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 22:39:39,534 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:39:39,541 [pool-92-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:39:39,541 [pool-92-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:39:39,544 [pool-92-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R: ans5[25,7]
2018-07-14 22:39:39,546 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1004303690_0032_r_000000_0 is done. And is in the process of committing
2018-07-14 22:39:39,547 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:39:39,547 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1004303690_0032_r_000000_0 is allowed to commit now
2018-07-14 22:39:39,547 [pool-92-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1004303690_0032_r_000000_0' to
file:/tmp/temp1426823468/tmp2119303889/_temporary/0/task_local1004303690_0032_r_000000
2018-07-14 22:39:39,548 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:39:39,548 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1004303690_0032_r_000000_0' done.
2018-07-14 22:39:39,548 [pool-92-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1004303690_0032_r_000000_0
2018-07-14 22:39:39,548 [Thread-915] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:39:39,696 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1004303690_0032
2018-07-14 22:39:39,697 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,details
2018-07-14 22:39:39,697 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R:
ans5[25,7]
2018-07-14 22:39:39,712 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,715 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,716 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,726 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:39:39,728 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
```

```
2.6.5   0.16.0  acadgild        2018-07-14 22:39:38     2018-07-14 22:39:39
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1004303690_0032        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,details    HASH_JOIN
file:/tmp/temp1426823468/tmp2119303889,
job_local212290170_0031 1       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp1426823468/tmp2119303889"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local212290170_0031 ->      job_local1004303690_0032,
job_local1004303690_0032


2018-07-14 22:39:39,729 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,730 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,731 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,733 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,733 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,734 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:39:39,735 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:39:39,735 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:39:39,736 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:39:39,736 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:39:39,745 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:39:39,745 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,101,300)
(102,Shahrukh,10000,2,102,500)
(103,Akshay,11000,3,,)
```

```
(104,Anubhav,5000,4,104,300)
(105,Pawan,2500,5,105,100)
(106,Aamir,25000,1,,)
(107,Salman,17500,2,,)
(108,Ranbir,14000,3,,)
(109,Katrina,1000,4,,)
(110,Priyanka,2000,5,110,400)
(111,Tushar,500,1,,)
(112,Ajay,5000,2,,)
(113,Jubeen,1000,1,,)
(114,Madhuri,2000,2,114,200)
grunt>
grunt>
grunt>
grunt> ans6 = foreach ans5 generate ..$3,$5;
grunt> dump ans6;
2018-07-14 22:41:34,938 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY
2018-07-14 22:41:34,952 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:41:34,952 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:41:34,952 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:41:34,952 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:41:34,957 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:41:34,960 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:41:34,961 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:41:34,961 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:41:34,971 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:41:34,971 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:41:34,972 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:34,973 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:41:34,973 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:41:34,973 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:41:34,973 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:41:34,974 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
```

```
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:41:34,974 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:41:34,975 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:41:34,976 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:41:34,976 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:41:34,976 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531588294975-0
2018-07-14 22:41:34,985 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:41:34,986 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:34,995 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:41:34,997 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:41:34,997 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:41:34,997 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:41:34,997 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:41:35,019 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:41:35,033 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1077675389_0033
2018-07-14 22:41:35,102 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:41:35,103 [Thread-945] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:41:35,106 [Thread-945] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:41:35,106 [Thread-945] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:41:35,106 [Thread-945] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:41:35,106 [Thread-945] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:41:35,106 [Thread-945] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:41:35,106 [Thread-945] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:41:35,108 [Thread-945] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:41:35,108 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
```

```
attempt_local1077675389_0033_m_000000_0
2018-07-14 22:41:35,111 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:41:35,112 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:41:35,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:41:35,114 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:41:35,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:41:35,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:41:35,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:41:35,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:41:35,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:41:35,154 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:41:35,155 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:41:35,155 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:41:35,161 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:41:35,162 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:41:35,162 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:41:35,162 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:41:35,162 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:41:35,162 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:41:35,170 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:41:35,170 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1077675389_0033_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:41:35,171 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:41:35,171 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1077675389_0033_m_000000_0' done.
2018-07-14 22:41:35,171 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1077675389_0033_m_000000_0
2018-07-14 22:41:35,171 [Thread-945] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:41:35,172 [Thread-945] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:41:35,177 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1077675389_0033_r_000000_0
2018-07-14 22:41:35,183 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:41:35,183 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@a9bdfb5
2018-07-14 22:41:35,184 [pool-95-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:41:35,185 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1077675389_0033_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:41:35,187 [localfetcher#29] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#29 about to shuffle
output of map attempt_local1077675389_0033_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:41:35,188 [localfetcher#29] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1077675389_0033_m_000000_0
2018-07-14 22:41:35,188 [localfetcher#29] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:41:35,189 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:41:35,190 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:41:35,190 [pool-95-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:41:35,190 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:41:35,190 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:41:35,194 [pool-95-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:41:35,194 [pool-95-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:41:35,194 [pool-95-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:41:35,194 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:41:35,194 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:41:35,194 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:41:35,203 [pool-95-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:41:35,203 [pool-95-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:41:35,206 [pool-95-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
```

