

Índice

Exercícios de laboratórios Hadoop.....	2
HACK50 – Listando arquivo do HDFS.....	2
HACK51 – Criando um diretório Filesystem do Hadoop.....	2
HACK53 – Copiando um arquivo para Filesystem do Hadoop (HDFS).....	2
HACK54 – Mostrando o conteúdo de arquivo no HDFS.....	3
HACK55 – Copiando arquivos entre diretórios no HDFS.....	3
HACK56 – Apagando um arquivo no HDFS.....	4
HACK57 – Apagando um diretório do HDFS.....	4
HACK58 – Interfaces Web do HDFS.....	4
HACK59 – Consultando a lista de comando do HDFS.....	5
Referências.....	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.

HACK50 – Listando arquivo do HDFS.

1. Para listar os arquivos que estão em nosso filesystem Hadoop use o comando a seguir:

```
hdfs dfs -ls /
```

2. lembre que nosso filesystem foi formatado e não temos nenhum arquivo nele apenas a estrutura básica pós formatação

```
Found 2 items
```

```
drwx----- - hadoop supergroup          0 2014-01-17 01:05 /tmp
drwxr-xr-x - hadoop supergroup          0 2014-01-17 01:05 /user
```

HACK51 – Criando um diretório Filesystem do Hadoop.

1. Vamos criar um diretório em /user/hadoop para armazenar nossos hacks de trabalho. crie um diretório chamado hacks.

```
hdfs dfs -mkdir /user/
hdfs dfs -mkdir /user/hadoop
hdfs dfs -mkdir /user/hadoop/hacks
```

2. Confira se o mesmo foi criado com ls

```
hdfs dfs -ls /user/hadoop/
```

HACK53 – Copiando um arquivo para Filesystem do Hadoop (HDFS).

1. Vamos criar um arquivo texto para copiar para o hadoop , acesse seu home e crie o mesmo.

```
cd ~
```

```
echo "O hadoop he um filesystem distribuido" >> hack_test.txt
```

2. Veja o conteúdo do arquivo com o comando cat

```
cat hack_test.txt  
"O hadoop he um filesystem distribuido"
```

3. Copie o arquivo para o diretório recém-criado (hacks)

```
hdfs dfs -copyFromLocal hack_test.txt /user/hadoop/hacks
```

4. Lista verificando se o arquivo está no hadoop

```
hdfs dfs -ls /user/hadoop/hacks  
Found 1 items  
-rw-r--r--    1 hadoop supergroup          44 2014-01-19 09:35  
/user/hadoop/hacks/hack_test.txt
```

HACK54 – Mostrando o conteúdo de arquivo no HDFS.

1. Para mostrar um conteúdo de arquivo use o comando cat do hadoop

```
hdfs dfs -cat /user/hadoop/hacks/hack_test.txt  
"O hadoop é um filesystem distribuido"
```

HACK55 – Copiando arquivos entre diretórios no HDFS.

1. Crie os diretórios **input** e **output** no hdfs /user/hadoop/hacks.

```
hdfs dfs -mkdir /user/hadoop/hacks/input  
hdfs dfs -mkdir /user/hadoop/hacks/output
```

2. Copie o arquivo hack_test.txt para o diretório input e para o output

```
hdfs dfs -cp /user/hadoop/hacks/hack_test.txt  
/user/hadoop/hacks/input/hack_test.txt  
  
hdfs dfs -cp /user/hadoop/hacks/hack_test.txt  
/user/hadoop/hacks/output/hack_test.txt
```

3. Usaremos posteriormente estes diretórios e arquivos para executar exemplos de mapreduce

HACK56 – Apagando um arquivo no HDFS.

1. Para apagar o arquivo `hack_test.txt` do diretório **output** use o comando `rm`

```
hdfs dfs -rm /user/hadoop/hacks/output/hack_test.txt
```

HACK57 – Apagando um diretório do HDFS.

1. Para apagar o diretório **output** use o comando `rmdir`

```
hdfs dfs -rmdir /user/hadoop/hacks/output
```

HACK58 – Interfaces Web do HDFS.

1. Cada serviço do do HDFS (NameNode e DataNode) disponibiliza uma interface web de acesso, para acessar use a porta 50070, mude o ip no link a seguir pelo ip da sua virtual machine.

```
http://192.168.0.102:50070
```

2. sera listada diversas informações sobre o hdfs

NameNode 'localhost:9000' (active)

Started:	Sat Jan 18 23:12:19 BRST 2014
Version:	2.2.0, 1529768
Compiled:	2013-10-07T06:28Z by hortonmu from branch-2.2.0
Cluster ID:	CID-cbe36208-7f3e-4790-9cff-186c2e30f908
Block Pool ID:	BP-1369347808-127.0.1.1-1389925992897

[Browse the filesystem](#)
[NameNode Logs](#)

Cluster Summary

Security is OFF

16 files and directories, 4 blocks = 20 total.

Heap Memory used 19.38 MB is 47% of Committed Heap Memory 40.77 MB.

Non Heap Memory used 32.48 MB is 96% of Committed Non Heap Memory 33.75 MB.

Configured Capacity	:	5.78 GB
DFS Used	:	168 KB
Non DFS Used	:	3.98 GB
DFS Remaining	:	1.80 GB
DFS Used%	:	0.00%

3. No link [Browse the filesystem](#) pode navegar pelo sistema de arquivos , **ATENÇÃO** porém quando clicar ele vai redirecionar para localhost altere para seu ip , conforme exemplo a seguir

```
http://192.168.0.102:50075/browseDirectory.jsp?namenodeInfoPort=50070&dir=/&nnaddr=127.0.0.1:9000
```

4. Nevegue entre seus diretórios criados.

HACK59 – Consultando a lista de comando do HDFS.

1. Na documentação é possível visualizar todos os comandos de manipulação do filesystem hadoop pelo link a seguir

<http://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-common/FileSystemShell.html>

Referências

Hadoop Real-World Solutions Cookbook

Author(s) : Jonathan R. Owens, Brian Femiano, Jon Lentz, PACKT Publishing

HDFS Users Guide

[http://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsUserGuide.html#Shell Commands](http://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsUserGuide.html#Shell_Commands)