Learning Hadoop 2

Randal Scott King

Section 4

Using MapReduce and Pig

[PACKT] Udemy VIDEO

In this Section, we are going to take a look at...

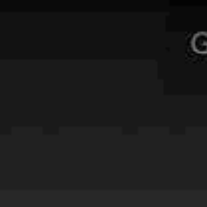
- Coding a “Word Count” application in MapReduce
- Coding a “Word Count” application in Pig
- Using Pig for common ETL (Extract, Transform, and Load) tasks
- Creating a User-defined Function (UDF) in Pig

Learning Hadoop 2

Randal Scott King

Video 4.1

Coding "Word Count" in MapReduce



In this Video, we are going to take a look at...

- The Eclipse IDE for Java
- Exporting a .jar
- Running jars in MapReduce

Applications Places System — Wed Sep 16, 10:33 AM cloudera — Go to Dashboard

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera | Hue | Hadoop | HBase | Impala | Spark | Solr | Oozie | Cloudera Manager | Getting Started

HUE | Query Editors | Data Browsers | Workflows | Search | Security | File Browser | Job Browser | cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera | History | Trash

Name	Size	User	Group	Permissions	Date
..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:32 AM
.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Show 45 of 3 items — Page 1 of 1

Applications Places System Wed Sep 16, 10:33 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera

History Trash

Name	Size	User	Group	Permissions	Date
..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:32 AM
.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Show 45 of 2 items

Page 1 of 1

quickstart.cloudera:8888/filebrowser/view/user/cloudera#

Hue - File Browser - M...

PACKT

Applications Places System Wed Sep 16, 10:33 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/fileb

Cloudera Hue Hadoop

HUE Query Editors cloudera

File Browser

Search for file name

Upload New

Home / user / clo

History Trash

	Name	Date
	..	September 06, 2015 06:19 PM
	.	September 16, 2015 10:32 AM
	.Trash	September 05, 2015 05:21 PM
	MoreTransactions.csv	September 05, 2015 11:04 AM
	Transactions.csv	September 05, 2015 11:04 AM

Show 45 of 3 items

1 of 1

Hue - File Browser - M...

File Upload

[PACKT]

Applications Places System

Wed Sep 16, 10:33 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/fileb

Cloudera Hue Hadoop

HUE Query Editors cloudera

File Browser

Search for file name

Upload New

Home / user / clo

History Trash

Name

J

.

.Trash

MoreTransactions.csv

Transactions.csv

Date

September 06, 2015 06:19 PM

September 16, 2015 10:32 AM

September 05, 2015 05:21 PM

September 05, 2015 11:04 AM

September 05, 2015 11:04 AM

Show 45 of 3 items

1 of 1

Hue - File Browser - M...

File Upload

cloudera **Downloads**

Places

- Search
- Recently Used
- cloudera
- Desktop
- File System
- Documents
- Music
- Pictures
- Videos
- Downloads

Name	Size	Modified
Customers.csv	600 bytes	09/05/2015
MoreTransactions.csv	272 bytes	09/05/2015
Names.txt	74 bytes	10:25
Transactions.csv	981 bytes	09/05/2015

Add Remove

All Files

Cancel Open

PACKT

Applications Places System Wed Sep 16, 10:33 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera | History | Trash

	Name	Size	User	Group	Permissions	Date
	..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
	.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:33 AM
	.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
	MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
	Names.txt	74 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:33 AM
	Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Show 45 of 4 items

Page 1 of 1

Hue - File Browser - M...

Applications Places System Wed Sep 16, 10:33 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / **cloudera** History Trash

Name	Size	User	Group	Permissions	Date
..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:33 AM
.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
Names.txt	74 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:33 AM
Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Show 45 of 4 items

Page 1 of 1

Hue - File Browser - M...

Applications Places System Wed Sep 16, 10:33 AM cloudera

Computer

cloudera's Home

Trash

Launch Cloudera Enterprise (trial)

Launch Cloudera Express

Eclipse

eclipse LUNA

cloudera® Ask Bigger Questions

[Hue - File Browser - M...

Applications Places System

Wed Sep 16, 10:34 AM cloudera

Java - training/src/StubDriver.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

New Shift+Alt+N

Open File...

Close Ctrl+W

Close All Shift+Ctrl+W

Save Ctrl+S

Save As...

Save All Shift+Ctrl+S

Revert

Move...

Rename... F2

Refresh F5

Convert Line Delimiters To

Print... Ctrl+P

Switch Workspace

Restart

Import...

Export...

Properties Alt+Enter

1 StubDriver.java [training/src]

2 WordCount.java [WordCount/src]

Exit

Java Project

Project...

Package

Class

Interface

Enum

Annotation

Source Folder

Java Working Set

Folder

File

Untitled Text File

JUnit Test Case

Task

Example...

Other... Ctrl+N

Quick Access

Java

b;

gs) throws Exception {

re passed from the command line.

river <input dir> <output dir>\n");

our job's configuration.

Javadoc Declaration Console

o display at this time.

Writable Smart Insert 1 : 1

[Hue - File Browser - ...

Java - training/src/Stub...

Applications Places System

Wed Sep 16, 10:34 AM cloudera

File Edit Source Refactor Navig

Package Explorer

training

Quick Access

Java

New Java Project

Create a Java Project

Create a Java project in the workspace or in an external location.

Project name: WordCount

☑ Use default location

Location: /home/cloudera/workspace/WordCount

Browse...

JRE

◉ Use an execution environment JRE: JavaSE-1.7

○ Use a project specific JRE: jdk1.7.0_67-cloudera

○ Use default JRE (currently 'jdk1.7.0_67-cloudera') Configure JREs...