```
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:41:35,206 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1077675389_0033_r_000000_0 is done. And is in the process of committing
2018-07-14 22:41:35,208 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 1 / 1 copied.
2018-07-14 22:41:35,208 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1077675389_0033_r_000000_0 is allowed to commit now
2018-07-14 22:41:35,210 [pool-95-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1077675389_0033_r_000000_0' to
file:/tmp/temp1426823468/tmp1994843870/_temporary/0/task_local1077675389_0033_r_000000
2018-07-14 22:41:35,210 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:41:35,210 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1077675389_0033_r_000000_0' done.
2018-07-14 22:41:35,210 [pool-95-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1077675389_0033_r_000000_0
2018-07-14 22:41:35,211 [Thread-945] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:41:35,406 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1077675389_0033
2018-07-14 22:41:35,406 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:41:35,406 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C:
ex_sum[22,8],ex_g[21,7] R: ex_sum[22,8]
2018-07-14 22:41:35,409 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:41:35,414 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,416 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,419 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,429 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:41:35,431 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:41:35,432 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:41:35,432 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:41:35,435 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 22:41:35,435 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:41:35,438 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:41:35,453 [main] INFO
```

```
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:41:35,455 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,459 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:41:35,459 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:41:35,465 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:41:35,467 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:41:35,467 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:41:35,467 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:41:35,467 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:41:35,474 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:41:35,474 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:41:35,474 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:41:35,495 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:41:35,508 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1801151887_0034
2018-07-14 22:41:35,572 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:41:35,573 [Thread-974] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:41:35,576 [Thread-974] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:41:35,576 [Thread-974] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:41:35,576 [Thread-974] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:41:35,576 [Thread-974] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:41:35,576 [Thread-974] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:41:35,576 [Thread-974] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:41:35,579 [Thread-974] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:41:35,579 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1801151887_0034_m_000000_0
```

```
2018-07-14 22:41:35,582 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:41:35,587 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:41:35,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:41:35,589 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:41:35,613 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:41:35,613 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:41:35,613 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:41:35,613 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:41:35,613 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:41:35,614 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:41:35,617 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:41:35,617 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:41:35,621 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R: ans5[25,7],ans6[26,7]
2018-07-14 22:41:35,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:41:35,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:41:35,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:41:35,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 322; bufvoid = 104857600
2018-07-14 22:41:35,623 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:41:35,624 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:41:35,627 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1801151887_0034_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:41:35,628 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:41:35,628 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1801151887_0034_m_000000_0' done.
2018-07-14 22:41:35,628 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1801151887_0034_m_000000_0
2018-07-14 22:41:35,628 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1801151887_0034_m_000001_0
2018-07-14 22:41:35,630 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:41:35,630 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:41:35,631 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp1994843870/part-r-00000:0+62
2018-07-14 22:41:35,649 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:41:35,649 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:41:35,649 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:41:35,649 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:41:35,649 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:41:35,650 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:41:35,651 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:41:35,652 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:41:35,654 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R: ans5[25,7],ans6[26,7]
2018-07-14 22:41:35,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:41:35,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:41:35,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:41:35,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 22:41:35,655 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:41:35,656 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:41:35,657 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1801151887_0034_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:41:35,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:41:35,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1801151887_0034_m_000001_0' done.
2018-07-14 22:41:35,658 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1801151887_0034_m_000001_0
2018-07-14 22:41:35,658 [Thread-974] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
```

```
task executor complete.
2018-07-14 22:41:35,659 [Thread-974] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:41:35,660 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Starting task: attempt_local1801151887_0034_r_000000_0
2018-07-14 22:41:35,664 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:41:35,665 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@df56d67
2018-07-14 22:41:35,665 [pool-98-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:41:35,668 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1801151887_0034_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:41:35,669 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#30 about to shuffle
output of map attempt_local1801151887_0034_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 22:41:35,669 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1801151887_0034_m_000001_0
2018-07-14 22:41:35,669 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 22:41:35,670 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#30 about to shuffle
output of map attempt_local1801151887_0034_m_000000_0 decomp: 352 len: 356 to MEMORY
2018-07-14 22:41:35,671 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 352 bytes from
map-output for attempt_local1801151887_0034_m_000000_0
2018-07-14 22:41:35,671 [localfetcher#30] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 352, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->451
2018-07-14 22:41:35,671 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:41:35,672 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:41:35,672 [pool-98-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:41:35,672 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:41:35,672 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 435 bytes
2018-07-14 22:41:35,673 [pool-98-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 451 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:41:35,673 [pool-98-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 453 bytes from
disk
2018-07-14 22:41:35,673 [pool-98-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:41:35,673 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:41:35,674 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 441 bytes
2018-07-14 22:41:35,674 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
```

```
2018-07-14 22:41:35,681 [pool-98-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:41:35,681 [pool-98-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:41:35,683 [pool-98-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R: ans5[25,7],ans6[26,7]
2018-07-14 22:41:35,684 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1801151887_0034_r_000000_0 is done. And is in the process of committing
2018-07-14 22:41:35,686 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- 2 / 2 copied.
2018-07-14 22:41:35,686 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1801151887_0034_r_000000_0 is allowed to commit now
2018-07-14 22:41:35,687 [pool-98-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1801151887_0034_r_000000_0' to
file:/tmp/temp1426823468/tmp-644673751/_temporary/0/task_local1801151887_0034_r_000000
2018-07-14 22:41:35,688 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- reduce > reduce
2018-07-14 22:41:35,688 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1801151887_0034_r_000000_0' done.
2018-07-14 22:41:35,688 [pool-98-thread-1] INFO  org.apache.hadoop.mapred.LocalJobRunner
- Finishing task: attempt_local1801151887_0034_r_000000_0
2018-07-14 22:41:35,688 [Thread-974] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:41:35,878 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1801151887_0034
2018-07-14 22:41:35,878 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans6,details
2018-07-14 22:41:35,878 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R:
ans5[25,7],ans6[26,7]
2018-07-14 22:41:35,882 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,884 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,884 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,889 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:41:35,890 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:41:34     2018-07-14 22:41:35
HASH_JOIN,GROUP_BY

