

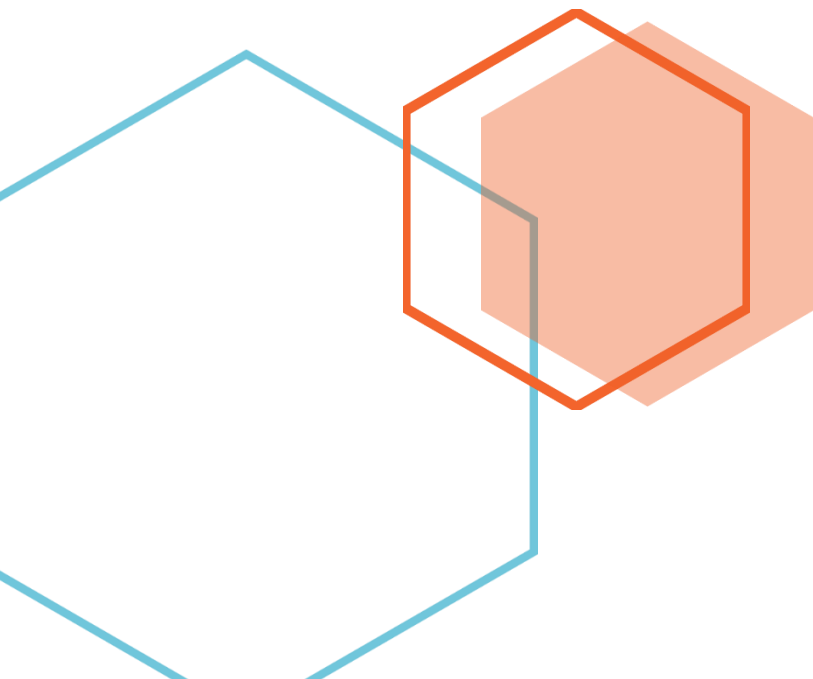


[Compte Rendu]

TP 1 : Manipulation du système de fichiers HDFS

Big Data: Fondements et Architectures de stockage

L'Ecole Normale Supérieure de l'Enseignement Technique de
Mohammedia (ENSET)



...

1. Démarrez les processus Hadoop avec les commandes start-dfs.sh et start-yarn.sh.

```
root@amina:/# ssh localhost
Welcome to Ubuntu 16.04 LTS (GNU/Linux 4.4.0-21-generic x86_64)

 * Documentation:  https://help.ubuntu.com/

775 packages can be updated.
518 updates are security updates.

New release '18.04.6 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Wed Feb 23 14:14:59 2022 from 127.0.0.1
```

```
root@amina:~# start-dfs.sh
22/02/23 16:29:16 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: starting namenode, logging to /root/hadoop-2.7.1/logs/hadoop-root-nam
enode-amina.out
localhost: starting datanode, logging to /root/hadoop-2.7.1/logs/hadoop-root-dat
anode-amina.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /root/hadoop-2.7.1/logs/hadoop-r
oot-secondarynamenode-amina.out
22/02/23 16:29:31 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
```

