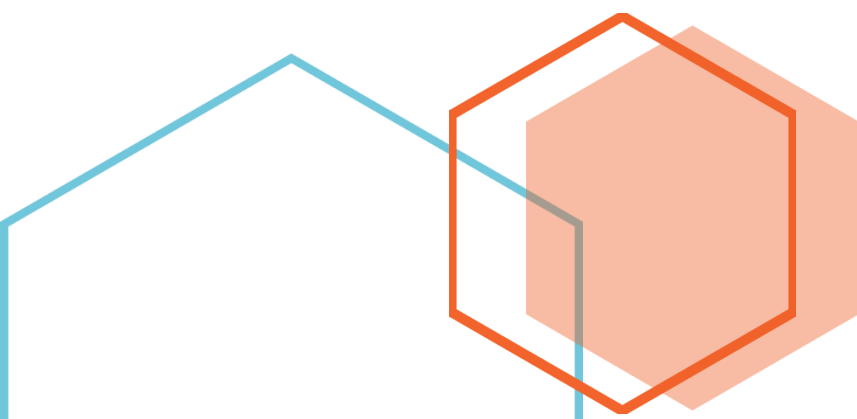# [Compte Rendu]

TP 2: API JAVA FileSystem

Big Data: Fondements et Architectures de stockage

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

1&2.

```
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
root@amina:~# hdfs dfs -ls -R
drwxr-xr-x   - root supergroup          0 2022-02-23 16:36 BDCC
drwxr-xr-x   - root supergroup          0 2022-02-23 16:39 BDCC/CPP
drwxr-xr-x   - root supergroup          0 2022-02-23 16:39 BDCC/CPP/Cours
drwxr-xr-x   - root supergroup          0 2022-02-23 16:36 BDCC/CPP/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 CoursCPP2
-rw-r--r--   1 root supergroup          0 2022-02-23 17:48 CoursCPP3
root@amina:~#
```

3.

```java
import ...
public class Q3 {
    public static void main(String[] args) {
        String[] pathValue = {"/BDCC/CPP/Cours/CoursCPP1.txt","/BDCC/CPP/Cours/CoursCPP2.txt","/BDCC/CPP/Cours/CoursCPP3.txt"};
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            for(int i=0;i<3;i++)
            {
                String url=pathValue[i];
                Path path=new Path(url);
                FSDataOutputStream fsdo= fs.create(path);
                BufferedWriter br=new BufferedWriter(new OutputStreamWriter(fsdo));
                br.write( "Bonjour ");
                br.newLine();
                br.write( "tout le monde");
                br.close();
            }
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

localhost:50070/explorer.html#/BDCC/CPP/Cours

Hadoop   Overview   Datanodes   Snapshot   Startup Progress   Utilities ▾

Browse Directory

/BDCC/CPP/Cours   Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	22 B	3/2/2022, 6:53:50 AM	1	128 MB	CoursCPP1.txt
-rw-r--r--	root	supergroup	22 B	3/2/2022, 6:53:50 AM	1	128 MB	CoursCPP2.txt
-rw-r--r--	root	supergroup	22 B	3/2/2022, 6:53:50 AM	1	128 MB	CoursCPP3.txt

Hadoop, 2015.

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

```java
import ...
public class Q4 {
    public static void main(String[] args) {
        String[] pathValue = {"/BDCC/CPP/Cours/CoursCPP1.txt","/BDCC/CPP/Cours/CoursCPP2.txt","/BDCC/CPP/Cours/CoursCPP3.txt"};
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            for(int i=0;i<3;i++)
            {
                String url=pathValue[i];
                Path path=new Path(url);
                FSDataInputStream fsi=fs.open(path);
                BufferedReader br=new BufferedReader(new InputStreamReader(fsi));
                String ligne=null;
                while((ligne=br.readLine())!=null){
                    System.out.println(ligne);
                }
                br.close();
            }
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

```
Bonjour
tout le monde
Bonjour
tout le monde
Bonjour
tout le monde

Process finished with exit code 0
```

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

```java
java ×    G Q4.java ×    G Q5.java ×
import ...
public class Q5 {
    public static void main(String[] args) {
        String[] pathValue = {"/BDCC/CPP/Cours/CoursCPP1.txt","/BDCC/CPP/Cours/CoursCPP2.txt","/BDCC/CPP/Cours/CoursCPP3.txt"};
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            for(int i=0;i<3;i++)
            {
                String url=pathValue[i];
                FileUtil.copy(fs,new Path(url),fs,new Path( pathString: "/BDCC/JAVA/Cours"), deleteSource: false,cf);
            }
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Hadoop   Overview   Datanodes   Snapshot   Startup Progress   Utilities ▾

Browse Directory

/BDCC/JAVA/Cours   Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursCPP1.txt
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursCPP2.txt
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursCPP3.txt

Hadoop, 2015.

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

```java
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
public class Q6 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            fs.delete(new Path( pathString: "/BDCC/JAVA/Cours/CoursCPP3.txt"), b: false);
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Hadoop   Overview   Datanodes   Snapshot   Startup Progress   Utilities ▾

Browse Directory

/BDCC/JAVA/Cours								Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursCPP1.txt
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursCPP2.txt

