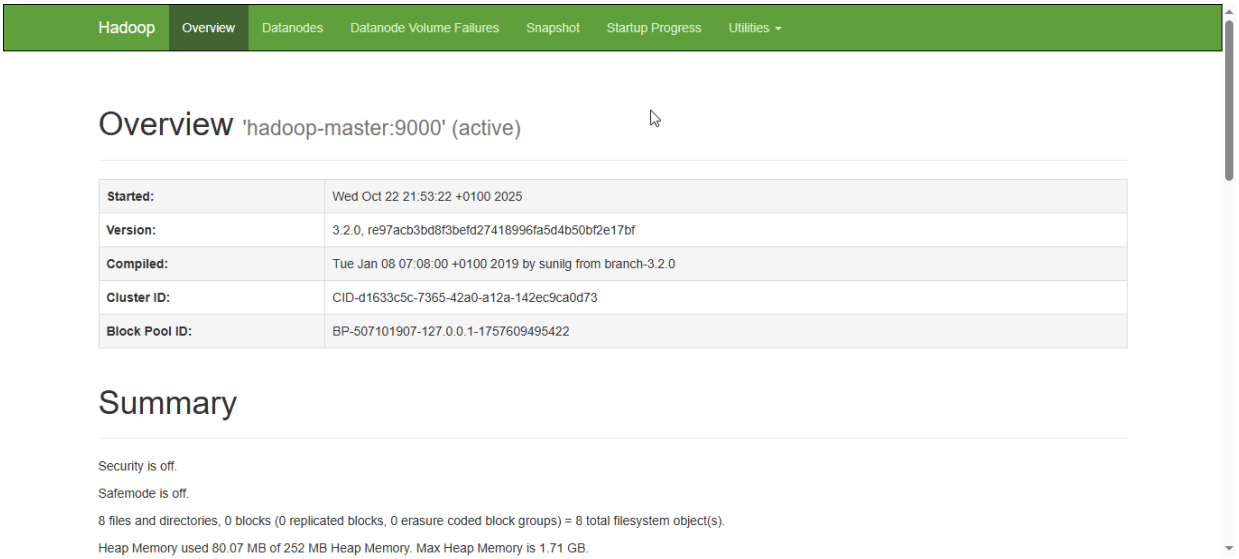
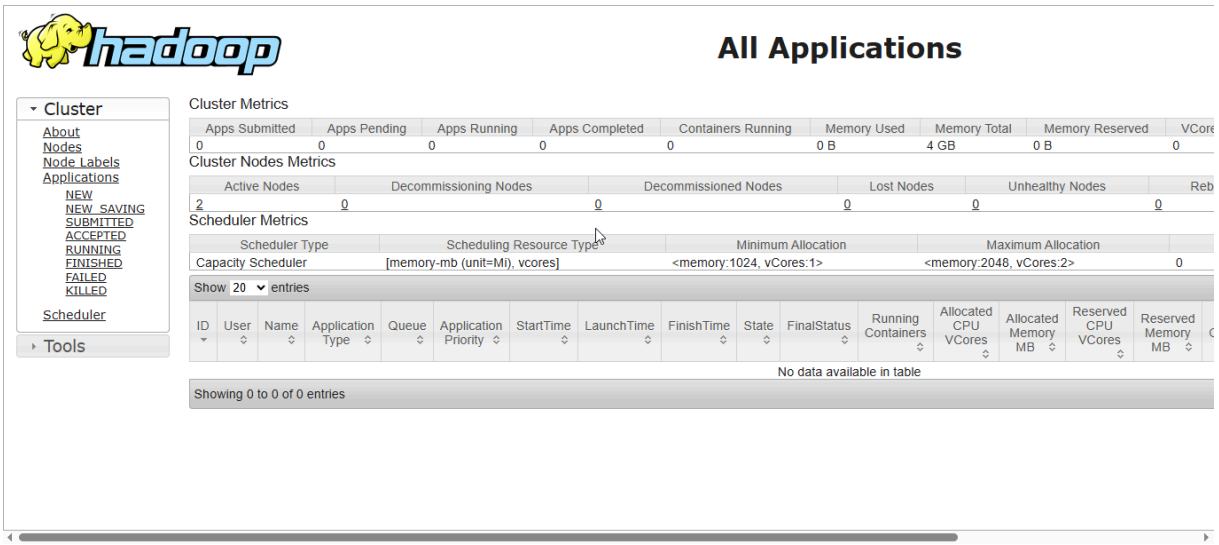
Compte rendu BIG DATA