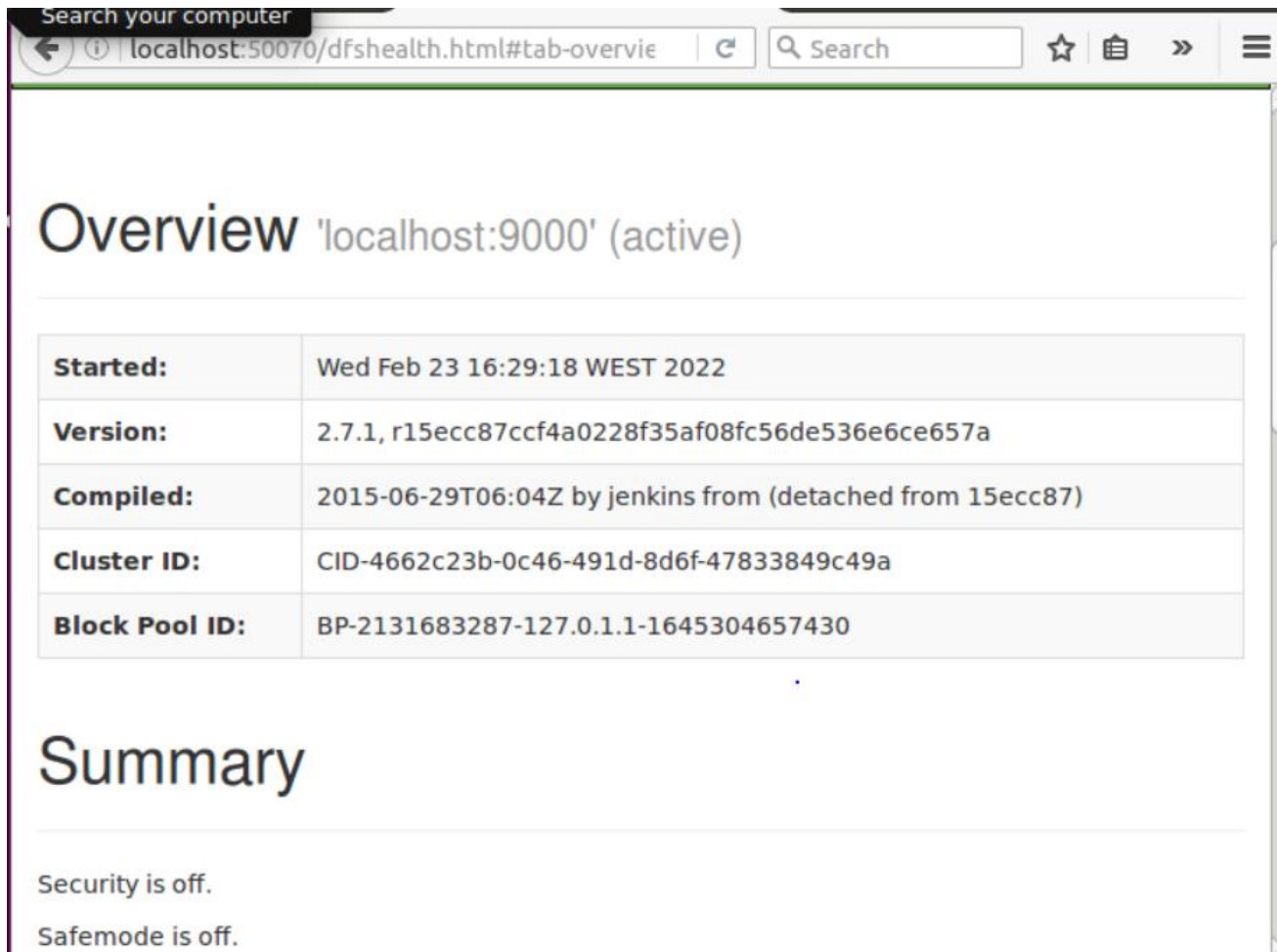
```
root@amina:~# start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /root/hadoop-2.7.1/logs/yarn-root-resourcem
anager-amina.out
localhost: starting nodemanager, logging to /root/hadoop-2.7.1/logs/yarn-root-no
demanager-amina.out
```

Ensuite, vérifiez qu'il est en cours d'exécution avec la commande jps

```
root@amina:~# jps
2211 NameNode
2548 SecondaryNameNode
2709 ResourceManager
2360 DataNode
3146 Jps
2955 NodeManager
```

...

Et accéder à l'interface web de NameNode avec <http://localhost:50070>.



The screenshot shows a web browser window with the URL `localhost:50070/dfshealth.html#tab-overview`. The page title is "Overview 'localhost:9000' (active)". Below the title is a table with the following information:

Started:	Wed Feb 23 16:29:18 WEST 2022
Version:	2.7.1, r15ecc87ccf4a0228f35af08fc56de536e6ce657a
Compiled:	2015-06-29T06:04Z by jenkins from (detached from 15ecc87)
Cluster ID:	CID-4662c23b-0c46-491d-8d6f-47833849c49a
Block Pool ID:	BP-2131683287-127.0.1.1-1645304657430

Below the table is a section titled "Summary" with the following text:

Security is off.
Safemode is off.

2. Créez dans la racine du HDFS l'arborescence suivante :

```
root@amina:~# hdfs dfs -mkdir -p /BDCC/JAVA/TPs
```

```
root@amina:~# hdfs dfs -mkdir -p /BDCC/JAVA/Cours
```

```
root@amina:~# hdfs dfs -mkdir -p /BDCC/CPP/TPs
```

```
root@amina:~# hdfs dfs -mkdir -p /BDCC/CPP/Cours
```

localhost:50070/explorer.html#/

Browse Directory

/ Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	232022/2/15:58:49	0	0 B	BDCC
drwxr-xr-x	root	supergroup	0 B	232022/2/14:12:13	0	0 B	rep
drwxr-xr-x	root	supergroup	0 B	232022/2/14:21:23	0	0 B	rep1
drwxr-xr-x	root	supergroup	0 B	232022/2/15:34:40	0	0 B	user