Project layout

○ Use project folder as root for sources and class files

◉ Create separate folders for sources and class files Configure default...

Working sets

☐ Add project to working sets

< Back Next > Cancel Finish

[Hue - File Browser - ...

Java - training/src/Stub...

Coding "Word Count" in MapReduce

Section 4, Lecture 12

Java - Eclipse

File Edit Sc ... w Help

Quick Access | Java

Package E:

training

WordCo

New
Go Into
Open in New Window
Open Type Hierarchy F4
Show In Shift+Alt+W
Copy Ctrl+C
Copy Qualified Name
Paste Ctrl+V
Delete Delete
Remove from Context Shift+Ctrl+Alt+Down
Build Path
Source Shift+Alt+S
Refactor Shift+Alt+T
Import...
Export...
Refresh F5
Close Project
Close Unrelated Projects
Assign Working Sets...
Debug As
Run As
Validate
Team
Compare With
Restore from Local History...
Properties Alt+Enter

Declaration Console

time.

1:59 /5:55

Properties for WordCount

type filter text

- Resource
- Builders
- Java Build Path
- Java Code Style
- Java Compiler
- Java Editor
- Javadoc Location
- Project References
- Run/Debug Settings
- Task Repository
- Task Tags
- Validation
- WikiText

Java Build Path

Source | Projects | Libraries | Order and Export

JARs and class folders on the build path:

- JRE System Library [JavaSE-1.7]

Add JARs...
Add External JARs...
Add Variable...
Add Library...
Add Class Folder...
Add External Class Folder...
Edit...
Remove
Migrate JAR File...

Cancel OK

Applications Places System
Wed Sep 16, 10:35 AM cloudera
Go to Dashboard
JAR Selection
usr lib hadoop client
Places
Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads
Name
Size
Modified
bin 06/09/2015
client 06/09/2015
client-0.20 06/09/2015
cloudera 06/09/2015
etc 06/09/2015
lib 06/09/2015
libexec 06/09/2015
sbin 06/09/2015
hadoop-annotations.jar 21.0 KB 05/19/2015
hadoop-annotations-2.6.0-cdh5.4.2.jar 21.0 KB 05/19/2015
hadoop-auth.jar 71.4 KB 05/19/2015
hadoop-auth-2.6.0-cdh5.4.2.jar 71.4 KB 05/19/2015
hadoop-aws.jar 13.7 MB 05/19/2015
hadoop-aws-2.6.0-cdh5.4.2.jar 13.7 MB 05/19/2015
hadoop-common.jar 3.2 MB 05/19/2015
hadoop-common-2.6.0-cdh5.4.2.jar 3.2 MB 05/19/2015
hadoop-common-2.6.0-cdh5.4.2-tests.jar 1.7 MB 05/19/2015
hadoop-common-tests.jar 1.7 MB 05/19/2015
Add
Remove
.jar;.zip
Cancel
OK
Package Expl
training
WordCoun
File Edit Sour
Java
2:09 / 5:55
[Hue - File Browser -
Java - Eclipse
Browse Q&A
Add Bookmark
Continue

Applications Places System

Wed Sep 16, 10:35 AM cloudera

JAR Selection

usr lib hadoop client

Places

- Search
- Recently Used
- cloudera
- Desktop
- File System
- Documents
- Music
- Pictures
- Videos
- Downloads

Name	Size	Modified
parquet-column.jar	871.4 KB	05/19/2015
parquet-common.jar	20.5 KB	05/19/2015
parquet-encoding.jar	272.4 KB	05/19/2015
parquet-format.jar	368.2 KB	05/19/2015
parquet-format-javadoc.jar	129.7 KB	05/19/2015
parquet-format-sources.jar	6.3 KB	05/19/2015
parquet-generator.jar	20.2 KB	05/19/2015
parquet-hadoop.jar	198.1 KB	05/19/2015
parquet-hadoop-bundle.jar	2.6 MB	05/19/2015
parquet-jackson.jar	893.1 KB	05/19/2015
parquet-pig.jar	82.6 KB	05/19/2015
parquet-pig-bundle.jar	2.6 MB	05/19/2015
parquet-protobuf.jar	48.0 KB	05/19/2015
parquet-scala_2.10.jar	32.4 KB	05/19/2015
parquet-scrooge_2.10.jar	20.4 KB	05/19/2015
parquet-test-hadoop2.jar	6.1 KB	05/19/2015
parquet-thrift.jar	181.4 KB	05/19/2015
parquet-tools.jar	61.5 KB	05/19/2015

Add Remove

.jar;.zip

Cancel OK

[Hue - File Browser - ...

Java - Eclipse

Properties for WordCount

type filter text

- Resource
- Builders
- Java Build Path
- Java Code Style
- Java Compiler
- Java Editor
- Javadoc Location
- Project References
- Run/Debug Settings
- Task Repository
- Task Tags
- Validation
- WikiText

Java Build Path

Source | Projects | Libraries | Order and Export

JARs and class folders on the build path:

- hadoop-annotations-2.6.0-cdh5.4.2.jar - /usr/lib/hadoop
- hadoop-annotations.jar - /usr/lib/hadoop
- hadoop-auth-2.6.0-cdh5.4.2.jar - /usr/lib/hadoop
- hadoop-auth.jar - /usr/lib/hadoop
- hadoop-aws-2.6.0-cdh5.4.2.jar - /usr/lib/hadoop
- hadoop-aws.jar - /usr/lib/hadoop
- hadoop-common-2.6.0-cdh5.4.2-tests.jar - /usr/lib/hadoop
- hadoop-common-2.6.0-cdh5.4.2.jar - /usr/lib/hadoop
- hadoop-common-tests.jar - /usr/lib/hadoop
- hadoop-common.jar - /usr/lib/hadoop
- hadoop-nfs-2.6.0-cdh5.4.2.jar - /usr/lib/hadoop
- hadoop-nfs.jar - /usr/lib/hadoop
- parquet-avro.jar - /usr/lib/hadoop
- parquet-cascading.jar - /usr/lib/hadoop
- parquet-column.jar - /usr/lib/hadoop

Add JARs...
Add External JARs...
Add Variable...
Add Library...
Add Class Folder...
Add External Class Folder...
Edit...
Remove
Migrate JAR File...

