**Next version of the NASA Log Aggregation MRv2 job**

For this job, we will use real Dates as reduce/sort keys -

and will run this job on **JDK8 , CentOS6.7 CDH5.5 VM as user 'joe'**

For some reason, Hadoop does not provide a Date type of keys, as it does for Strings, Booleans and some others. So in order to sort and aggregate using the real Date types, we have to create a custom key type.

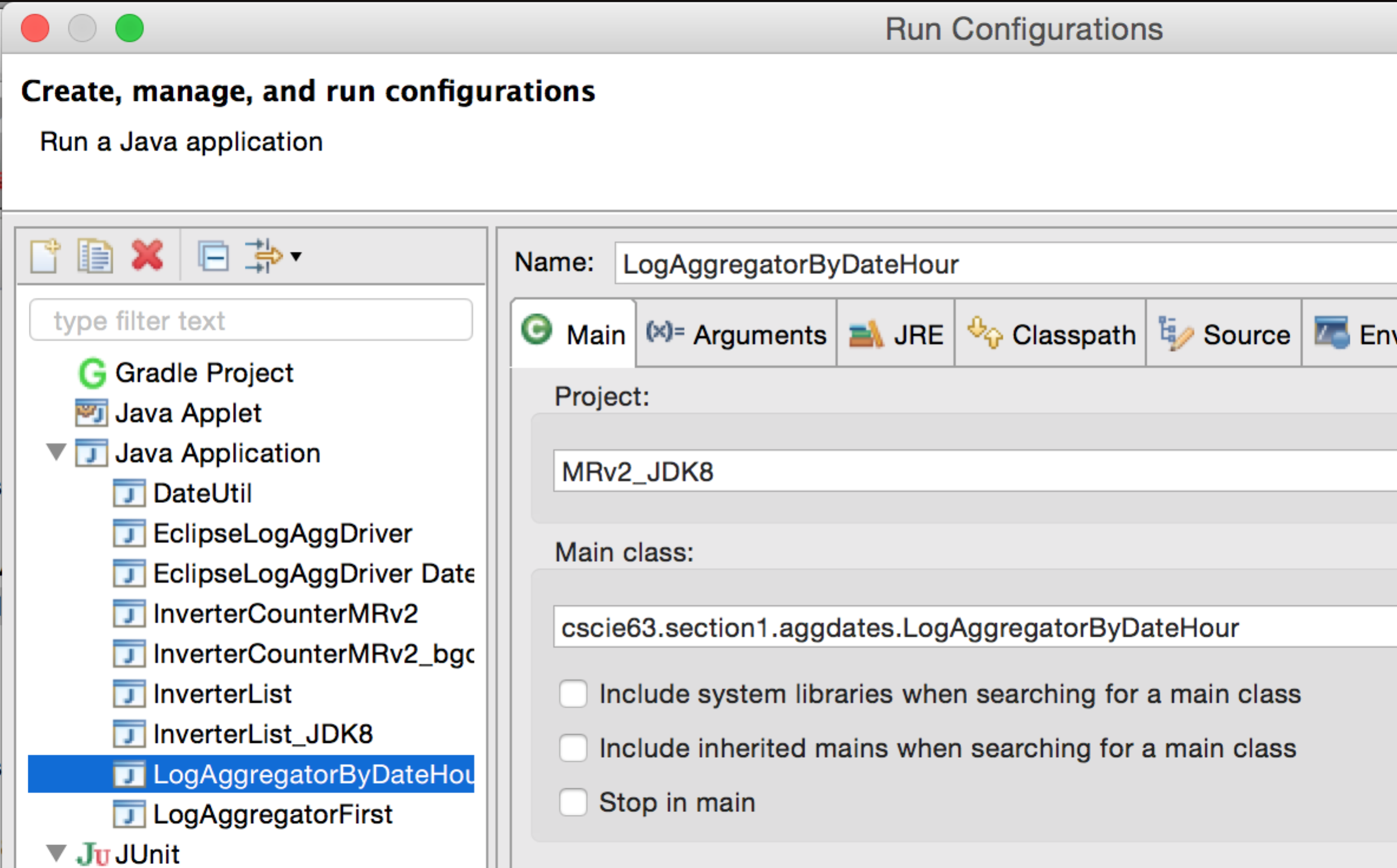
We will call it DateWritable.