/BDCC Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	232022/2/15:58:05	0	0 B	CPP
drwxr-xr-x	root	supergroup	0 B	232022/2/15:59:06	0	0 B	JAVA

/BDCC/JAVA Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	232022/2/15:58:49	0	0 B	Cours
drwxr-xr-x	root	supergroup	0 B	232022/2/15:59:06	0	0 B	TPs

/BDCC/CPP							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	232022/2/15:56:24	0	0 B	Cours
drwxr-xr-x	root	supergroup	0 B	232022/2/15:58:05	0	0 B	TPs

3. Créez dans le répertoire Cours de CPP les fichiers CoursCPP1, CoursCPP2 et CoursCPP3. Puis ajoutez du contenu dans les fichiers créés.

```
hdfs dfs -touchz /BDCC/CPP/Cours/CoursCPP1
```

```
hdfs dfs -touchz /BDCC/CPP/Cours/CoursCPP2
```

```
root@amina:~# hdfs dfs -touchz /BDCC/CPP/Cours/CoursCPP3
```

Browse Directory

/BDCC/CPP/Cours							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	232022/2/16:48:45	1	128 MB	CoursCPP1
-rw-r--r--	root	supergroup	0 B	232022/2/16:49:19	1	128 MB	CoursCPP2
-rw-r--r--	root	supergroup	0 B	232022/2/16:49:24	1	128 MB	CoursCPP3

- Affichez le contenu des fichiers CoursCPP1, CoursCPP2 et CoursCPP3.

```
root@amina:~# hdfs dfs -appendToFile - /BDCC/JAVA/Cours/CoursCPP1
22/02/23 20:15:53 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Amina Dahmouni
^Croot@amina:~# hdfs dfs -cat /BDCC/JAVA/Cours/CoursCPP1
22/02/23 20:16:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Amina Dahmouni
```

```
root@amina:~# hdfs dfs -appendToFile - /BDCC/JAVA/Cours/CoursCPP2
22/02/23 20:16:56 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Big data and cloud computing
^Croot@amina:~# hdfs dfs cat /BDCC/JAVA/Cours/CoursCPP2
cat: Unknown command
Did you mean -cat? This command begins with a dash.
root@amina:~# hdfs dfs -cat /BDCC/JAVA/Cours/CoursCPP2
22/02/23 20:18:02 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Big data and cloud computing
```

- Copiez les fichiers CoursCPP1, CoursCPP2 et CoursCPP3 dans le répertoire Cours de l'JAVA.

```
root@amina:~# hdfs dfs -mv /BDCC/CPP/Cours/CoursCPP1 /BDCC/JAVA/Cours
```

```
root@amina:~# hdfs dfs -mv /BDCC/CPP/Cours/CoursCPP2 /BDCC/JAVA/Cours
```

```
root@amina:~# hdfs dfs -mv /BDCC/CPP/Cours/CoursCPP3 /BDCC/JAVA/Cours
```

/BDCC/JAVA/Cours							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	23/02/2022 16:48:45	1	128 MB	CoursCPP1
-rw-r--r--	root	supergroup	0 B	23/02/2022 16:49:19	1	128 MB	CoursCPP2
-rw-r--r--	root	supergroup	0 B	23/02/2022 16:49:24	1	128 MB	CoursCPP3

- Supprimer le fichier CoursCPP3 dans le répertoire Cours de l'JAVA, et renommer CoursCPP1 et CoursCPP2 par CoursJAVA1 et CoursJAVA2 respectivement.

```

root@amina:~# hdfs dfs -rm /BDCC/JAVA/Cours/CoursCPP3
22/02/23 20:22:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
22/02/23 20:22:55 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /BDCC/JAVA/Cours/CoursCPP3
root@amina:~#

```

/BDCC/JAVA/Cours								Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
-rw-r--r--	root	supergroup	15 B	23/02/2022 19:16:17	1	128 MB	CoursCPP1	
-rw-r--r--	root	supergroup	29 B	23/02/2022 19:17:27	1	128 MB	CoursCPP2	

```






root@amina:~# hdfs dfs -mv /BDCC/JAVA/Cours/CoursCPP1 /BDCC/JAVA/Cours/CoursJAVA1
22/02/23 20:32:12 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
root@amina:~# hdfs dfs -mv /BDCC/JAVA/Cours/CoursCPP2 /BDCC/JAVA/Cours/CoursJAVA2
22/02/23 20:32:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
root@amina:~#