Cancel OK

Applications Places System

Wed Sep 16, 10:35 AM cloudera

JAR Selection

usr lib hadoop client

Places

Search
Recently Used
cloudera
Desktop
File System
Documents
Music
Pictures
Videos
Downloads

Name	Size	Modified
paranamer.jar	28.9 KB	05/19/2015
paranamer-2.3.jar	28.9 KB	05/19/2015
protobuf-java.jar	521.0 KB	05/19/2015
protobuf-java-2.5.0.jar	521.0 KB	05/19/2015
servlet-api.jar	102.6 KB	05/19/2015
servlet-api-2.5.jar	102.6 KB	05/19/2015
slf4j-api.jar	25.5 KB	05/19/2015
slf4j-api-1.7.5.jar	25.5 KB	05/19/2015
slf4j-log4j12.jar	8.7 KB	05/19/2015
snappy-java.jar	972.6 KB	05/19/2015
snappy-java-1.0.4.1.jar	972.6 KB	05/19/2015
stax-api.jar	22.8 KB	05/19/2015
stax-api-1.0-2.jar	22.8 KB	05/19/2015
xmlenc.jar	14.7 KB	05/19/2015
xmlenc-0.52.jar	14.7 KB	05/19/2015
xz.jar	92.5 KB	05/19/2015
xz-1.0.jar	92.5 KB	05/19/2015
zookeeper.jar	1.3 MB	05/19/2015

Add Remove

.jar;.zip

Cancel OK

File Edit Sour

Package Expl

training

WordCoun

Java

[Hue - File Browser - ...

Java - Eclipse

Applications Places System

Wed Sep 16, 10:39 AM cloudera

Java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

New Shift+Alt+N
Open File...

Close Ctrl+W
Close All Shift+Ctrl+W

Save Ctrl+S
Save As...
Save All Shift+Ctrl+S
Revert

Move...
Rename... F2
Refresh F5
Convert Line Delimiters To

Print... Ctrl+P

Switch Workspace
Restart

Import...
Export...

Properties Alt+Enter

1 StubDriver.java [training/src]
2 WordCount.java [WordCount/src]

Exit

Java Project
Project...

Package
Class
Interface
Enum
Annotation
Source Folder
Java Working Set
Folder
File
Untitled Text File
JUnit Test Case
Task

Example...

Other... Ctrl+N

Quick Access

Java

Javadoc Declaration Console

o display at this time.

WordCount

[Hue - File Browser - ...

Java - Eclipse

Applications Places System Wed Sep 16, 10:39 AM cloudera

File Edit Source Refactor Navi

Quick Access Java

Package Explorer

training

WordCount

WordCount

New Java Class

Java Class

The use of the default package is discouraged.

Source folder: WordCount/src Browse...

Package: (default) Browse...

Enclosing type: Browse...

Name: WordCount

Modifiers: public package private protected

abstract final static

Superclass: java.lang.Object Browse...

Interfaces: Add...

Which method stubs would you like to create?

public static void main(String[] args)

Constructors from superclass

Inherited abstract methods

Do you want to add comments? (Configure templates and default value here)

Generate comments

Cancel Finish

[Hue - File Browser - ...] Java - Eclipse

Applications Places System Wed Sep 16, 10:39 AM cloudera

Java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

Problems Javadoc Declaration Console

No consoles to display at this time.

WordCount.java - WordCount/src

[Hue - File Browser - ... Java - Eclipse

Applications Places System Wed Sep 16, 10:39 AM cloudera

Java - WordCount/src/WordCount.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

WordCount.java

```
1
2 public class WordCount {
3
4 }
5
```

Problems Javadoc Declaration Console

No consoles to display at this time.

Writable Smart Insert 1 : 1

[Hue - File Browser - ... Java - WordCount/src/...

Applications Places System Wed Sep 16, 10:39 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | WordCount - Hadoop ... | LearningHadoop2/Wo...

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera | History | Trash

Name	Size	User	Group	Permissions	Date
..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:33 AM
.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
Names.txt	74 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:33 AM
Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Page 1 of 1

Hue - File Browser - Mozilla Firefox

Hue - File Browser - M... | Java - WordCount/src/...

Hadoop Wiki Login

Search Titles Text

WordCount

FrontPage RecentChanges FindPage HelpContents **WordCount**

Immutable Page Info Attachments More Actions:

WordCount Example

WordCount example reads text files and counts how often words occur. The input is text files and the output is text files, each line of which contains a word and the count of how often it occured, separated by a tab.

Each mapper takes a line as input and breaks it into words. It then emits a key/value pair of the word and 1. Each reducer sums the counts for each word and emits a single key/value with the word and sum.

As an optimization, the reducer is also used as a combiner on the map outputs. This reduces the amount of data sent across the network by combining each word into a single record.