Hadoop, 2015.

```java
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
public class Q6 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            // fs.delete(new Path("/BDCC/JAVA/Cours/CoursCPP3.txt"),false);
            fs.rename(new Path( pathString: "/BDCC/JAVA/Cours/CoursCPP1.txt"),new Path( pathString: "/BDCC/JAVA/Cours/CoursJAVA1.txt"));
            fs.rename(new Path( pathString: "/BDCC/JAVA/Cours/CoursCPP2.txt"),new Path( pathString: "/BDCC/JAVA/Cours/CoursJAVA2.txt"));
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Browse Directory

/BDCC/JAVA/Cours								Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursJAVA1.txt
-rw-r--r--	root	supergroup	22 B	3/2/2022, 7:23:54 AM	1	128 MB	CoursJAVA2.txt

Hadoop, 2015.

7. Créer un répertoire dans le système de fichier local les fichiers :

TP1CPP,TP2CPP,TP1JAVA,TP2JAVA,TP3JAVA.

Home	Documents	BDCC4	BigData	Rep_Hadoop_TP1 ⌄							
Name							Size	Modified	Star		
TP1CPP.txt							0 bytes	16:28	☆		
TP1JAVA.txt							0 bytes	16:29	☆		
TP2CPP.txt							0 bytes	16:29	☆		
TP2JAVA.txt							0 bytes	16:29	☆		
TP3JAVA.txt							0 bytes	16:29	☆		

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

```java
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
public class Q8 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            fs.copyFromLocalFile(new Path( pathString: "/home/amina/Documents/BDCC4/BigData/Rep_Hadoop_TP1/TP1CPP.txt"),
                new Path( pathString: "/BDCC/CPP/TPs"));
            fs.copyFromLocalFile(new Path( pathString: "/home/amina/Documents/BDCC4/BigData/Rep_Hadoop_TP1/TP2CPP.txt"),
                new Path( pathString: "/BDCC/CPP/TPs"));
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Browse Directory

/BDCC/CPP/TPs   Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	3/2/2022, 4:39:51 PM	1	128 MB	TP1CPP.txt
-rw-r--r--	root	supergroup	0 B	3/2/2022, 4:39:52 PM	1	128 MB	TP2CPP.txt

Hadoop, 2015.

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

```java
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
public class Q9 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            fs.copyFromLocalFile(new Path( pathString: "/home/amina/Documents/BDCC4/BigData/Rep_Hadoop_TP1/TP1JAVA.txt"),
                new Path( pathString: "/BDCC/JAVA/TPs"));
            fs.copyFromLocalFile(new Path( pathString: "/home/amina/Documents/BDCC4/BigData/Rep_Hadoop_TP1/TP2JAVA.txt"),
                new Path( pathString: "/BDCC/JAVA/TPs"));
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Browse Directory

/BDCC/JAVA/TPs   Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	3/2/2022, 4:45:40 PM	1	128 MB	TP1JAVA.txt
-rw-r--r--	root	supergroup	0 B	3/2/2022, 4:45:40 PM	1	128 MB	TP2JAVA.txt

Hadoop, 2015.

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

```java
import ...
public class Q10 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            List<String> tabFiles = getAllFilePath(new Path( pathString: "/BDCC"),fs);
            fs.close();
            for(int i=0;i<tabFiles.size();i++){
                System.out.println(tabFiles.get(i));
            }
        }catch (Exception e){
            e.printStackTrace();
        }
    }
    public static List<String> getAllFilePath(Path filePath, FileSystem fs) throws FileNotFoundException, IOException {
        List<String> fileList = new ArrayList<String>();
        FileStatus[] fileStatus = fs.listStatus(filePath);
        for (FileStatus fileStat : fileStatus) {
            if (fileStat.isDirectory()) {
                fileList.addAll(getAllFilePath(fileStat.getPath(), fs));
            } else {fileList.add(fileStat.getPath().toString());}
        }
        return fileList;
    }
}
```

```
hdfs://localhost:9000/BDCC/CPP/Cours/CoursCPP1.txt
hdfs://localhost:9000/BDCC/CPP/Cours/CoursCPP2.txt
hdfs://localhost:9000/BDCC/CPP/TPs/TP1CPP.txt
hdfs://localhost:9000/BDCC/CPP/TPs/TP2CPP.txt
hdfs://localhost:9000/BDCC/JAVA/Cours/CoursJAVA1.txt
hdfs://localhost:9000/BDCC/JAVA/Cours/CoursJAVA2.txt
hdfs://localhost:9000/BDCC/JAVA/TPs/TP1JAVA.txt
hdfs://localhost:9000/BDCC/JAVA/TPs/TP2JAVA.txt

Process finished with exit code 0
```

11. Supprimer le fichier TP1CPP du répertoire TPs.

```java
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
public class Q11 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        //cf.set("dfs.replication","1");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            fs.delete(new Path( pathString: "/BDCC/CPP/TPs/TP1CPP.txt"), b: false);
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Hadoop	Overview	Datanodes	Snapshot	Startup Progress	Utilities ⌄

Browse Directory

/BDCC/CPP/TPs								Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	root	supergroup	0 B	3/2/2022, 11:14:38 PM	1	128 MB	TP2CPP.txt

Hadoop, 2015.

12. Supprimer le répertoire Java avec son contenu.

```java
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.FileSystem;
import org.apache.hadoop.fs.Path;
public class Q12 {
    public static void main(String[] args) {
        Configuration cf=new Configuration();
        cf.set("fs.defaultFS","hdfs://localhost:9000");
        System.setProperty("HADOOP_USER_NAME","root");
        try {
            FileSystem fs=FileSystem.get(cf);
            fs.delete(new Path( pathString: "/BDCC/JAVA"), b: true);
            fs.close();
        }catch (Exception e){
            e.printStackTrace();
        }
    }
}
```

Hadoop   Overview   Datanodes   Snapshot   Startup Progress   Utilities ▾

Browse Directory

| /BDCC | | | | | | | | Go! |

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	root	supergroup	0 B	3/2/2022, 6:53:49 AM	0	0 B	CPP

Hadoop, 2015.