```

/BDCC/JAVA/Cours								Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
-rw-r--r--	root	supergroup	15 B	23/02/2022 19:16:17	1	128 MB	CoursJAVA1	
-rw-r--r--	root	supergroup	29 B	23/02/2022 19:17:27	1	128 MB	CoursJAVA2	

- Créer un répertoire dans le système de fichier local les fichiers TP1CPP, TP2CPP, TP1JAVA, TP2JAVA, TP3JAVA.

```
root@amina:/home/RepHadoopTP1# touch TP1CPP TP2CPP TP1JAVA TP2JAVA TP3JAVA
root@amina:/home/RepHadoopTP1# ls
TP1CPP TP1JAVA TP2CPP TP2JAVA TP3JAVA
root@amina:/home/RepHadoopTP1#
```

Name	Size	Type	Modified
 TP1CPP	0 bytes	Text	20:44
 TP1JAVA	0 bytes	Text	20:44
 TP2CPP	0 bytes	Text	20:44
 TP2JAVA	0 bytes	Text	20:44
 TP3JAVA	0 bytes	Text	20:44

- Copier les fichiers TP1CPP, TP2CPP à partir du système de fichier local vers le répertoire TPs de répertoire CPP.

```
root@amina:/home/RepHadoopTP1# hdfs dfs -put /home/RepHadoopTP1/TP1CPP /BDCC/CPP
/TPs
22/02/23 20:58:27 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
root@amina:/home/RepHadoopTP1# hdfs dfs -put /home/RepHadoopTP1/TP2CPP /BDCC/CPP
/TPs
22/02/23 20:59:21 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
root@amina:/home/RepHadoopTP1#
```

- Copier les fichiers TP1CJAVA, TP2JAVA à partir du système de fichier local vers le répertoire TPs de répertoire JAVA.

```
root@amina:/home/RepHadoopTP1# hdfs dfs -put /home/RepHadoopTP1/TP1JAVA /BDCC/JA
VA/TPs
22/02/23 21:16:38 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
root@amina:/home/RepHadoopTP1# hdfs dfs -put /home/RepHadoopTP1/TP2JAVA /BDCC/JA
VA/TPs
22/02/23 21:17:22 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
root@amina:/home/RepHadoopTP1#
```


...

/BDCC/PHP/TPs							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	23/02/2022 20:49:25	1	128 MB	TP1PHP
-rw-r--r--	root	supergroup	0 B	23/02/2022 20:49:44	1	128 MB	TP2PHP

/BDCC/JAVA/TPs							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	23/02/2022 20:51:56	1	128 MB	TP1JAVA
-rw-r--r--	root	supergroup	0 B	23/02/2022 20:52:13	1	128 MB	TP2JAVA

10. Afficher tout le contenu du répertoire BDCC d'une manière récursive.

```
root@amina:~# hdfs dfs -ls -R
22/02/23 21:34:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
drwxr-xr-x - root supergroup 0 2022-02-23 16:36 BDCC
drwxr-xr-x - root supergroup 0 2022-02-23 16:39 BDCC/PHP
drwxr-xr-x - root supergroup 0 2022-02-23 16:39 BDCC/PHP/Cours
drwxr-xr-x - root supergroup 0 2022-02-23 16:36 BDCC/PHP/TPs
drwxr-xr-x - root supergroup 0 2022-02-23 16:36 BDCC/JAVA
drwxr-xr-x - root supergroup 0 2022-02-23 16:36 BDCC/JAVA/Cours
drwxr-xr-x - root supergroup 0 2022-02-23 16:34 BDCC/JAVA/TPs
-rw-r--r-- 1 root supergroup 0 2022-02-23 17:48 CoursPHP2
-rw-r--r-- 1 root supergroup 0 2022-02-23 17:48 CoursPHP3
root@amina:~#
```

11. Supprimer le fichier TP1PHP du répertoire TPs.

```
root@amina:/home/RepHadoopTP1# hdfs dfs -rm /BDCC/PHP/TPs/TP1PHP
22/02/23 21:57:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
22/02/23 21:57:44 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Empty interval = 0 minutes.
Deleted /BDCC/PHP/TPs/TP1PHP
root@amina:/home/RepHadoopTP1#
```

12. Supprimer le répertoire Java avec son contenu.

```
root@amina:/home/RepHadoopTP1# hdfs dfs -rmr /BDCC/JAVA/
rmr: DEPRECATED: Please use 'rm -r' instead.
22/02/23 22:00:50 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
22/02/23 22:00:51 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /BDCC/JAVA
root@amina:/home/RepHadoopTP1#
```

/BDCC							Go!
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	23/02/2022/15:58:05	0	0 B	CPP