To run the example, the command syntax is
bin/hadoop jar hadoop-*-examples.jar wordcount [-m <#maps>] [-r <#reducers>] <in-dir> <out-dir>

All of the files in the input directory (called in-dir in the command line above) are read and the counts of words in the input are written to the output directory (called out-dir above). It is assumed that both inputs and outputs are stored in HDFS (see ImportantConcepts). If your input is not already in HDFS, but is rather in a local file system somewhere, you need to copy the data into HDFS using a command like this:

bin/hadoop dfs -mkdir <hdfs-dir>
bin/hadoop dfs -copyFromLocal <local-dir> <hdfs-dir>

Applications Places System Wed Sep 16, 10:40 AM cloudera
WordCount - Hadoop Wiki - Mozilla Firefox
Hue - File Browser × WordCount - Hadoop ... × LearningHadoop2/Wo... ×
wiki.apache.org/hadoop/WordCount Search
Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

each word into a single record.

To run the example, the command syntax is
bin/hadoop jar hadoop-*-examples.jar wordcount [-m <#maps>] [-r <#reducers>] <in-dir> <out-dir>

All of the files in the input directory (called in-dir in the command line above) are read and the counts of words in the input are written to the output directory (called out-dir above). It is assumed that both inputs and outputs are stored in HDFS (see ImportantConcepts). If your input is not already in HDFS, but is rather in a local file system somewhere, you need to copy the data into HDFS using a command like this:

bin/hadoop dfs -mkdir <hdfs-dir>
bin/hadoop dfs -copyFromLocal <local-dir> <hdfs-dir>

As of version 0.17.2.1, you only need to run a command like this:
bin/hadoop dfs -copyFromLocal <local-dir> <hdfs-dir>

Word count supports generic options : see DevelopmentCommandLineOptions

Below is the standard wordcount example implemented in Java:

Toggle line numbers

```
1 package org.myorg;
2
3 import java.io.IOException;
4 import java.util.*;
5
6 import org.apache.hadoop.fs.Path;
7 import org.apache.hadoop.conf.*;
8 import org.apache.hadoop.io.*;
9 import org.apache.hadoop.mapreduce.*;
```

WordCount - Hadoop ... Java - WordCount/src/...
[PACKT]

```
1  package org.myorg;
2
3  import java.io.IOException;
4  import java.util.*;
5
6  import org.apache.hadoop.fs.Path;
7  import org.apache.hadoop.conf.*;
8  import org.apache.hadoop.io.*;
9  import org.apache.hadoop.mapreduce.*;
10 import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
11 import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
12 import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
13 import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;
14
15 public class WordCount {
16
17  public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> {
18     private final static IntWritable one = new IntWritable(1);
19     private Text word = new Text();
20
21     public void map(LongWritable key, Text value, Context context) throws IOException, InterruptedException {
22         String line = value.toString();
23         StringTokenizer tokenizer = new StringTokenizer(line);
24         while (tokenizer.hasMoreTokens()) {
25             word.set(tokenizer.nextToken());
26             context.write(word, one);
27         }
28     }
```

```
    private final static IntWritable one = new IntWritable(1);
    private Text word = new Text();

    public void map(LongWritable key, Text value, Context context) throws IOException, InterruptedException {
        String line = value.toString();
        StringTokenizer tokenizer = new StringTokenizer(line);
        while (tokenizer.hasMoreTokens()) {
            word.set(tokenizer.nextToken());
            context.write(word, one);
        }
    }

  public void run (Context context) throws IOException, InterruptedException {
        setup(context);
        while (context.nextKeyValue()) {
            map(context.getCurrentKey(), context.getCurrentValue(), context);
        }
        cleanup(context);
  }
}

public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {

    public void reduce(Text key, Iterable<IntWritable> values, Context context)
      throws IOException, InterruptedException {
        int sum = 0;
        for (IntWritable val : values) {
            sum += val.get();
        }
        context.write(key, new IntWritable(sum));
```

```
        int sum = 0;
        for (IntWritable val : values) {
            sum += val.get();
        }
        context.write(key, new IntWritable(sum));
    }
}

public int run(String[] args) throws Exception {

    Job job = Job.getInstance(new Configuration());

    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);

    job.setMapperClass(Map.class);
    job.setReducerClass(Reduce.class);

    job.setInputFormatClass(TextInputFormat.class);
    job.setOutputFormatClass(TextOutputFormat.class);

    FileInputFormat.setInputPaths(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));

    job.setJarByClass(WordCount.class);

    job.submit();
    return 0;
    }
```

Applications Places System Wed Sep 16, 10:40 AM cloudera

Java - WordCount/src/WordCount.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

WordCount.java

```
1
2 public class WordCount {
3
4 }
5
```

Problems Javadoc Declaration Console

No consoles to display at this time.

Writable Smart Insert 1 : 1

LearningHadoop2/Wor... Java - WordCount/src/...

Applications Places System Wed Sep 16, 10:40 AM cloudera

Java - WordCount/src/WordCount.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

*WordCount.java

```
66      job.setReducerClass(Reduce.class);
67
68      job.setInputFormatClass(TextInputFormat.class);
69      job.setOutputFormatClass(TextOutputFormat.class);
70
71      FileInputFormat.setInputPaths(job, new Path(args[0]));
72      FileOutputFormat.setOutputPath(job, new Path(args[1]));
73
74      job.setJarByClass(WordCount.class);
75
76      job.submit();
77      return 0;
78      }
79
80  public static void main(String[] args) throws Exception {
81     Configuration conf = new Configuration();
82     String[] otherArgs = new GenericOptionsParser(conf, args).getRemainingArgs();
83     ToolRunner.run(new WordCount(), otherArgs);
84  }
85  }
```

Problems @ Javadoc Declaration Console

No consoles to display at this time.