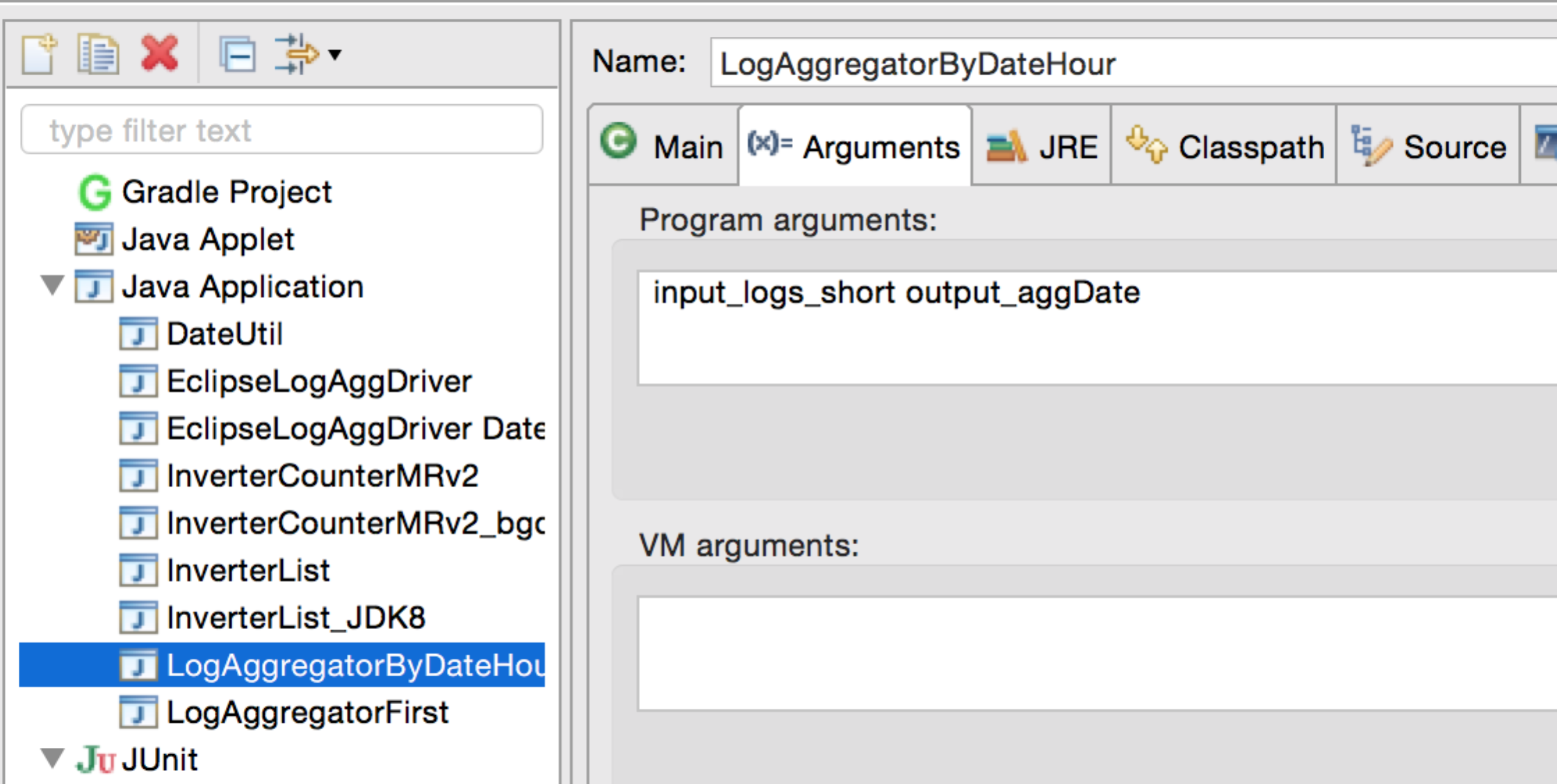
Project: MRv2\_JDK8

Relevant Code:

achine generated alternative text:
v MRv2_JDK8 
.DS Store 
input_logs_short 
src 
.DS Store 
cscie63 
.DS Store 
sectionl 
.DS Store 
aggdates 
DateWritable.java 
EclipseLogAggDriver.java 
LogAggregatorByDateHour.java 
LogParsingUtil.java 

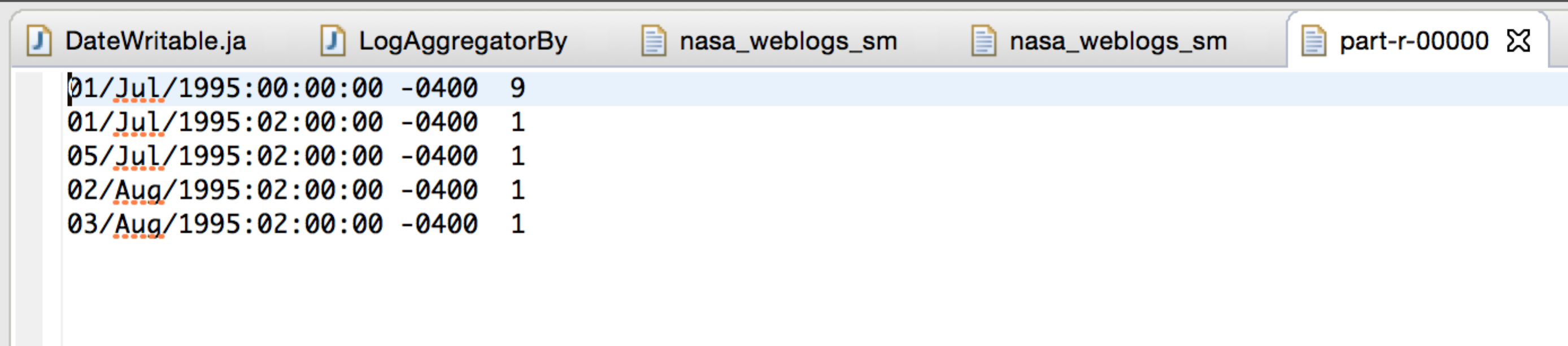
Create a new Run Configuration:





Run the job:

Notice how the results are correctly ordered now:



**Now lets create a JAR and make sure we can run this job as user ‘joe’ on the CentOS6.7 VM:**

Create a new JAR with the relevant code:

Right-click on the project and select “Export”

achine generated alternative text:
Java - 
MRv2 JDK8/src/cscie63/section1/a 
u:] DateW 
Package Explorer 
CSClE63 MRv2 
-J 
MRv2_J 
test 
*JRE 
CDH 
Binput 
Binput 
outp 
outp 
_s 
pa 
New 
Go Into 
Open in New Window 
Open Type Hierarchy 
Show In 
[C Copy 
Copy Qualified Name 
Paste 
Delete 
Remove from Context 
Build Path 
Source 
Refactor 
Import... 
Export... 
Refresh 
Close Project 
Assign Working Sets... 

achine generated alternative text:
Export 
Select 
Export resources into a JAR file on the local file system. 
Select an export destination: 
type filter text 
General 
B Install 
v BJava 
"JAR file 
Javadoc 
Runnable JAR file 
B Run/Debug 
Tasks 
Team 
BXML 
Other 
< Back 

Select only relevant code:

achine generated alternative text:
JAR Export 
JAR File Specification 
Define which resources should be exported into the JAR. 
Select the resources to export: 
MRv2_JDK8 
6: cscie63.section1 .aggdates 
edu.hu.bgd 
test 
.settings 
Binput_cited_short 
output_aggDate 
utoutÄnverterAjs 
O Export generated class files and resources 
n 
Export all output folders for checked projects 
n 
Export Java source files and resources 
DateWritable.java 
@ EclipseLogAggDriver.java 
LogAggregatorByDateHour.java 
LogParsingUtil.java 
n 
Select refactorings... 
Export refactorings for checked projects. 
Select the export destination: 
JAR file: /Users/marina o ova/Marina/VM_shared/lo date_a 
Options: 
o 
Compress the contents of the JAR file 
n 
Add directory entries 
n 
Overwrite existing files without warning 
ecl.•ar 
Cancel 
Browse... 
< Back 
Next > 

achine generated alternative text:
JAR Export 
JAR Packaging Options 
Define the options for the JAR export. 
Select options for handling problems: 
o 
Export class files with compile errors 
o 
Export class files with compile warnings 
Create source folder structure 
VI Build projects if not built automatically 
n 
Save the description of this JAR in the workspace 
Description file: 

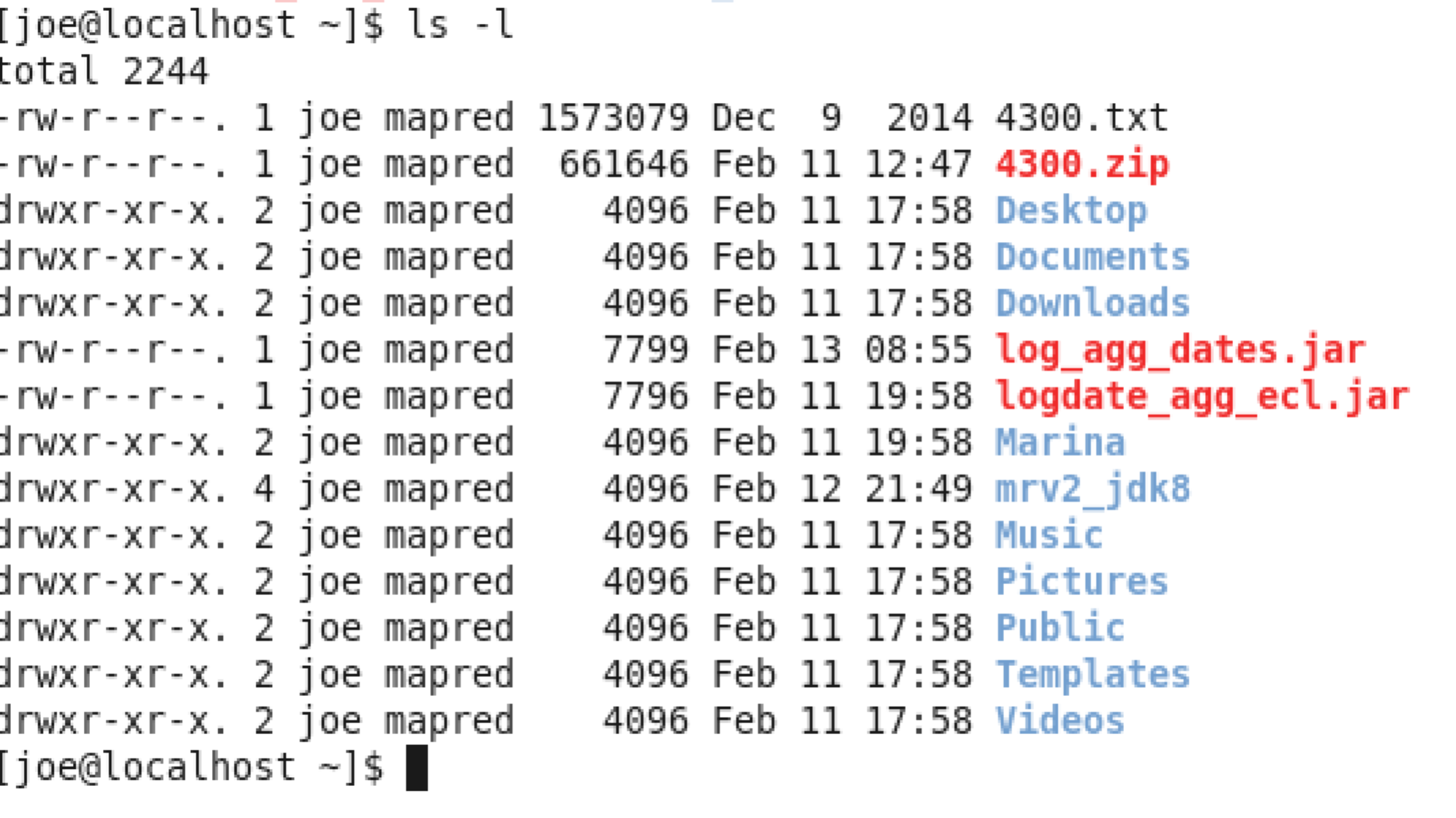
achine generated alternative text:
JAR Export 
JAR Manifest Specification 
Customize the manifest file for the JAR file. 
Specify the manifest: 
O 
Generate the manifest file 
n 
Save the manifest in the workspace 
171 use the saved manifest in the generated JAR description file 
Manifest file: 
C) 
Use existing manifest from workspace 
Manifest file: 
Seal contents: 
C) 
Seal the JAR 
Seal some packages 
Select the class of the application entry point: 
Main class: 
< Back 
Browse... 
Browse... 
Details... 
Details... 
Browse... 
Next > 
Nothing sealed 
Cancel 