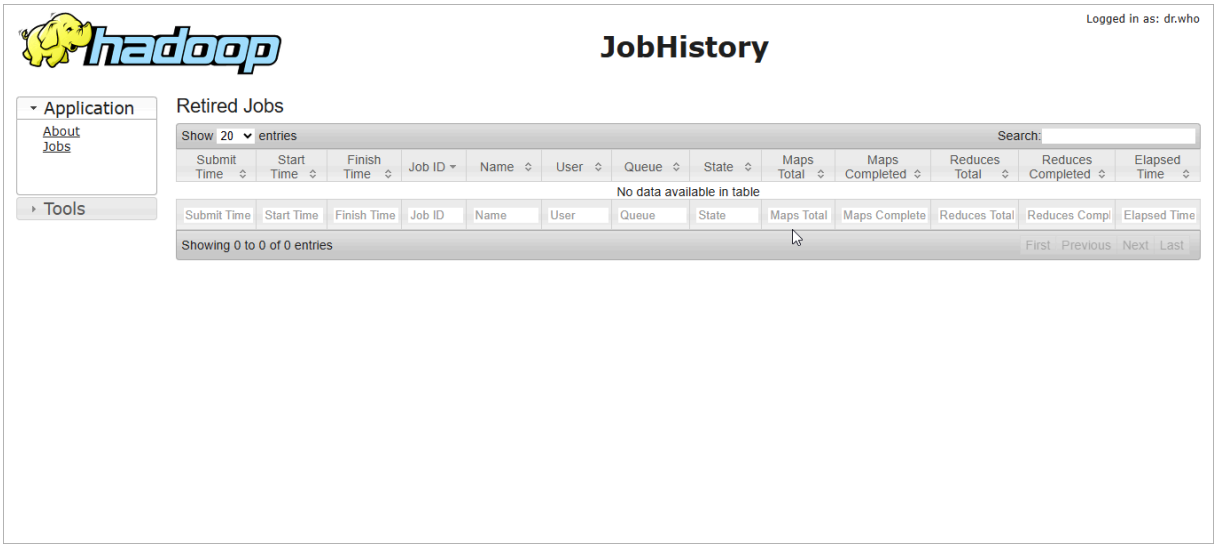
I. Installation et configuration :

1. Vérification d'installation de Hadoop et yarn.

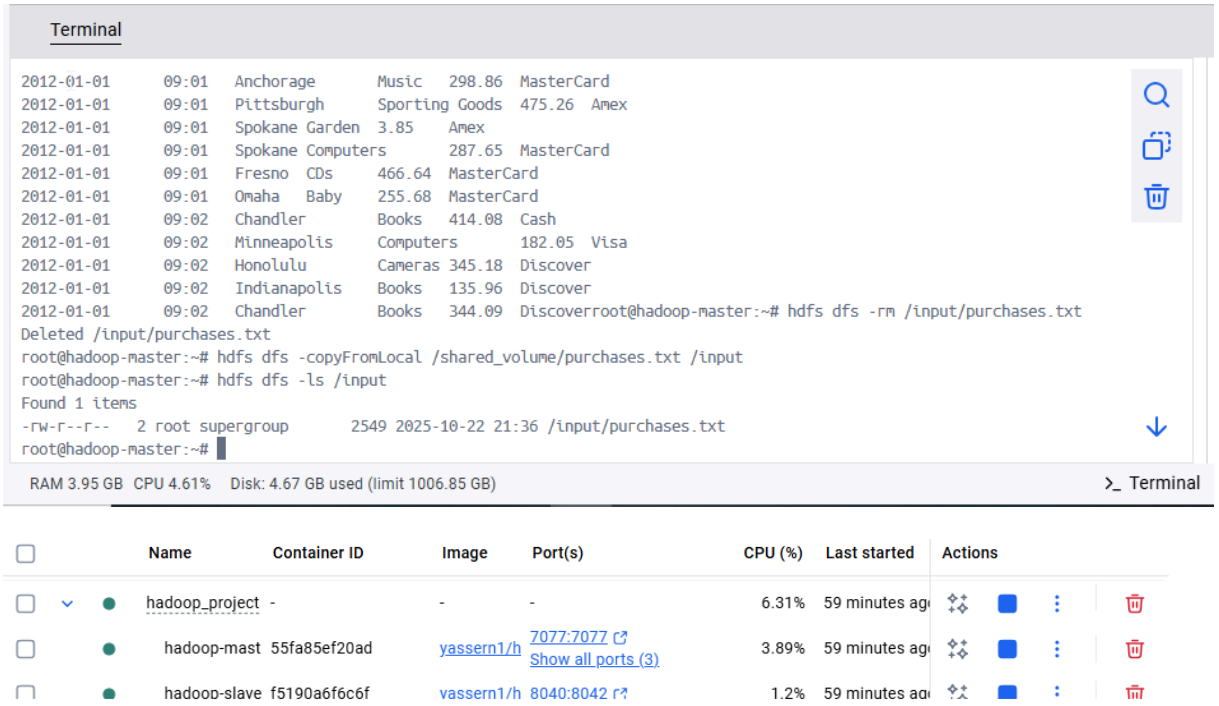
Hadoop ainsi que YARN démarrent et s'exécutent correctement.



