

## Índice

Exercícios de laboratórios Hadoop.....	2
HACK70 – Rodando um MapReduce no Hadoop ( PI ).....	2
HACK71 – Rodando um MapReduce Wordcount.....	2
HACK72 – Listando os Jobs.....	3
HACK73 – Instalando a IDE Eclipse.....	4
HACK74 – Adquirindo bibliotecas de desenvolvimento Hadoop.....	4
HACK75 – Instalando o Plugin de Hadoop para Eclipse.....	4
HACK76 – Criando um Projeto Java Hadoop e importando Bibliotecas.....	4
HACK77 – Crie uma Classe MAP .....	5

## Exercícios de laboratórios Hadoop

Cria uma pasta em seu ambiente para armazenar estes Hacks ( Exercícios práticos de laboratórios ), os Hacks serão inicialmente usados e criados para a prática dos conhecimentos adquiridos no curso e posteriormente podem ser uma base de conhecimento para consulta.

### ***HACK70 – Rodando um MapReduce no Hadoop ( PI ).***

1. Acesse sua VM com o usuário hadoop
2. A base do hadoop traz alguns exemplos de mapreduce desenvolvidos , vamos aprender como executar um map reduce, para teste vamos usar o calculo de PI
3. com o usuário hadoop execute o comando abaixo

```
cd /usr/local/hadoop
hadoop jar ./share/hadoop/mapreduce/hadoop-mapreduce-examples-2.2.0.jar
pi 2 5
```

4. Será apresentado o processo de MapReduce

```
Wrote input for Map #0
Wrote input for Map #1
Starting Job
14/01/19 11:28:20 INFO client.RMProxy: Connecting to ResourceManager at /
0.0.0.0:8032
14/01/19 11:28:21 INFO input.FileInputFormat: Total input paths to process : 2
14/01/19 11:28:22 INFO mapreduce.JobSubmitter: number of splits:2
.....
Job Finished in 71.477 seconds
Estimated value of Pi is 3.60000000000000000000
```

### ***HACK71 – Rodando um MapReduce Wordcount.***

1. Este MapReduce de exemplo calcula quantas palavras existem em um determinado arquivo
2. Vamos utilizar o arquivo hack\_test.txt que esta no diretório /user/hadoop/hacks/input que copiamos em hacks anteriores. O WordCount precisa de um diretório de “dados” que é

nosso input e outro para “imprimir” os resultados que usaremos o /user/hadoop/hacks/output

3. Para executar o Wordcount execute os comandos a seguir

```
cd /usr/local/hadoop
hadoop jar ./share/hadoop/mapreduce/hadoop-mapreduce-examples-2.2.0.jar
wordcount /user/hadoop/hacks/input /user/hadoop/hacks/output
```

**Atenção:** Caso já tenha um diretório com o nome output será emitido erro (exclua o mesmo pois estava programado em hacks anteriores para fazer isso)

4. Este MapReduce vai criar o diretório output e adicionar o resultado nele, você pode acessar via browser para facilitar sua visualização:

http://192.168.0.102:50075/browseBlock.jsp?  
blockId=1073741860&blockSize=56&genstamp=1036&filename=%2Fuser%2Fhadoop%2Fhacks  
%2Foutput  
%2Fpart-r-000000&datanodePort=50010&namenodeInfoPort=50070&nnaddr=127.0.0.1:9000

**File: /user/hadoop/hacks/output/part-r-00000**

Goto :

[Go back to dir listing](#)  
[Advanced view/download options](#)

```
distribuido" 1
filesystem 1
hadoop 1
um 1
é 1
"0 1
```

5. Veja que ele contou a ocorrência de palavras em nosso arquivo.

## **HACK72 – Listando os Jobs.**

1. O comando abaixo lista os jobs que estão em execução no hadoop

```
hadoop job -list
```

### ***HACK73 – Instalando a IDE Eclipse.***

5. O Eclipse pode ser adquirido em <http://www.eclipse.org/>
6. O instalador já se encontra em nosso material de treinamento ( hadoop-fundamental-ambientelivre.zip ) no diretório eclipse.
7. Crie um diretório em seu sistema /opt/hadoop-dev
8. copie o arquivo eclipse-standard-XXXX-linuxXXXXtar.gz ( se linux ) para o diretório /opt/hadoop-dev , e o arquivo
9. Descompacte ou instale o Eclipse em /opt/hadoop-dev/eclipse
10. rode o eclipse.exe em windows e /eclipse.sh em Linux