achine generated alternative text:
JAR Export 
JAR export finished with warnings. See details for additional information. 
Details 
OK 

Click OK.

 Copy the resulting jar to your CentOS 6.7 + JDK8 VM, into user ‘joe’s home dir:

**logdate\_agg\_ecl.jar**



Put the input data into the HDFS **nasa\_logs** dir: (or any other dir you want)

[joe@localhost ~]$ hadoop fs -ls

Found 3 items

drwxr-xr-x   - joe supergroup          0 2016-02-11 13:46 count\_output

drwxr-xr-x   - joe supergroup          0 2016-02-11 12:53 joe\_input

drwxr-xr-x   - joe supergroup          0 2016-02-11 19:59 nasa\_logs

[joe@localhost ~]$ **hadoop fs -copyFromLocal /mnt/hgfs/VM\_shared/NASA\_access\_log\_Jul95 nasa\_logs**

[joe@localhost ~]$ **hadoop fs -ls nasa\_logs**

Found 1 items

-rw-r--r--   1 joe supergroup  205242368 2016-02-11 20:07 nasa\_logs/NASA\_access\_log\_Jul95

Run the job:

[joe@localhost ~]$ **hadoop jar logdate\_agg\_ecl.jar  cscie63.section1.aggdates.LogAggregatorByDateHour nasa\_logs agg\_output**

*16/02/11 20:07:53 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032*

*16/02/11 20:07:54 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.*

*16/02/11 20:07:54 INFO input.FileInputFormat: Total input paths to process : 1*

*16/02/11 20:07:54 INFO mapreduce.JobSubmitter: number of splits:2*

*16/02/11 20:07:55 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1455242135426\_0001*

*16/02/11 20:07:55 INFO impl.YarnClientImpl: Submitted application application\_1455242135426\_0001*

*16/02/11 20:07:55 INFO mapreduce.Job: The url to track the job:*[*http://localhost:8088/proxy/application\_1455242135426\_0001/*](http://localhost:8088/proxy/application_1455242135426_0001/)

*16/02/11 20:07:55 INFO mapreduce.Job: Running job: job\_1455242135426\_0001*

*16/02/11 20:08:06 INFO mapreduce.Job: Job job\_1455242135426\_0001 running in uber mode : false*

*16/02/11 20:08:06 INFO mapreduce.Job:  map 0% reduce 0%*

*16/02/11 20:08:22 INFO mapreduce.Job:  map 4% reduce 0%*

*16/02/11 20:08:25 INFO mapreduce.Job:  map 9% reduce 0%*

*16/02/11 20:08:28 INFO mapreduce.Job:  map 15% reduce 0%*

*16/02/11 20:08:31 INFO mapreduce.Job:  map 27% reduce 0%*

*16/02/11 20:08:34 INFO mapreduce.Job:  map 37% reduce 0%*

*16/02/11 20:08:37 INFO mapreduce.Job:  map 50% reduce 0%*

*16/02/11 20:08:40 INFO mapreduce.Job:  map 54% reduce 0%*

*16/02/11 20:08:43 INFO mapreduce.Job:  map 59% reduce 0%*

*16/02/11 20:08:46 INFO mapreduce.Job:  map 63% reduce 0%*

*16/02/11 20:08:49 INFO mapreduce.Job:  map 67% reduce 0%*

*16/02/11 20:09:05 INFO mapreduce.Job:  map 83% reduce 0%*

*16/02/11 20:09:19 INFO mapreduce.Job:  map 83% reduce 17%*

*16/02/11 20:09:25 INFO mapreduce.Job:  map 100% reduce 17%*

*16/02/11 20:09:28 INFO mapreduce.Job:  map 100% reduce 36%*

*16/02/11 20:09:31 INFO mapreduce.Job:  map 100% reduce 51%*

*16/02/11 20:09:34 INFO mapreduce.Job:  map 100% reduce 74%*

*16/02/11 20:09:37 INFO mapreduce.Job:  map 100% reduce 87%*

*16/02/11 20:09:40 INFO mapreduce.Job:  map 100% reduce 100%*

*16/02/11 20:09:40 INFO mapreduce.Job: Job job\_1455242135426\_0001 completed successfully*