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1077675389_0033        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local1801151887_0034        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,ans6,details       HASH_JOIN
```

```
file:/tmp/temp1426823468/tmp-644673751,

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 14 records in: "file:/tmp/temp1426823468/tmp-644673751"

Counters:
Total records written : 14
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1077675389_0033        ->      job_local1801151887_0034,
job_local1801151887_0034


2018-07-14 22:41:35,890 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,891 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,892 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,894 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,895 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,895 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:41:35,898 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:41:35,898 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:41:35,899 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:41:35,899 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:41:35,907 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:41:35,907 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(101,Amitabh,20000,1,300)
(102,Shahrukh,10000,2,500)
(103,Akshay,11000,3,)
(104,Anubhav,5000,4,300)
(105,Pawan,2500,5,100)
(106,Aamir,25000,1,)
(107,Salman,17500,2,)
(108,Ranbir,14000,3,)
(109,Katrina,1000,4,)
(110,Priyanka,2000,5,400)
(111,Tushar,500,1,)
(112,Ajay,5000,2,)
(113,Jubeen,1000,1,)
(114,Madhuri,2000,2,200)
grunt>
grunt>
```

```
grunt>
grunt> ans7 = filter ans6 BY sum is null;
grunt> ans8 = foreach ans7 generate EmpId,Name;
grunt> dump ans8;
2018-07-14 22:43:00,871 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 22:43:00,900 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:43:00,900 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:43:00,900 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:43:00,903 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:43:00,909 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 22:43:00,916 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:43:00,919 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:43:00,921 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:43:00,921 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:43:00,931 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:43:00,931 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:43:00,932 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:00,933 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:43:00,933 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:43:00,933 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:43:00,933 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:43:00,934 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:43:00,934 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:43:00,934 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:43:00,935 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
```

```
2018-07-14 22:43:00,935 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:43:00,935 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531588380935-0
2018-07-14 22:43:00,945 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:43:00,947 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:00,959 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:43:00,961 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:43:00,961 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:43:00,961 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:43:00,965 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:43:00,989 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:43:01,006 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local2141184433_0035
2018-07-14 22:43:01,097 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:43:01,097 [Thread-1004] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:43:01,100 [Thread-1004] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:43:01,100 [Thread-1004] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:43:01,100 [Thread-1004] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:43:01,100 [Thread-1004] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:43:01,100 [Thread-1004] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:43:01,100 [Thread-1004] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:43:01,103 [Thread-1004] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:43:01,103 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2141184433_0035_m_000000_0
2018-07-14 22:43:01,106 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:43:01,107 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
```

```
   Locations:

-----------------------

2018-07-14 22:43:01,108 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:43:01,109 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:43:01,122 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:43:01,122 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:43:01,122 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:43:01,122 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:43:01,122 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:43:01,123 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:43:01,124 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:43:01,124 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:43:01,126 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:43:01,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:43:01,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:43:01,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:43:01,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:43:01,130 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:43:01,140 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:43:01,141 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local2141184433_0035_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:43:01,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:43:01,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local2141184433_0035_m_000000_0' done.
2018-07-14 22:43:01,142 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2141184433_0035_m_000000_0
2018-07-14 22:43:01,142 [Thread-1004] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:43:01,142 [Thread-1004] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:43:01,142 [pool-101-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local2141184433_0035_r_000000_0
```

```
2018-07-14 22:43:01,148 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:43:01,148 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@2c592e61
2018-07-14 22:43:01,150 [pool-101-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:43:01,151 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local2141184433_0035_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:43:01,152 [localfetcher#31] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#31 about to shuffle
output of map attempt_local2141184433_0035_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:43:01,152 [localfetcher#31] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local2141184433_0035_m_000000_0
2018-07-14 22:43:01,152 [localfetcher#31] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:43:01,152 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:43:01,153 [pool-101-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 22:43:01,153 [pool-101-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:43:01,154 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:43:01,154 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:43:01,155 [pool-101-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:43:01,155 [pool-101-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:43:01,155 [pool-101-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:43:01,155 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:43:01,155 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:43:01,155 [pool-101-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 22:43:01,162 [pool-101-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:43:01,162 [pool-101-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:43:01,165 [pool-101-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:43:01,165 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local2141184433_0035_r_000000_0 is done. And is in the process of committing
2018-07-14 22:43:01,167 [pool-101-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 22:43:01,167 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local2141184433_0035_r_000000_0 is allowed to commit now
2018-07-14 22:43:01,167 [pool-101-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local2141184433_0035_r_000000_0' to
file:/tmp/temp1426823468/tmp-1046586739/_temporary/0/task_local2141184433_0035_r_00000
0
2018-07-14 22:43:01,168 [pool-101-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 22:43:01,168 [pool-101-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local2141184433_0035_r_000000_0' done.
2018-07-14 22:43:01,168 [pool-101-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local2141184433_0035_r_000000_0
2018-07-14 22:43:01,168 [Thread-1004] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:43:01,298 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local2141184433_0035
2018-07-14 22:43:01,298 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:43:01,298 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C:
ex_sum[22,8],ex_g[21,7] R: ex_sum[22,8]
2018-07-14 22:43:01,300 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:43:01,304 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,307 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,309 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,320 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:43:01,321 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:43:01,322 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:43:01,322 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:43:01,325 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 22:43:01,325 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:43:01,328 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:43:01,353 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:43:01,355 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,357 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
```