Writable Smart Insert 85 : 2

LearningHadoop2/Wor... Java - WordCount/src/...

Applications Places System Wed Sep 16, 10:41 AM cloudera

Java - WordCount/src/WordCount.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

*WordCount.java

```
27
28    public void map(LongWritable key, Text value, Context context) throws IOException, InterruptedException
29        String line = value.toString();
30        StringTokenizer tokenizer = new StringTokenizer(line);
31        while (tokenizer.hasMoreTokens()) {
32            word.set(tokenizer.nextToken());
33            context.write(word, one);
34        }
35    }
36
37  public void run (Context context) throws IOException, InterruptedException {
38        setup(context);
39        while (context.nextKeyValue()) {
40            map(context.getCurrentKey(), context.getCurrentValue(), context);
41        }
42        cleanup(context);
43    }
44  }
45
46  public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
```

Problems @ Javadoc Declaration Console

No consoles to display at this time.

Writable Smart Insert 35 : 6

LearningHadoop2/Wor... Java - WordCount/src/...

Applications Places System Wed Sep 16, 10:41 AM cloudera

Java - WordCount/src/WordCount.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

*WordCount.java

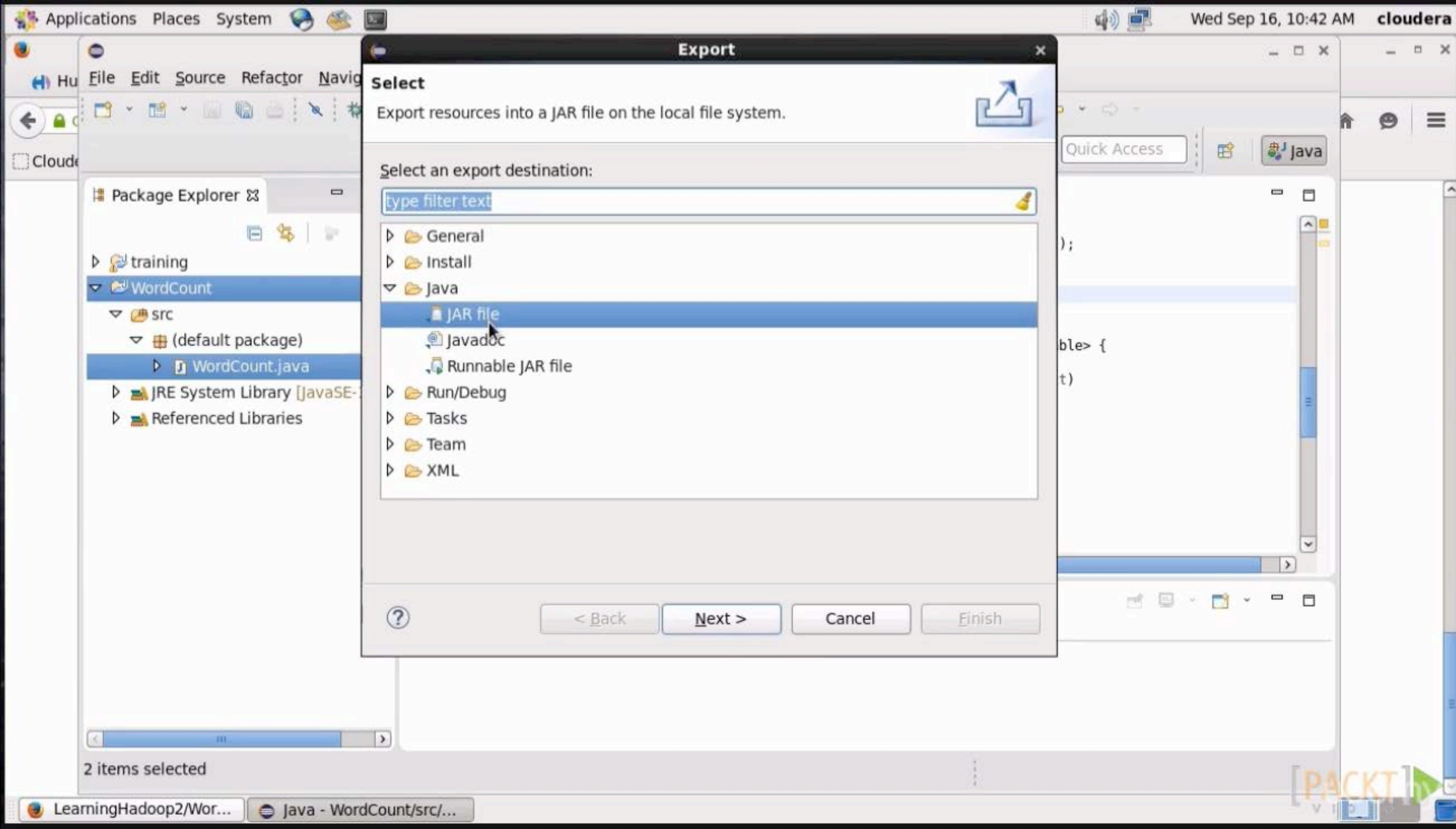
```
39          while (context.nextKeyValue()) {
40              map(context.getCurrentKey(), context.getCurrentValue(), context);
41          }
42          cleanup(context);
43      }
44  }
45
46  public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
47
48      public void reduce(Text key, Iterable<IntWritable> values, Context context)
49        throws IOException, InterruptedException {
50          int sum = 0;
51          for (IntWritable val : values) {
52              sum += val.get();
53          }
54          context.write(key, new IntWritable(sum));
55      }
56  }
57
58  public int run(String[] args) throws Exception {
```

Problems @ Javadoc Declaration Console

No consoles to display at this time.

