

HDFS 实验

一. 实验目的

熟悉 HDFS 操作常用的 Shell 命令

参考资料: <https://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/FileSystemShell.html>

二. 实验内容

Hadoop 已预先安装好, 同学们如需要在本地练习安装使用 Hadoop, 可参考

https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-common/SingleCluster.html#Prepare_to_Start_the_Hadoop_Cluster

- 1) 启动 Hadoop, 开启 NameNode 和 DataNode 守护进程, 启动时可能会出现如下 WARN 提示: WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable 提示可以忽略, 并不会影响正常使用。

```
[guesta@server1 ~]$ start-dfs.sh
```

- 2) 使用 mkdir 命令在 HDFS 中创建目录, 以供之后存放文件, 例: (其中, /user/input 为要创建的目录, 注意当前用户是否有在该目录下写入的权限)

```
[guesta@server1 ~]$ hdfs dfs -mkdir -p /user/input
```

- 3) 使用 ls 命令查看 HDFS 中某一个目录, 输出该目录下的所有文件的读写权限、大小、创建时间、路径等信息, 例:

```
[guesta@server1 ~]$ hdfs dfs -ls /user
```

- 4) 使用 put 命令将本地文件上传至 HDFS, 例:

```
[guesta@server1 ~]$ hdfs dfs -put ./etc/hadoop/core-site.xml /user/input
```

若文件在 HDFS 中已经存在, 可覆盖 HDFS 已有文件, 例:

```
[guesta@server1 ~]$ hdfs dfs -put -f ./etc/hadoop/core-site.xml /user/input
```

若文件在 HDFS 中已经存在, 可在 HDFS 已有文件基础上追加内容, 例:

```
[guesta@server1 ~]$ hdfs dfs -put -appendToFile ./etc/hadoop/core-site.xml /user/input/core-site.xml
```

- 5) 使用 rm 命令删除 HDFS 中文件夹下文件, 例:

```
[guesta@server1 ~]$ hdfs dfs -rm /user/input/core-site.xml
```

- 6) 使用 `rm` 命令删除 HDFS 中的文件夹，例：

```
[guesta@server1 ~]$ hdfs dfs -rm -r /user/input
```

- 7) 运行 Hadoop 自带 word count 实例并查看结果，例：

先在本地创建测试文件 `hello.txt`，并上传至 HDFS



```
[guesta@server1 ~]$ hdfs -dfs put hello.txt /user/input
```

运行 Hadoop 自带 word count 程序

```
[guesta@server1 ~]$ hadoop jar /usr/local/hadoop/share/hadoop/mapreduce/  
hadoop-mapreduce-examples-2.7.7.jar wordcount /user/input/hello.txt  
/user/output
```

将结果文件输出至终端，查看 word count 结果，例：

```
[guesta@server1 ~]$ hdfs dfs -cat /user/output/part-r-00000
```

- 8) 关闭 Hadoop

```
[guesta@server1 ~]$ stop-dfs.sh
```