*16/02/11 20:09:40 INFO mapreduce.Job: Counters: 49*

*File System Counters*

*FILE: Number of bytes read=68101746*

*FILE: Number of bytes written=136538909*

*FILE: Number of read operations=0*

*FILE: Number of large read operations=0*

*FILE: Number of write operations=0*

*HDFS: Number of bytes read=205246718*

*HDFS: Number of bytes written=21166*

*HDFS: Number of read operations=9*

*HDFS: Number of large read operations=0*

*HDFS: Number of write operations=2*

*Job Counters*

*Launched map tasks=2*

*Launched reduce tasks=1*

*Data-local map tasks=2*

*Total time spent by all maps in occupied slots (ms)=133027*

*Total time spent by all reduces in occupied slots (ms)=32056*

*Total time spent by all map tasks (ms)=133027*

*Total time spent by all reduce tasks (ms)=32056*

*Total vcore-seconds taken by all map tasks=133027*

*Total vcore-seconds taken by all reduce tasks=32056*

*Total megabyte-seconds taken by all map tasks=136219648*

*Total megabyte-seconds taken by all reduce tasks=32825344*

*Map-Reduce Framework*

*Map input records=1891715*

*Map output records=1891715*

*Map output bytes=64318310*

*Map output materialized bytes=68101752*

*Input split bytes=254*

*Combine input records=0*

*Combine output records=0*

*Reduce input groups=663*

*Reduce shuffle bytes=68101752*

*Reduce input records=1891715*

*Reduce output records=663*

*Spilled Records=3783430*

*Shuffled Maps =2*

*Failed Shuffles=0*

*Merged Map outputs=2*

*GC time elapsed (ms)=2033*

*CPU time spent (ms)=76330*

*Physical memory (bytes) snapshot=756191232*

*Virtual memory (bytes) snapshot=8160714752*

*Total committed heap usage (bytes)=549806080*

*Shuffle Errors*

*BAD\_ID=0*

*CONNECTION=0*

*IO\_ERROR=0*

*WRONG\_LENGTH=0*

*WRONG\_MAP=0*

*WRONG\_REDUCE=0*

*File Input Format Counters*

*Bytes Read=205246464*

*File Output Format Counters*

*Bytes Written=21166*

Check the results:

[joe@localhost ~]$ **hadoop fs -ls agg\_output**

Found 2 items

-rw-r--r--   1 joe supergroup          0 2016-02-11 20:09 agg\_output/\_SUCCESS

-rw-r--r--   1 joe supergroup      21166 2016-02-11 20:09 agg\_output/part-r-00000

[joe@localhost ~]$ **hadoop fs -cat agg\_output/part-r-00000 | head -10**

01/Jul/1995:00:00:00 -0400    3565

01/Jul/1995:01:00:00 -0400    3004

01/Jul/1995:02:00:00 -0400    2268

01/Jul/1995:03:00:00 -0400    1734

01/Jul/1995:04:00:00 -0400    1482

01/Jul/1995:05:00:00 -0400    1343

01/Jul/1995:06:00:00 -0400    1528

01/Jul/1995:07:00:00 -0400    1557

01/Jul/1995:08:00:00 -0400    1927

01/Jul/1995:09:00:00 -0400    2096

cat: Unable to write to output stream.

[joe@localhost ~]$ hadoop fs -cat agg\_output/part-r-00000 | tail -10

28/Jul/1995:05:00:00 -0400    916

28/Jul/1995:06:00:00 -0400    1086

28/Jul/1995:07:00:00 -0400    1616

28/Jul/1995:08:00:00 -0400    2543

28/Jul/1995:09:00:00 -0400    2741

28/Jul/1995:10:00:00 -0400    3555

28/Jul/1995:11:00:00 -0400    3570

28/Jul/1995:12:00:00 -0400    3505

28/Jul/1995:13:00:00 -0400    1905

11/Feb/2016:20:08:16 -0800    1

[joe@localhost ~]$

**Compiling and running your own MR job on CDH**

You do not have to use Eclipse to develop and compile your Java code and create JARs. You can do all that from a command line on your VM or local machine.