Writable Smart Insert 35 : 6

LearningHadoop2/Wor... Java - WordCount/src/...


Applications Places System
Wed Sep 16, 10:42 AM cloudera
File Edit Source Refactor Navig
Export
Select
Export resources into a JAR file on the local file system.
Select an export destination:
type filter text
General
Install
Java
JAR file
Javadoc
Runnable JAR file
Run/Debug
Tasks
Team
XML
< Back
Next >
Cancel
Finish
Package Explorer
training
WordCount
src
(default package)
WordCount.java
JRE System Library [JavaSE-
Referenced Libraries
Quick Access
Java
2 items selected
LearningHadoop2/Wor...
Java - WordCount/src/...

Applications Places System Wed Sep 16, 10:42 AM cloudera

File Edit Source Refactor Navig

Package Explorer

training
WordCount
src
(default package)
WordCount.java
JRE System Library [JavaSE-
Referenced Libraries

2 items selected

Quick Access Java

);

ble> {

t)

JAR Export

JAR File Specification

Define which resources should be exported into the JAR.

Select the resources to export:

training
WordCount

.classpath
.project

Export generated class files and resources
Export all output folders for checked projects
Export Java source files and resources
Export refactorings for checked projects. Select refactorings...

Select the export destination:

JAR file: /home/cloudera/WordCount.jar Browse...

Options:

Compress the contents of the JAR file
Add directory entries
Overwrite existing files without warning

< Back Next > Cancel Finish

LearningHadoop2/Wor... Java - WordCount/src/...

[PACKT]

Applications Places System Wed Sep 16, 10:42 AM cloudera

File Edit Source Refactor Navig

Package Explorer

training
WordCount
src
(default package)
WordCount.java
JRE System Library [JavaSE-
Referenced Libraries

2 items selected

Quick Access Java

);

ble> {

t)

JAR Export

JAR File Specification

Define which resources should be exported into the JAR.

Select the resources to export:

- training
- WordCount

- .classpath
- .project

- [x] Export generated class files and resources
- [ ] Export all output folders for checked projects
- [ ] Export Java source files and resources
- [ ] Export refactorings for checked projects. Select refactorings...

Select the export destination:

JAR file: /home/cloudera/WordCount.jar Browse...

Options:

- [x] Compress the contents of the JAR file
- [ ] Add directory entries
- [ ] Overwrite existing files without warning

< Back Next > Cancel Finish

LearningHadoop2/Wor... Java - WordCount/src/...

[PACKT]

Applications Places System

Wed Sep 16, 10:42 AM cloudera

File Edit Source Refactor Navig

Package Explorer

training

WordCount

src

(default package)

WordCount.java

JRE System Library [JavaSE-

Referenced Libraries

2 items selected

JAR Export

JAR File Specification

Define whic

Select the r

Select the e

JAR file:

owse...

Options:

Compre

Add directory entries

Overwrite existing files without warning

< Back Next > Cancel Finish

Save Modified Resources

Some modified resources must be saved before this operation.

WordCount.java

Cancel OK

Quick Access

Java

);

ble> {

t)

LearningHadoop2/Wor...

Java - WordCount/src/...

Applications Places System

Wed Sep 16, 10:42 AM cloudera

File Edit Source Refactor Navig

Package Explorer

training

WordCount

src

(default package)

WordCount.java

JRE System Library [JavaSE

Referenced Libraries

JAR Export

JAR File Specification

Define which resources should be exported into the JAR.

Select the resources to export:

training

WordCount

.classpath

.project

JAR Export

JAR file: /home/cloudera/WordCount.jar Browse...

Options:

Compress the contents of the JAR file

Add directory entries

Overwrite existing files without warning

< Back Next > Cancel Finish

Quick Access Java

);

le> {

)

LearningHadoop2/Wor...

Java - WordCount/src/...

[PACKT]

Applications Places System Wed Sep 16, 10:43 AM cloudera

File Edit Source Refactor Navig

Package Explorer

training
WordCount
src
(default package)
WordCount.java
JRE System Library [JavaSE
Referenced Libraries

Quick Access Java

);

le> {

)

JAR Export

JAR File Specification

Define which resources should be exported into the JAR.

Select the resources to export:

training
WordCount

.classpath
.project

JAR Export

JAR export finished with warnings. See details for additional information.

Details >> OK

JAR file: /home/cloudera/WordCount.jar Browse...

Options:

Compress the contents of the JAR file
Add directory entries
Overwrite existing files without warning

< Back Next > Cancel Finish

LearningHadoop2/Wor... Java - WordCount/src/...

PACKT

Applications Places System

Wed Sep 16, 10:43 AM cloudera

Java - WordCount/src/WordCount.java - Eclipse

File Edit Source Refactor Navigate Search Project Run Window Help

Quick Access | Java

Package Explorer

- training
- WordCount
  - src
    - (default package)
      - WordCount.java
  - JRE System Library [JavaSE-1.7]
  - Referenced Libraries

WordCount.java

```
39          while (context.nextKeyValue()) {
40              map(context.getCurrentKey(), context.getCurrentValue(), context);
41          }
42          cleanup(context);
43      }
44  }
45
46  public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
47
48      public void reduce(Text key, Iterable<IntWritable> values, Context context)
49        throws IOException, InterruptedException {
50          int sum = 0;
51          for (IntWritable val : values) {
52              sum += val.get();
53          }
54          context.write(key, new IntWritable(sum));
55      }
56  }
57
58  public int run(String[] args) throws Exception {
```

Problems @ Javadoc Declaration Console

No consoles to display at this time.