```
Instead, use dfs.bytes-per-checksum
2018-07-14 22:43:01,357 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:43:01,363 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:43:01,366 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:43:01,366 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:43:01,366 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:43:01,366 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:43:01,389 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:43:01,389 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:43:01,389 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:43:01,421 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:43:01,436 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local287438152_0036
2018-07-14 22:43:01,504 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:43:01,504 [Thread-1033] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:43:01,507 [Thread-1033] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:43:01,507 [Thread-1033] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:43:01,507 [Thread-1033] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:43:01,507 [Thread-1033] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:43:01,507 [Thread-1033] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:43:01,507 [Thread-1033] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:43:01,509 [Thread-1033] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:43:01,509 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local287438152_0036_m_000000_0
2018-07-14 22:43:01,514 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:43:01,515 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
```

```
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:43:01,518 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:43:01,518 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:43:01,535 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:43:01,535 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:43:01,535 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:43:01,535 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:43:01,535 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:43:01,536 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:43:01,537 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:43:01,538 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:43:01,541 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:43:01,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:43:01,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:43:01,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:43:01,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 22:43:01,548 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:43:01,550 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:43:01,551 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local287438152_0036_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:43:01,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:43:01,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local287438152_0036_m_000000_0' done.
2018-07-14 22:43:01,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local287438152_0036_m_000000_0
2018-07-14 22:43:01,556 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local287438152_0036_m_000001_0
2018-07-14 22:43:01,558 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:43:01,558 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:43:01,559 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-1046586739/part-r-00000:0+62
2018-07-14 22:43:01,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:43:01,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:43:01,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:43:01,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:43:01,684 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:43:01,685 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:43:01,686 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:43:01,686 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:43:01,687 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:43:01,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:43:01,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:43:01,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:43:01,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 22:43:01,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:43:01,688 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:43:01,689 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local287438152_0036_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:43:01,690 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:43:01,690 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local287438152_0036_m_000001_0' done.
2018-07-14 22:43:01,690 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local287438152_0036_m_000001_0
2018-07-14 22:43:01,690 [Thread-1033] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:43:01,691 [Thread-1033] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:43:01,691 [pool-104-thread-1] INFO
```

```
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local287438152_0036_r_000000_0
2018-07-14 22:43:01,698 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:43:01,698 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@5de4cc4c
2018-07-14 22:43:01,699 [pool-104-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:43:01,700 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local287438152_0036_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:43:01,701 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#32 about to shuffle
output of map attempt_local287438152_0036_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 22:43:01,701 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local287438152_0036_m_000000_0
2018-07-14 22:43:01,701 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->286
2018-07-14 22:43:01,702 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#32 about to shuffle
output of map attempt_local287438152_0036_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 22:43:01,702 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local287438152_0036_m_000001_0
2018-07-14 22:43:01,702 [localfetcher#32] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 2, commitMemory -> 286, usedMemory
->385
2018-07-14 22:43:01,702 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:43:01,703 [pool-104-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 22:43:01,703 [pool-104-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:43:01,703 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:43:01,703 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 369 bytes
2018-07-14 22:43:01,704 [pool-104-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 385 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:43:01,704 [pool-104-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 387 bytes from
disk
2018-07-14 22:43:01,704 [pool-104-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:43:01,704 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:43:01,704 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 375 bytes
2018-07-14 22:43:01,705 [pool-104-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 22:43:01,712 [pool-104-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
```

```
2018-07-14 22:43:01,712 [pool-104-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:43:01,715 [pool-104-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:43:01,717 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local287438152_0036_r_000000_0 is done. And is in the process of committing
2018-07-14 22:43:01,718 [pool-104-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 22:43:01,718 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local287438152_0036_r_000000_0 is allowed to commit now
2018-07-14 22:43:01,719 [pool-104-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local287438152_0036_r_000000_0' to
file:/tmp/temp1426823468/tmp483650944/_temporary/0/task_local287438152_0036_r_000000
2018-07-14 22:43:01,723 [pool-104-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 22:43:01,723 [pool-104-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local287438152_0036_r_000000_0' done.
2018-07-14 22:43:01,723 [pool-104-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local287438152_0036_r_000000_0
2018-07-14 22:43:01,723 [Thread-1033] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:43:01,855 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local287438152_0036
2018-07-14 22:43:01,855 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans7,ans8,details
2018-07-14 22:43:01,855 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: ans5[25,7],details[18,10],details[-1,-1],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:43:01,865 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,868 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,871 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,901 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:43:01,903 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:43:00     2018-07-14 22:43:01
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local2141184433_0035        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER
job_local287438152_0036 2       1       n/a     n/a     n/a     n/a     n/a     n/a
n/a     n/a     ans5,ans7,ans8,details  HASH_JOIN
file:/tmp/temp1426823468/tmp483650944,
```