Lets do it on the **CentOS6.7 VM as the user ‘joe’**

Copy all code to a shared drive or scp to your VM

achine generated alternative text:
mrv2_jdk8 
.DS_Store 
cite75_99_short.txt 
src 
cscie63 
edu 
NASA_access_log_Ju195 

achine generated alternative text:
[j oe@localhost 
/home/ j oe/mrv2 
[j oe@localhost 
[j oe@localhost 
Feb 12 
Feb 12 
Feb 12 
Feb 12 
mrv2 
j dk8 
mrv2 
mrv2 
jdk8]$ 
jdk8]$ 
jdk8]$ 
pwd 
cp 
Is 
-r /mnt/hgfs/W_shared/mrv2 jdk8/src/ 
-l src/cscie63/section1/aggdates/ 
total 
16 
. 1 joe mapred 1288 
. 1 joe mapred 703 
. 1 joe mapred 2887 
. 1 joe mapred 2279 
[joe@localhost mrv2 jdk8]$ 
21 
21 
21 
21 
: 40 
: 40 
: 40 
: 40 
DateWritable. j ava 
EclipseLogAggD river. j ava 
LogAgg regato r8yDateHour. j ava 
LogPa rsingUtil . j ava 

Create 'classes' dir

**mkdir classes**

Compile your classes using Hadoop classpath:

*[joe@localhost mrv2\_jdk8]$* ***javac -classpath `hadoop classpath` src/cscie63/section1/aggdates/\*.java -d classes***

*[joe@localhost mrv2\_jdk8]$ ls -l classes*

*total 4*

*drwxr-xr-x. 3 joe mapred 4096 Feb 12 21:45 cscie63*

*[joe@localhost mrv2\_jdk8]$ ls -l classes/cscie63/section1/aggdates/*

*total 24*

*-rw-r--r--. 1 joe mapred 2474 Feb 12 21:45 DateWritable.class*

*-rw-r--r--. 1 joe mapred 766 Feb 12 21:45 EclipseLogAggDriver.class*

*-rw-r--r--. 1 joe mapred 2152 Feb 12 21:45 LogAggregatorByDateHour.class*

*-rw-r--r--. 1 joe mapred 2471 Feb 12 21:45 LogAggregatorByDateHour$LogAggMapper.class*

*-rw-r--r--. 1 joe mapred 1826 Feb 12 21:45 LogAggregatorByDateHour$LogAggReducer.class*

*-rw-r--r--. 1 joe mapred 2539 Feb 12 21:45 LogParsingUtil.class*

*[joe@localhost mrv2\_jdk8]$*

Create JAR:

cd classes

*[joe@localhost classes]$* ***jar -cvf ../logdates\_agg\_vm.jar cscie63/\****

*added manifest*

*adding: cscie63/section1/(in = 0) (out= 0)(stored 0%)*

*adding: cscie63/section1/aggdates/(in = 0) (out= 0)(stored 0%)*

*adding: cscie63/section1/aggdates/LogAggregatorByDateHour$LogAggMapper.class(in = 2471) (out= 1067)(deflated 56%)*

*adding: cscie63/section1/aggdates/LogAggregatorByDateHour.class(in = 2152) (out= 1005)(deflated 53%)*

*adding: cscie63/section1/aggdates/EclipseLogAggDriver.class(in = 766) (out= 460)(deflated 39%)*

*adding: cscie63/section1/aggdates/LogAggregatorByDateHour$LogAggReducer.class(in = 1826) (out= 718)(deflated 60%)*

*adding: cscie63/section1/aggdates/DateWritable.class(in = 2474) (out= 1036)(deflated 58%)*

*adding: cscie63/section1/aggdates/LogParsingUtil.class(in = 2539) (out= 1160)(deflated 54%)*

*[joe@localhost classes]$*

cd ..

Run the job (make sure output dir does not exist!):

**[joe@localhost** classes]$ cd ..

[joe@localhost mrv2\_jdk8]$ **hadoop jar logdates\_agg\_vm.jar cscie63.section1.aggdates.LogAggregatorByDateHour nasa\_logs nasa\_output**

*16/02/12 21:54:49 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032*

*16/02/12 21:54:50 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.*

*16/02/12 21:54:51 INFO input.FileInputFormat: Total input paths to process : 1*

*16/02/12 21:54:51 INFO mapreduce.JobSubmitter: number of splits:2*

*16/02/12 21:54:52 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1455341753307\_0001*

*16/02/12 21:54:52 INFO impl.YarnClientImpl: Submitted application application\_1455341753307\_0001*

*16/02/12 21:54:52 INFO mapreduce.Job: The url to track the job:* [*http://localhost:8088/proxy/application\_1455341753307\_0001/*](http://localhost:8088/proxy/application_1455341753307_0001/)

*16/02/12 21:54:52 INFO mapreduce.Job: Running job: job\_1455341753307\_0001*

*16/02/12 21:55:04 INFO mapreduce.Job: Job job\_1455341753307\_0001 running in uber mode : false*