LearningHadoop2/Wor... | Java - WordCount/src/...

Applications Places System Wed Sep 16, 10:45 AM cloudera

Computer

cloudera's Home

Trash

Launch Cloudera Enterprise (trial)

Launch Cloudera Express

Eclipse

cloudera@quickstart:~

File Edit View Search Terminal Help

```
[cloudera@quickstart ~]$ hadoop jar
```

cloudera®

Ask Bigger Questions

[LearningHadoop2/Wor... [Java - WordCount/src/... cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:45 AM cloudera

Computer

cloudera's Home

Trash

Launch Cloudera Enterprise (trial)

Launch Cloudera Express

Eclipse

cloudera@quickstart:~

File Edit View Search Terminal Help

```
[cloudera@quickstart ~]$ hadoop jar /home/cloudera/WordCount.jar W
```

cloudera®

Ask Bigger Questions

[LearningHadoop2/Wor... [Java - WordCount/src/... cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:45 AM cloudera

Computer

cloudera's Home

Trash

Launch Cloudera Enterprise (trial)

Launch Cloudera Express

Eclipse

cloudera@quickstart:~

File Edit View Search Terminal Help

```
[cloudera@quickstart ~]$ hadoop jar /home/cloudera/WordCount.jar WordCount Names
.txt output
```

cloudera®

Ask Bigger Questions

[LearningHadoop2/Wor... [Java - WordCount/src/... cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:46 AM cloudera

Computer

cloudera's Home

Trash

Launch Cloudera Enterprise (trial)

Launch Cloudera Express

Eclipse

cloudera@quickstart:~

File Edit View Search Terminal Help

```
[cloudera@quickstart ~]$ hadoop jar /home/cloudera/WordCount.jar WordCount Names
.txt output
15/09/16 10:46:05 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0
:8032
15/09/16 10:46:06 WARN mapreduce.JobSubmitter: Hadoop command-line option parsin
g not performed. Implement the Tool interface and execute your application with
ToolRunner to remedy this.
15/09/16 10:46:06 INFO input.FileInputFormat: Total input paths to process : 1
15/09/16 10:46:06 INFO mapreduce.JobSubmitter: number of splits:1
15/09/16 10:46:07 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_14
42364086378_0007
15/09/16 10:46:07 INFO impl.YarnClientImpl: Submitted application application_14
42364086378_0007
15/09/16 10:46:07 INFO mapreduce.Job: The url to track the job: http://quickstar
t.cloudera:8088/proxy/application_1442364086378_0007/
[cloudera@quickstart ~]$
```

cloudera®

Ask Bigger Questions

[LearningHadoop2/Wor... [Java - WordCount/src/... cloudera@quickstart:~

```
65      job.setMapperClass(Map.class);
66      job.setReducerClass(Reduce.class);
67
68      job.setInputFormatClass(TextInputFormat.class);
69      job.setOutputFormatClass(TextOutputFormat.class);
70
71      FileInputFormat.setInputPaths(job, new Path(args[0]));
72      FileOutputFormat.setOutputPath(job, new Path(args[1]));
73
74      job.setJarByClass(WordCount.class);
75
76      job.submit();
77      return 0;
78      }
79
80  public static void main(String[] args) throws Exception {
81      Configuration conf = new Configuration();
82      String[] otherArgs = new GenericOptionsParser(conf, args).getRemainingArgs();
83      ToolRunner.run(new WordCount(), otherArgs);
84  }
85  }
```

Applications Places System Wed Sep 16, 10:46 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | Hue - Quick Start

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser 1 cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / **cloudera** History Trash

Name	Size	User	Group	Permissions	Date
..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:33 AM
.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
Names.txt	74 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:33 AM
Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Show 45 of 4 items

Page 1 of 1

Hue - File Browser - M... | [Java - WordCount/src/... | cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:46 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | Hue - Quick Start

quickstart.cloudera:8888/filebrowser/view/user/cloudera

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser 1 cloudera

Manage jobs

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera | History | Trash

Name	Size	User	Group	Permissions	Date
..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:33 AM
.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
Names.txt	74 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:33 AM
Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM

Show 45 of 4 items

Page 1 of 1

quickstart.cloudera:8888/jobbrowser

Hue - File Browser - M... | [Java - WordCount/src/... | cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:46 AM cloudera

Hue - Job Browser - Mozilla Firefox

Hue - Job Browser | Hue - Quick Start

quickstart.cloudera:8888/jobbrowser/ Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser 1 cloudera

Manage jobs

Job Browser

Username cloudera Text Search for text

Succeeded Running Failed Killed

Logs	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted	
	1442364086378_0007	WordCount.jar	RUNNING	cloudera	5%	5%	root.cloudera	N/A	17s	09/16/15 10:46:07	Kill
	1442364086378_0005	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	28s	09/16/15 06:00:56	
	1442364086378_0004	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	27s	09/16/15 05:22:09	
	1442364086378_0003	wordcount	FAILED	cloudera	100%	100%	root.cloudera	N/A	43s	09/15/15 18:13:34	

Showing 1 to 4 of 4 entries ← Previous 1 Next →

quickstart.cloudera:8888/jobbrowser

Hue - Job Browser - M... [Java - WordCount/src/... cloudera@quickstart:~

[PACKT]