```
Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 8 records in: "file:/tmp/temp1426823468/tmp483650944"

Counters:
Total records written : 8
Total bytes written : 0
Spillable Memory Manager spill count : 0
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local2141184433_0035        ->      job_local287438152_0036,
job_local287438152_0036


2018-07-14 22:43:01,905 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,907 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,909 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,912 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,913 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,915 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:43:01,921 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:43:01,922 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:43:01,922 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:43:01,922 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:43:01,934 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:43:01,934 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(103,Akshay)
(106,Aamir)
(107,Salman)
(108,Ranbir)
(109,Katrina)
(111,Tushar)
(112,Ajay)
(113,Jubeen)
grunt>
grunt> dump ans8;
2018-07-14 22:45:08,610 [main] INFO  org.apache.pig.tools.pigstats.ScriptState - Pig
features used in the script: HASH_JOIN,GROUP_BY,FILTER
2018-07-14 22:45:08,626 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:45:08,626 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:45:08,626 [main] INFO  org.apache.pig.data.SchemaTupleBackend - Key
```

```
[pig.schematuple] was not set... will not generate code.
2018-07-14 22:45:08,626 [main] INFO
org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer -
{RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator,
GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter,
MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer,
PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2018-07-14 22:45:08,629 [main] INFO
org.apache.pig.newplan.logical.rules.ColumnPruneVisitor - Columns pruned for details: $2,
$3
2018-07-14 22:45:08,631 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File
concatenation threshold: 100 optimistic? false
2018-07-14 22:45:08,631 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.CombinerOptimizerUtil - Choosing to
move algebraic foreach to combiner
2018-07-14 22:45:08,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size before optimization: 2
2018-07-14 22:45:08,632 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR
plan size after optimization: 2
2018-07-14 22:45:08,641 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:45:08,641 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:45:08,641 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:08,643 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:45:08,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:45:08,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:45:08,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:45:08,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=81
2018-07-14 22:45:08,644 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:45:08,645 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:45:08,645 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Key
[pig.schematuple] is false, will not generate code.
2018-07-14 22:45:08,645 [main] INFO  org.apache.pig.data.SchemaTupleFrontend - Starting
process to move generated code to distributed cacche
2018-07-14 22:45:08,645 [main] INFO  org.apache.pig.data.SchemaTupleFrontend -
Distributed cache not supported or needed in local mode. Setting key
[pig.schematuple.local.dir] with code temp directory: /tmp/1531588508645-0
2018-07-14 22:45:08,652 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:45:08,653 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:08,664 [JobControl] WARN
```

```
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:45:08,669 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:45:08,669 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:45:08,669 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:45:08,670 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:45:08,697 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:1
2018-07-14 22:45:08,712 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1680842245_0037
2018-07-14 22:45:08,778 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:45:08,778 [Thread-1063] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:45:08,781 [Thread-1063] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:45:08,781 [Thread-1063] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:45:08,781 [Thread-1063] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:45:08,781 [Thread-1063] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:45:08,781 [Thread-1063] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:45:08,781 [Thread-1063] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:45:08,783 [Thread-1063] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:45:08,784 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1680842245_0037_m_000000_0
2018-07-14 22:45:08,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:45:08,787 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 81
Input split[0]:
   Length = 81
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:45:08,789 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:45:08,789 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_expenses.txt:0+81
2018-07-14 22:45:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
```

```
2018-07-14 22:45:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:45:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:45:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:45:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:45:08,804 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:45:08,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:45:08,805 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not
generate code.
2018-07-14 22:45:08,809 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:45:08,812 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:45:08,812 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:45:08,812 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:45:08,812 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 114; bufvoid = 104857600
2018-07-14 22:45:08,812 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214364(104857456); length = 33/6553600
2018-07-14 22:45:08,818 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:45:08,819 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1680842245_0037_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:45:08,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:45:08,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1680842245_0037_m_000000_0' done.
2018-07-14 22:45:08,820 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1680842245_0037_m_000000_0
2018-07-14 22:45:08,820 [Thread-1063] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:45:08,821 [Thread-1063] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:45:08,822 [pool-107-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1680842245_0037_r_000000_0
2018-07-14 22:45:08,827 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:45:08,827 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@102facfe
2018-07-14 22:45:08,828 [pool-107-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:45:08,829 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
attempt_local1680842245_0037_r_000000_0 Thread started: EventFetcher for fetching Map
```