*16/02/12 21:55:04 INFO mapreduce.Job: map 0% reduce 0%*

*16/02/12 21:55:19 INFO mapreduce.Job: map 1% reduce 0%*

*16/02/12 21:55:20 INFO mapreduce.Job: map 3% reduce 0%*

*16/02/12 21:55:22 INFO mapreduce.Job: map 4% reduce 0%*

*16/02/12 21:55:23 INFO mapreduce.Job: map 5% reduce 0%*

*16/02/12 21:55:25 INFO mapreduce.Job: map 7% reduce 0%*

*16/02/12 21:55:26 INFO mapreduce.Job: map 9% reduce 0%*

*16/02/12 21:55:28 INFO mapreduce.Job: map 10% reduce 0%*

*16/02/12 21:55:29 INFO mapreduce.Job: map 13% reduce 0%*

*16/02/12 21:55:31 INFO mapreduce.Job: map 14% reduce 0%*

*16/02/12 21:55:32 INFO mapreduce.Job: map 17% reduce 0%*

*16/02/12 21:55:34 INFO mapreduce.Job: map 20% reduce 0%*

*16/02/12 21:55:35 INFO mapreduce.Job: map 24% reduce 0%*

*16/02/12 21:55:37 INFO mapreduce.Job: map 27% reduce 0%*

*16/02/12 21:55:38 INFO mapreduce.Job: map 33% reduce 0%*

*16/02/12 21:55:40 INFO mapreduce.Job: map 36% reduce 0%*

*16/02/12 21:55:41 INFO mapreduce.Job: map 43% reduce 0%*

*16/02/12 21:55:43 INFO mapreduce.Job: map 47% reduce 0%*

*16/02/12 21:55:44 INFO mapreduce.Job: map 52% reduce 0%*

*16/02/12 21:55:46 INFO mapreduce.Job: map 54% reduce 0%*

*16/02/12 21:55:49 INFO mapreduce.Job: map 56% reduce 0%*

*16/02/12 21:55:52 INFO mapreduce.Job: map 58% reduce 0%*

*16/02/12 21:55:55 INFO mapreduce.Job: map 60% reduce 0%*

*16/02/12 21:55:58 INFO mapreduce.Job: map 64% reduce 0%*

*16/02/12 21:56:02 INFO mapreduce.Job: map 67% reduce 0%*

*16/02/12 21:56:19 INFO mapreduce.Job: map 83% reduce 0%*

*16/02/12 21:56:32 INFO mapreduce.Job: map 83% reduce 17%*

*16/02/12 21:56:40 INFO mapreduce.Job: map 100% reduce 17%*

*16/02/12 21:56:44 INFO mapreduce.Job: map 100% reduce 38%*

*16/02/12 21:56:47 INFO mapreduce.Job: map 100% reduce 60%*

*16/02/12 21:56:50 INFO mapreduce.Job: map 100% reduce 76%*

*16/02/12 21:56:53 INFO mapreduce.Job: map 100% reduce 88%*

*16/02/12 21:56:56 INFO mapreduce.Job: map 100% reduce 100%*

*16/02/12 21:56:57 INFO mapreduce.Job: Job job\_1455341753307\_0001 completed successfully*

*16/02/12 21:56:58 INFO mapreduce.Job: Counters: 49*

*File System Counters*

*FILE: Number of bytes read=68101746*

*FILE: Number of bytes written=136538912*

*FILE: Number of read operations=0*

*FILE: Number of large read operations=0*

*FILE: Number of write operations=0*

*HDFS: Number of bytes read=205246718*

*HDFS: Number of bytes written=21166*

*HDFS: Number of read operations=9*

*HDFS: Number of large read operations=0*

*HDFS: Number of write operations=2*

*Job Counters*

*Launched map tasks=2*

*Launched reduce tasks=1*

*Data-local map tasks=2*

*Total time spent by all maps in occupied slots (ms)=167820*

*Total time spent by all reduces in occupied slots (ms)=35025*

*Total time spent by all map tasks (ms)=167820*

*Total time spent by all reduce tasks (ms)=35025*

*Total vcore-seconds taken by all map tasks=167820*

*Total vcore-seconds taken by all reduce tasks=35025*

*Total megabyte-seconds taken by all map tasks=171847680*

*Total megabyte-seconds taken by all reduce tasks=35865600*

*Map-Reduce Framework*

*Map input records=1891715*

*Map output records=1891715*

*Map output bytes=64318310*

*Map output materialized bytes=68101752*

*Input split bytes=254*

*Combine input records=0*

*Combine output records=0*

*Reduce input groups=663*

*Reduce shuffle bytes=68101752*

*Reduce input records=1891715*

*Reduce output records=663*

*Spilled Records=3783430*

*Shuffled Maps =2*

*Failed Shuffles=0*

*Merged Map outputs=2*

*GC time elapsed (ms)=2162*

*CPU time spent (ms)=81320*

*Physical memory (bytes) snapshot=759541760*

*Virtual memory (bytes) snapshot=8160854016*

*Total committed heap usage (bytes)=549523456*

*Shuffle Errors*

*BAD\_ID=0*

*CONNECTION=0*

*IO\_ERROR=0*

*WRONG\_LENGTH=0*

*WRONG\_MAP=0*

*WRONG\_REDUCE=0*

*File Input Format Counters*

*Bytes Read=205246464*

*File Output Format Counters*

*Bytes Written=21166*

[joe@localhost mrv2\_jdk8]$

Notice the URL of the Job tracker:

<http://localhost:8088/proxy/application_1455341753307_0001/>

Check it out:

achine generated alternative text:
MapReduce Application application_1455341753307 
MapReduce Application 
x Welcome to Centos 
0001 - Mozilla Firefox 
Search 
Cluster 
Active Jobs 
MapReduce Application 
application 1455341753307 0001 
Search: 
Application 
Show 20 
entries 
About 
Jobs 
Job ID 
Tools 
job 1455341753307 0001 
Showing 1 to 1 of 1 entries 
Name 
LogAggregatorByHour 
Maps 
Map 
State C 
Total 
Progress C 
Maps 
Completed 
Reduce 
Progress C 
RUNNING 
2 
Re 
1 
@ Firefox automatically sends some data to Mozilla so that we can improve your experience. 
First Previo 
Choose What I Share x 