Applications Places System Wed Sep 16, 10:46 AM cloudera

Hue - Job Browser - Mozilla Firefox

Hue - Job Browser | Hue - Quick Start

quickstart.cloudera:8888/jobbrowser/ Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser 1 cloudera

Job Browser

Username cloudera Text Search for text

Succeeded Running Failed Killed

Logs	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted	
	1442364086378_0007	WordCount.jar	RUNNING	cloudera	5%	5%	root.cloudera	N/A	21s	09/16/15 10:46:07	Kill
	1442364086378_0005	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	28s	09/16/15 06:00:56	
	1442364086378_0004	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	27s	09/16/15 05:22:09	
	1442364086378_0003	wordcount	FAILED	cloudera	100%	100%	root.cloudera	N/A	43s	09/15/15 18:13:34	

Showing 1 to 4 of 4 entries ← Previous 1 Next →

Hue - Job Browser - M... [Java - WordCount/src/... cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:47 AM cloudera

Hue - Job Browser - Mozilla Firefox

Hue - Job Browser | Hue - Quick Start

quickstart.cloudera:8888/jobbrowser/ Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

Job Browser

Username cloudera Text Search for text

Succeeded Running Failed Killed

Logs	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted
	1442364086378_0007	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	35s	09/16/15 10:46:07
	1442364086378_0005	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	28s	09/16/15 06:00:56
	1442364086378_0004	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	27s	09/16/15 05:22:09
	1442364086378_0003	wordcount	FAILED	cloudera	100%	100%	root.cloudera	N/A	43s	09/15/15 18:13:34

Showing 1 to 4 of 4 entries ← Previous 1 Next →

Hue - Job Browser - M... [Java - WordCount/src/... cloudera@quickstart:~

[PACKT]

Applications Places System Wed Sep 16, 10:47 AM cloudera

Hue - Job Browser - Mozilla Firefox

Hue - Job Browser | Hue - Quick Start

quickstart.cloudera:8888/jobbrowser/ Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

Manage HDFS

Job Browser

Username cloudera Text Search for text

Succeeded Running Failed Killed

Logs	ID	Name	Status	User	Maps	Reduces	Queue	Priority	Duration	Submitted
	1442364086378_0007	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	35s	09/16/15 10:46:07
	1442364086378_0005	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	28s	09/16/15 06:00:56
	1442364086378_0004	WordCount.jar	SUCCEEDED	cloudera	100%	100%	root.cloudera	N/A	27s	09/16/15 05:22:09
	1442364086378_0003	wordcount	FAILED	cloudera	100%	100%	root.cloudera	N/A	43s	09/15/15 18:13:34

Showing 1 to 4 of 4 entries ← Previous 1 Next →

quickstart.cloudera:8888/filebrowser

Hue - Job Browser - M... [Java - WordCount/src/... cloudera@quickstart:~

[PACKT]

Applications Places System Wed Sep 16, 10:47 AM cloudera

Hue - File Browser - Mozilla Firefox

Hue - File Browser | Hue - Quick Start

quickstart.cloudera:8888/filebrowser/

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera | History | Trash

	Name	Size	User	Group	Permissions	Date
	..		hdfs	supergroup	drwxr-xr-x	September 06, 2015 06:19 PM
	.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:46 AM
	.Trash		cloudera	cloudera	drwxr-xr-x	September 05, 2015 05:21 PM
	MoreTransactions.csv	272 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
	Names.txt	74 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:33 AM
	Transactions.csv	981 bytes	cloudera	cloudera	-rw-r--r--	September 05, 2015 11:04 AM
	output		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:46 AM

Show 45 of 5 items

Page 1 of 1

quickstart.cloudera:8888/filebrowser/#

Hue - File Browser - M... | [Java - WordCount/src/... | cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:47 AM cloudera

Hue - File Browser - Mozilla Firefox

Connecting... | Hue - Quick Start

quickstart.cloudera:8888/filebrowser/#/user/cloudera/output

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

Search for file name | Actions | Move to trash | Upload | New

Home / user / cloudera / **output** History Trash

Name	Size	User	Group	Permissions	Date
..		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:46 AM
.		cloudera	cloudera	drwxr-xr-x	September 16, 2015 10:46 AM
_SUCCESS	0 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:46 AM
part-r-00000	35 bytes	cloudera	cloudera	-rw-r--r--	September 16, 2015 10:46 AM

Show 45 of 2 items

Page 1 of 1

quickstart.cloudera:8888/filebrowser/#

Hue - File Browser - M... | [Java - WordCount/src/... | cloudera@quickstart:~

Applications Places System Wed Sep 16, 10:47 AM cloudera

Hue - File Browser - part-r-00000 - File Viewer - Mozilla Firefox

Hue - File Browser - ... | Hue - Quick Start

quickstart.cloudera:8888/filebrowser/view/user/cloudera/output/part-r-00000

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

HUE Query Editors Data Browsers Workflows Search Security File Browser Job Browser cloudera

File Browser

ACTIONS

View as binary

Edit file

Download

View file location

Refresh

INFO

Last modified
Sept. 16, 2015 10:46 a.m.

User
cloudera

Group
cloudera

Home / user / cloudera / output / **part-r-00000**

Page 1 of 1

```
Aaron   4
Laura   1
Meredith        3
Sarah   3
```

Hue - File Browser - pa... | [Java - WordCount/src/... | cloudera@quickstart:~

Next Video

Coding "Word Count" in Pig

[PACKT] Udemy VIDEO