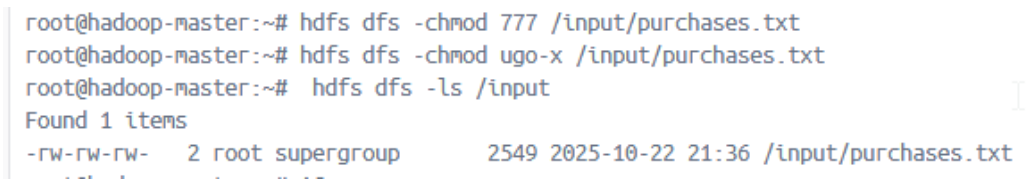




2. Manipulations HDFS :



3. Gérer les droits :



4. Déplacer et copier le fichier purchases.txt dans un autre emplacement HDFS (ex. racine) :

```
root@hadoop-master:~# hdfs dfs -ls -R /
drwxr-xr-x   - root supergroup          0 2025-10-23 01:00 /input
-rw-rw-rw-   2 root supergroup       2549 2025-10-22 21:36 /input/purchases.txt
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging/history
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging/history/done
drwxrwxrwt   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging/history/done_intermediate
```

```
root@hadoop-master:~# hdfs dfs -cp /input/purchases.txt /purchases.txt
root@hadoop-master:~# hdfs dfs -ls -R /
drwxr-xr-x   - root supergroup          0 2025-10-23 01:00 /input
-rw-rw-rw-   2 root supergroup       2549 2025-10-22 21:36 /input/purchases.txt
-rw-r--r--   2 root supergroup       2549 2025-10-23 01:02 /purchases.txt
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging/history
drwxrwx---   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging/history/done
drwxrwxrwt   - root supergroup          0 2025-10-02 09:06 /tmp/hadoop-yarn/staging/history/done_intermediate
```

```
root@hadoop-master:~# hdfs dfs -get /purchases.txt /shared_volume/achat.txt
root@hadoop-master:~# hdfs dfs -cp /purchases.txt /input/purchases.txt
cp: `/input/purchases.txt': File exists
```

Dans cette dernière image, nous avons essayé de créer une copie du fichier **purchases.txt** sous un autre nom (**achat.txt**) dans la machine locale. Ensuite, nous avons utilisé une commande pour copier le fichier dans la racine de HDFS, mais comme ce fichier existe déjà dans la racine (selon les étapes précédentes), le système a renvoyé le message **“File exists”**. Ainsi, le fait de copier le fichier dans le dossier **/shared_volume** avec un nom différent **ne change pas le nom du fichier dans HDFS**.

5. télécharger un fichier sur hdfs :

```
root@hadoop-master:~# hdfs dfs -mkdir /web_input
root@hadoop-master:~# hdfs dfs -ls /
Found 1 items
drwxr-xr-x   - root supergroup          0 2025-11-15 15:11 /web_input
root@hadoop-master:~# wget http://www.textfiles.com/etext/FICTION/alice.txt
--2025-11-15 15:16:46--  http://www.textfiles.com/etext/FICTION/alice.txt
Resolving www.textfiles.com (www.textfiles.com)... 208.86.224.90
Connecting to www.textfiles.com (www.textfiles.com)|208.86.224.90|:80... connected.
HTTP request sent, awaiting response... 200 OK
```

```
root@hadoop-master:~# hdfs dfs -put /shared_volume/alice.txt /web_input
root@hadoop-master:~# hdfs dfs -ls /web_input
Found 1 items
-rw-r--r--   2 root supergroup     150886 2025-11-15 15:36 /web_input/alice.txt
```