```
Completion Events
2018-07-14 22:45:08,830 [localfetcher#33] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#33 about to shuffle
output of map attempt_local1680842245_0037_m_000000_0 decomp: 106 len: 110 to MEMORY
2018-07-14 22:45:08,830 [localfetcher#33] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 106 bytes from
map-output for attempt_local1680842245_0037_m_000000_0
2018-07-14 22:45:08,830 [localfetcher#33] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 106, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->106
2018-07-14 22:45:08,831 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:45:08,831 [pool-107-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 22:45:08,831 [pool-107-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 1
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:45:08,832 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:45:08,833 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:45:08,833 [pool-107-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 1 segments, 106 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:45:08,833 [pool-107-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 110 bytes from
disk
2018-07-14 22:45:08,833 [pool-107-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:45:08,833 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:45:08,834 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 98 bytes
2018-07-14 22:45:08,834 [pool-107-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 22:45:08,840 [pool-107-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:45:08,840 [pool-107-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:45:08,842 [pool-107-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C: ex_sum[22,8],ex_g[21,7] R:
ex_sum[22,8]
2018-07-14 22:45:08,843 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1680842245_0037_r_000000_0 is done. And is in the process of committing
2018-07-14 22:45:08,846 [pool-107-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 1 / 1 copied.
2018-07-14 22:45:08,846 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
attempt_local1680842245_0037_r_000000_0 is allowed to commit now
2018-07-14 22:45:08,847 [pool-107-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1680842245_0037_r_000000_0' to
file:/tmp/temp1426823468/tmp-414403587/_temporary/0/task_local1680842245_0037_r_000000
2018-07-14 22:45:08,848 [pool-107-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 22:45:08,848 [pool-107-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1680842245_0037_r_000000_0' done.
2018-07-14 22:45:08,848 [pool-107-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
```

```
attempt_local1680842245_0037_r_000000_0
2018-07-14 22:45:08,848 [Thread-1063] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:45:09,125 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1680842245_0037
2018-07-14 22:45:09,125 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ex_g,ex_sum,expenses
2018-07-14 22:45:09,125 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: expenses[9,11],expenses[-1,-1],ex_sum[22,8],ex_g[21,7] C:
ex_sum[22,8],ex_g[21,7] R: ex_sum[22,8]
2018-07-14 22:45:09,128 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 50%
complete
2018-07-14 22:45:09,132 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,136 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,139 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,149 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.MRScriptState - Pig script settings are added to
the job
2018-07-14 22:45:09,151 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
mapred.job.reduce.markreset.buffer.percent is not set, set to default 0.3
2018-07-14 22:45:09,152 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Reduce phase detected, estimating # of required reducers.
2018-07-14 22:45:09,153 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler - Using
reducer estimator:
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
2018-07-14 22:45:09,156 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.InputSizeReducerEstimator
- BytesPerReducer=1000000000 maxReducers=999 totalInputFileSize=337
2018-07-14 22:45:09,156 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting Parallelism to 1
2018-07-14 22:45:09,161 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.JobControlCompiler -
Setting up single store job
2018-07-14 22:45:09,176 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 1
map-reduce job(s) waiting for submission.
2018-07-14 22:45:09,181 [JobControl] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics -
Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,185 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:45:09,185 [JobControl] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:45:09,192 [JobControl] WARN
org.apache.hadoop.mapreduce.JobResourceUploader - No job jar file set.  User classes may
not be found. See Job or Job#setJar(String).
2018-07-14 22:45:09,200 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:45:09,200 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
```

```
2018-07-14 22:45:09,200 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:45:09,201 [JobControl] INFO  org.apache.pig.builtin.PigStorage - Using
PigTextInputFormat
2018-07-14 22:45:09,201 [JobControl] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:45:09,201 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
2018-07-14 22:45:09,201 [JobControl] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths
(combined) to process : 1
2018-07-14 22:45:09,223 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
number of splits:2
2018-07-14 22:45:09,237 [JobControl] INFO  org.apache.hadoop.mapreduce.JobSubmitter -
Submitting tokens for job: job_local1127792275_0038
2018-07-14 22:45:09,308 [JobControl] INFO  org.apache.hadoop.mapreduce.Job - The url to
track the job: http://localhost:8080/
2018-07-14 22:45:09,308 [Thread-1092] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter set in config null
2018-07-14 22:45:09,311 [Thread-1092] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated.
Instead, use mapreduce.jobtracker.address
2018-07-14 22:45:09,311 [Thread-1092] INFO
org.apache.hadoop.conf.Configuration.deprecation -
mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use
mapreduce.reduce.markreset.buffer.percent
2018-07-14 22:45:09,311 [Thread-1092] INFO
org.apache.hadoop.conf.Configuration.deprecation - io.bytes.per.checksum is deprecated.
Instead, use dfs.bytes-per-checksum
2018-07-14 22:45:09,312 [Thread-1092] INFO
org.apache.hadoop.conf.Configuration.deprecation - mapred.reduce.tasks is deprecated.
Instead, use mapreduce.job.reduces
2018-07-14 22:45:09,312 [Thread-1092] INFO
org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated.
Instead, use fs.defaultFS
2018-07-14 22:45:09,312 [Thread-1092] INFO  org.apache.hadoop.mapred.LocalJobRunner -
OutputCommitter is
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigOutputCommitter
2018-07-14 22:45:09,313 [Thread-1092] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for map tasks
2018-07-14 22:45:09,314 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1127792275_0038_m_000000_0
2018-07-14 22:45:09,317 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:45:09,318 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 275
Input split[0]:
   Length = 275
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:45:09,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.builtin.PigStorage - Using PigTextInputFormat
2018-07-14 22:45:09,320 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/home/acadgild/employee_details.txt:0+275
2018-07-14 22:45:09,336 [LocalJobRunner Map Task Executor #0] INFO
```