Click on the job link

achine generated alternative text:
MapReduce Job - Mozilla 
MapReduce Job job 145... x Welcome to Centos 
localhost 
Firefox 
Search 
MapReduce Job 
job 1455341753307 
Application 
Job 
Overview 
Counters 
Configuration 
Map tasks 
Reduce tasks 
Tools 
Job Name: 
User Name: 
Queue: 
State: 
Uberized: 
Submitted: 
Started: 
Finished: 
Elapsed: 
Diagnostics: 
Average Map Time 
Average Shuffle Time 
Average Merge Time 
Average Reduce Time 
LogAggregatorByHour 
joe 
root.joe 
SUCCEEDED 
ApplicationMaster 
Attempt Number 
1 
false 
Fri Feb 12 PST 2016 
Fri Feb 12 PST 2016 
Fri Feb 12 PST 2016 
Imins, 
53sec 
Imins, 
23sec 
19sec 
6sec 
8sec 
Start Time 
0001 
Node 
Fri Feb 12 PST 2016 
localhost:8042 
Firefox automatically sends some data to Mozilla so that we can improve your experience. 
Logged in as: dr.who 
Job Overview 
Logs 
logs 
Choose What I Share 

Check the logs - click on the logs link:

achine generated alternative text:
Mozilla Firefox 
http://local...307 0001/joe x Welcome to CentOS 
Åter 
localhost 
q Search 
Application 
Tools 
Configuration 
Local Iogs 
Server 
nal 
stacks 
Server 
metrics 
Log Type: stderr 
Log Upload Time: Fri Feb 12 -0800 2016 
Log Length: 2257 
Feb 12, 21316 PM com.google. inject. serv let. InternalServletModule$BackwardsCompatibleServletContextProvider get 
WARNING: You are attempting to use a deprecated API (specifically, attempting to @lnject ServletContext inside an eagerly created 
Feb 12, 2016 PM com. sun.jersey.guice. spi. container.GuiceComponentProviderFactory register 
INFO: Register ing org.apache. hadoop.mapreduce.v2.app.webapp.JAXBContextResolver as a provider c lass 
Feb 12, 21316 PM com. sun.jersey.guice. spi. container.GuiceComponentProviderFactory register 
INFO: Register ing org.apache. hadoop.yarn.webapp.GenericExceptionHandler as a provider c lass 
Feb 12, 21316 PM com. sun.jersey.guice. spi. container.GuiceComponentProviderFactory register 
INFO: Register ing org.apache. hadoop.mapreduce.v2.app.webapp.AMWebServices as a root resource c lass 
Feb 12, 21316 PM com. sum. jersey. server. impl.application.WebApplicationImpl _initiate 
INFO: Initiating Jersey application, version '-lersey: 1.9 09/02/2011 11:17 AM' 
Feb 12, 21316 PM com. sun.jersey.guice. spi. container.GuiceComponentProviderFactory getComponentProvider 
INFO: Binding org.apache. hadoop.mapreduce.v2.app.webapp.JAXBContextResolver to GuiceManagedComponentProvider Wit 
Feb 12, 2016 PM com. sun.jersey.guice. spi. container.GuiceComponentProviderFactory getComponentProvider 
INFO: Binding org.apache. hadoop.yarn.webapp.GenericExceptionHandler to GuiceManagedComponentProvider with the sc 
Feb 12, 2016 PM com. sun.jersey.guice. spi. container.GuiceComponentProviderFactory getComponentProvider 
INFO: Binding org.apache. hadoop.mapreduce.v2.app.webapp.AMWebServices to GuiceManagedComponentProvider with the 
log4j :WARN No appenders could be found for logger (org.apache. hadoop. i pc. Server). 
log4j :WARN Please initialize the log4j system proper ly. 
log4j :WARN See http://logging.apache.org/log4j/1.2/faq. html#noconfig for more info. 
Log Type: stdout 
Log Upload Time: Fri Feb 12 -0800 2016 
Log Length: O 
Save Page As 
View Backgrc 
Select All 
View Page so 
View Page Inf 
Inspect Elem 

Check the output:

[joe@localhost mrv2\_jdk8]$ **hadoop fs -cat nasa\_output/part-r-00000 | tail -10**

28/Jul/1995:05:00:00 -0400        916

28/Jul/1995:06:00:00 -0400        1086

28/Jul/1995:07:00:00 -0400        1616

28/Jul/1995:08:00:00 -0400        2543

28/Jul/1995:09:00:00 -0400        2741

28/Jul/1995:10:00:00 -0400        3555

28/Jul/1995:11:00:00 -0400        3570

28/Jul/1995:12:00:00 -0400        3505

28/Jul/1995:13:00:00 -0400        1905

12/Feb/2016:21:55:14 -0800        1

[joe@localhost mrv2\_jdk8]$

Make sure you understand why there is one date from today! :-)