### ***HACK74 – Adquirindo bibliotecas de desenvolvimento Hadoop.***

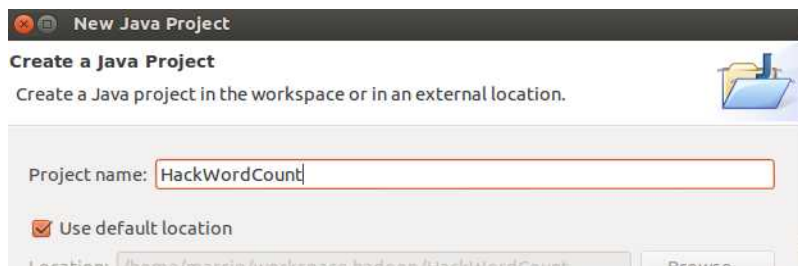
1. Para o desenvolvimento de MapReduce vamos necessitar das bibliotecas do Hadoop ( JARs) que podem ser baixadas de <http://hadoop.apache.org/releases.html>
2. Os mesmos se encontra no zip disponibilizado para o treinamento no diretório hadoop.
3. Crie o diretório opt/hadoop-dev/libs e descompacte as bibliotecas em hadoop-2.2.0.tar.gz e hadoop-2.2.0-src.tar.gz no mesmo.

### ***HACK75 – Instalando o Plugin de Hadoop para Eclipse.***

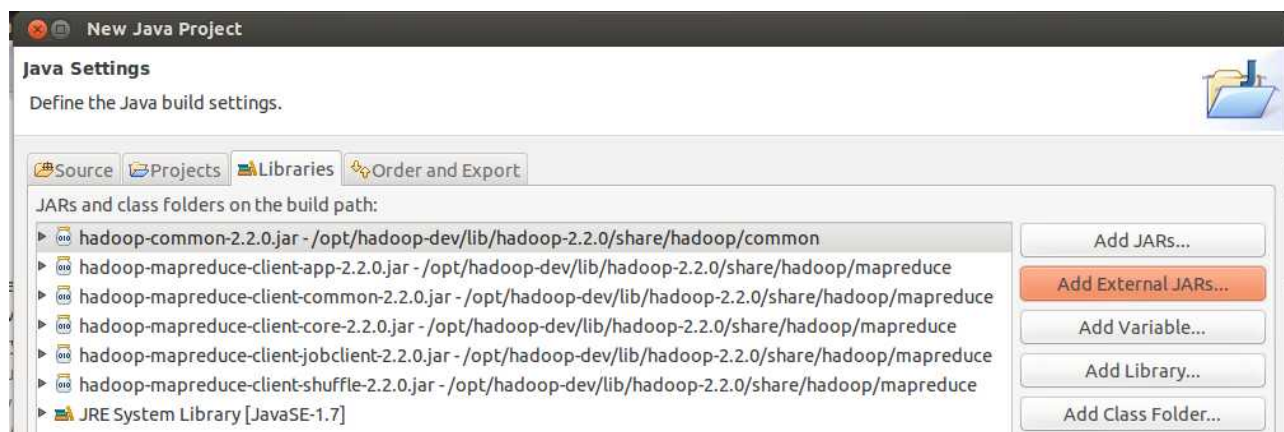
1. Opcional.
2. O mesmo se encontra no zip disponibilizado para o treinamento no diretório plugin-eclipse com o nome hadoop-eclipse-plugin-1.0.4.jar
3. <http://www.youtube.com/watch?v=IBzcRpN9Zo4>

### ***HACK76 – Criando um Projeto Java Hadoop e importando Bibliotecas.***

1. Inicie o Eclipse
2. Crie um novo projeto nas opções File → New – Java Project
3. De o nome de nosso Project como **HackWordCount**



4. Clique em Next e depois na Aba Libraries para referenciamos as bibliotecas do hadoop.
5. Clique em Add External Jars e selecione os Jar apontados abaixo :



6. Clique em Finish.

## **HACK77 – Crie uma Classe MAP e Reduce do WordCount2 .**

1. Crie uma classe Java com o nome WordCount2.Java
2. Inclua a mesma o código Abaixo
3. <http://wiki.apache.org/hadoop/WordCount>
4. Salve o Arquivo