```
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:45:09,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:45:09,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:45:09,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:45:09,336 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:45:09,337 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:45:09,339 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:45:09,339 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:45:09,342 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:45:09,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:45:09,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:45:09,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:45:09,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 256; bufvoid = 104857600
2018-07-14 22:45:09,345 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214344(104857376); length = 53/6553600
2018-07-14 22:45:09,346 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:45:09,347 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1127792275_0038_m_000000_0 is done. And
is in the process of committing
2018-07-14 22:45:09,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:45:09,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1127792275_0038_m_000000_0' done.
2018-07-14 22:45:09,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1127792275_0038_m_000000_0
2018-07-14 22:45:09,349 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1127792275_0038_m_000001_0
2018-07-14 22:45:09,350 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task -  Using ResourceCalculatorProcessTree : [ ]
2018-07-14 22:45:09,351 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Processing split: Number of splits :1
Total Length = 62
Input split[0]:
   Length = 62
   ClassName: org.apache.hadoop.mapreduce.lib.input.FileSplit
   Locations:

-----------------------

2018-07-14 22:45:09,353 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigRecordReader - Current
split being processed file:/tmp/temp1426823468/tmp-414403587/part-r-00000:0+62
```

```
2018-07-14 22:45:09,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - (EQUATOR) 0 kvi 26214396(104857584)
2018-07-14 22:45:09,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - mapreduce.task.io.sort.mb: 100
2018-07-14 22:45:09,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - soft limit at 83886080
2018-07-14 22:45:09,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufvoid = 104857600
2018-07-14 22:45:09,369 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396; length = 6553600
2018-07-14 22:45:09,370 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Map output collector class =
org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2018-07-14 22:45:09,371 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:45:09,371 [LocalJobRunner Map Task Executor #0] WARN
org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2018-07-14 22:45:09,373 [LocalJobRunner Map Task Executor #0] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigGenericMapReduce$Map -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:45:09,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner -
2018-07-14 22:45:09,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Starting flush of map output
2018-07-14 22:45:09,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Spilling map output
2018-07-14 22:45:09,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - bufstart = 0; bufend = 83; bufvoid = 104857600
2018-07-14 22:45:09,374 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - kvstart = 26214396(104857584); kvend =
26214372(104857488); length = 25/6553600
2018-07-14 22:45:09,375 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.MapTask - Finished spill 0
2018-07-14 22:45:09,376 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task:attempt_local1127792275_0038_m_000001_0 is done. And
is in the process of committing
2018-07-14 22:45:09,377 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - map
2018-07-14 22:45:09,377 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.Task - Task 'attempt_local1127792275_0038_m_000001_0' done.
2018-07-14 22:45:09,377 [LocalJobRunner Map Task Executor #0] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1127792275_0038_m_000001_0
2018-07-14 22:45:09,377 [Thread-1092] INFO  org.apache.hadoop.mapred.LocalJobRunner - map
task executor complete.
2018-07-14 22:45:09,378 [Thread-1092] INFO  org.apache.hadoop.mapred.LocalJobRunner -
Waiting for reduce tasks
2018-07-14 22:45:09,380 [pool-110-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Starting task:
attempt_local1127792275_0038_r_000000_0
2018-07-14 22:45:09,384 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Task -  Using
ResourceCalculatorProcessTree : [ ]
2018-07-14 22:45:09,384 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.ReduceTask -
Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@306d9bda
2018-07-14 22:45:09,385 [pool-110-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - MergerManager:
memoryLimit=652528832, maxSingleShuffleLimit=163132208, mergeThreshold=430669056,
ioSortFactor=10, memToMemMergeOutputsThreshold=10
2018-07-14 22:45:09,386 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher -
```

```
attempt_local1127792275_0038_r_000000_0 Thread started: EventFetcher for fetching Map
Completion Events
2018-07-14 22:45:09,387 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#34 about to shuffle
output of map attempt_local1127792275_0038_m_000001_0 decomp: 99 len: 103 to MEMORY
2018-07-14 22:45:09,387 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 99 bytes from map-output
for attempt_local1127792275_0038_m_000001_0
2018-07-14 22:45:09,387 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 99, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory ->99
2018-07-14 22:45:09,388 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.LocalFetcher - localfetcher#34 about to shuffle
output of map attempt_local1127792275_0038_m_000000_0 decomp: 286 len: 290 to MEMORY
2018-07-14 22:45:09,388 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.InMemoryMapOutput - Read 286 bytes from
map-output for attempt_local1127792275_0038_m_000000_0
2018-07-14 22:45:09,388 [localfetcher#34] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - closeInMemoryFile ->
map-output of size: 286, inMemoryMapOutputs.size() -> 2, commitMemory -> 99, usedMemory
->385
2018-07-14 22:45:09,388 [EventFetcher for fetching Map Completion Events] INFO
org.apache.hadoop.mapreduce.task.reduce.EventFetcher - EventFetcher is interrupted..
Returning
2018-07-14 22:45:09,389 [pool-110-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 22:45:09,389 [pool-110-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - finalMerge called with 2
in-memory map-outputs and 0 on-disk map-outputs
2018-07-14 22:45:09,391 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 2 sorted segments
2018-07-14 22:45:09,392 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 2 segments left of total size: 369 bytes
2018-07-14 22:45:09,395 [pool-110-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merged 2 segments, 385 bytes
to disk to satisfy reduce memory limit
2018-07-14 22:45:09,395 [pool-110-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 1 files, 387 bytes from
disk
2018-07-14 22:45:09,395 [pool-110-thread-1] INFO
org.apache.hadoop.mapreduce.task.reduce.MergeManagerImpl - Merging 0 segments, 0 bytes
from memory into reduce
2018-07-14 22:45:09,395 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Merger -
Merging 1 sorted segments
2018-07-14 22:45:09,395 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Merger - Down
to the last merge-pass, with 1 segments left of total size: 375 bytes
2018-07-14 22:45:09,398 [pool-110-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 22:45:09,406 [pool-110-thread-1] INFO
org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size
699400192 to monitor. collectionUsageThreshold = 489580128, usageThreshold = 489580128
2018-07-14 22:45:09,406 [pool-110-thread-1] WARN  org.apache.pig.data.SchemaTupleBackend
- SchemaTupleBackend has already been initialized
2018-07-14 22:45:09,409 [pool-110-thread-1] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.PigMapReduce$Reduce -
Aliases being processed per job phase (AliasName[line,offset]): M:
details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:45:09,415 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Task -
Task:attempt_local1127792275_0038_r_000000_0 is done. And is in the process of committing
2018-07-14 22:45:09,417 [pool-110-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - 2 / 2 copied.
2018-07-14 22:45:09,417 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
```

```
attempt_local1127792275_0038_r_000000_0 is allowed to commit now
2018-07-14 22:45:09,417 [pool-110-thread-1] INFO
org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter - Saved output of task
'attempt_local1127792275_0038_r_000000_0' to
file:/tmp/temp1426823468/tmp677879175/_temporary/0/task_local1127792275_0038_r_000000
2018-07-14 22:45:09,418 [pool-110-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - reduce > reduce
2018-07-14 22:45:09,418 [pool-110-thread-1] INFO  org.apache.hadoop.mapred.Task - Task
'attempt_local1127792275_0038_r_000000_0' done.
2018-07-14 22:45:09,418 [pool-110-thread-1] INFO
org.apache.hadoop.mapred.LocalJobRunner - Finishing task:
attempt_local1127792275_0038_r_000000_0
2018-07-14 22:45:09,418 [Thread-1092] INFO  org.apache.hadoop.mapred.LocalJobRunner -
reduce task executor complete.
2018-07-14 22:45:09,611 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
HadoopJobId: job_local1127792275_0038
2018-07-14 22:45:09,611 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Processing aliases ans5,ans7,ans8,details
2018-07-14 22:45:09,611 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
detailed locations: M: details[18,10],details[-1,-1],ans5[25,7],ans5[25,7] C:  R:
ans5[25,7],ans5[-1,-1],ans7[27,7],ans8[-1,-1]
2018-07-14 22:45:09,622 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,631 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,634 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,645 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - 100%
complete
2018-07-14 22:45:09,649 [main] INFO
org.apache.pig.tools.pigstats.mapreduce.SimplePigStats - Script Statistics:

HadoopVersion   PigVersion      UserId  StartedAt       FinishedAt      Features
2.6.5   0.16.0  acadgild        2018-07-14 22:45:08     2018-07-14 22:45:09
HASH_JOIN,GROUP_BY,FILTER

Success!

Job Stats (time in seconds):
JobId   Maps    Reduces MaxMapTime      MinMapTime      AvgMapTime      MedianMapTime
MaxReduceTime   MinReduceTime   AvgReduceTime   MedianReducetime        Alias  Feature
Outputs
job_local1127792275_0038        2       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ans5,ans7,ans8,details  HASH_JOIN
file:/tmp/temp1426823468/tmp677879175,
job_local1680842245_0037        1       1       n/a     n/a     n/a     n/a     n/a
n/a     n/a     n/a     ex_g,ex_sum,expenses    GROUP_BY,COMBINER

Input(s):
Successfully read 9 records from: "file:///home/acadgild/employee_expenses.txt"
Successfully read 14 records from: "file:///home/acadgild/employee_details.txt"

Output(s):
Successfully stored 8 records in: "file:/tmp/temp1426823468/tmp677879175"

Counters:
Total records written : 8
Total bytes written : 0
Spillable Memory Manager spill count : 0
```

```
Total bags proactively spilled: 0
Total records proactively spilled: 0

Job DAG:
job_local1680842245_0037        ->      job_local1127792275_0038,
job_local1127792275_0038


2018-07-14 22:45:09,650 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,651 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,652 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,658 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,661 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,662 [main] INFO  org.apache.hadoop.metrics.jvm.JvmMetrics - Cannot
initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
2018-07-14 22:45:09,663 [main] INFO
org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher -
Success!
2018-07-14 22:45:09,664 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
io.bytes.per.checksum is deprecated. Instead, use dfs.bytes-per-checksum
2018-07-14 22:45:09,664 [main] INFO  org.apache.hadoop.conf.Configuration.deprecation -
fs.default.name is deprecated. Instead, use fs.defaultFS
2018-07-14 22:45:09,664 [main] WARN  org.apache.pig.data.SchemaTupleBackend -
SchemaTupleBackend has already been initialized
2018-07-14 22:45:09,677 [main] INFO
org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2018-07-14 22:45:09,678 [main] INFO
org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to
process : 1
(103,Akshay)
(106,Aamir)
(107,Salman)
(108,Ranbir)
(109,Katrina)
(111,Tushar)
(112,Ajay)
(113,Jubeen)
grunt>
grunt>
grunt>
grunt>
grunt>
```