Fault Recovery in HDFS

Group 1

Mayank Singh Chauhan 2016CS50394
Atishya Jain 2016CS50393
Avaljot Singh 2016CS50389
Mankaran Singh 2016CS50391

August 29, 2019

1 Introduction

In this assignment, first we installed *Hadoop* file system on 4 VMs on Baadal, uploaded a few large files and analyzed how the data blocks get replicated on various data nodes.

2 Steps to install HDFS

1. Install JAVA:
   - sudo apt-get install default-jdk
2. Add a user to install and run HDFS. It also needs to be added to the sudo group.
   - sudo addgroup hadoop
   - sudo adduser –ingroup hadoop hadoopuser
   - sudo usermod -aG sudo hadoopuser
3. Generate an SSH key for the new user so as to allow the new user to access the local machine via SSH, with this key.
   - su - hadoopuser
   - ssh-keygen -t rsa
   - cat /.ssh/id_rsa.pub >> /.ssh/authorized_keys
4. Add the IP addresses corresponding to the master and the slaves of the Hadoop file system into /etc/hosts file.

- sudo vim /etc/hosts
- Add the following lines :
  (a) 10.17.50.25 mayank
  (b) 10.17.50.156 atishya
  (c) 10.17.3.64 mankaran
  (d) 10.17.6.25 avaljot

5. Now master has to be given access to all the slave VMs. The barrier here was the passwords that we had to enter during ssh. To circumvent this, we setup the ssh keys of the slaves on the master.

   - ssh-copy-id -i /.ssh/id_rsa.pub hadoopuser@mayank
   - ssh-copy-id -i /.ssh/id_rsa.pub hadoopuser@mankaran
   - ssh-copy-id -i /.ssh/id_rsa.pub hadoopuser@avaljot
   - ssh-copy-id -i /.ssh/id_rsa.pub hadoopuser@atishya

6. Now start the actual downloading of hadoop. We'll download it on the master node ans scp on the other slaves.

   - sudo mkdir /opt/hadoop
   - cd /opt/hadoop
   - wget https://archive.apache.org/dist/hadoop/core/hadoop-1.2.0/hadoop-1.2.0.tar.gz
   - tar -xzf hadoop-1.2.0.tar.gz
   - sudo mv hadoop-1.2.0 hadoop
   - scp -r /opt/hadoop/hadoop hadoopuser@atishya:/opt/hadoop
   - scp -r /opt/hadoop/hadoop hadoopuser@mankaran:/opt/hadoop
   - scp -r /opt/hadoop/hadoop hadoopuser@avaljot:/opt/hadoop

7. Change ownership of the hadoop folder on all the VMs to hadoopuser.

   - sudo chown -R hadoopuser /opt/hadoop

8. Now we configured the various configuration files resident inside /opt/hadoop/conf folder on all VMs as follows:

   - Add the following to the core-site.xml file:

```
<configuration>
    <property>
        <name>fs.default.name</name>
        <value>hdfs://mayank:9000/</value>
    </property>
    <property>
```

```
        <name>dfs.permissions</name>
        <value>false</value>
    </property>
</configuration>
```

- Add the following to the hdfs-site.xml file:

```
<configuration>
   <property>
      <name>dfs.data.dir</name>
      <value>/opt/hadoop/hadoop/dfs/name/data</value>
      <final>true</final>
   </property>
   <property>
      <name>dfs.name.dir</name>
      <value>/opt/hadoop/hadoop/dfs/name</value>
      <final>true</final>
   </property>
   <property>
      <name>dfs.replication</name>
      <value>3</value>
   </property>
</configuration>
```

- Add the following to the mapred-site.xml file:

```
<configuration>
   <property>
      <name>mapred.job.tracker</name>
      <value>mayank:9001</value>
   </property>
</configuration>
```

- Add the following to the hadoop-env.sh file:

```
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true
export HADOOP_CONF_DIR=/opt/hadoop/hadoop/conf
```

9. Now do the final configuration on the master VM (mayank).

    - Add this line to the /opt/hadoop/hadoop/conf/masters:
      mayank

- Add these lines to the file /opt/hadoop/hadoop/conf/slaves:
  atishya
  mankaran
  avaljot
- Format master node and start all nodes
  /opt/hadoop/hadoop/bin/hadoop namenode format
  /opt/hadoop/hadoop/bin/start-all.sh

3 Block Report

3.1 Initialisation

We added 4 files File1, File2, File3 and File4 each of approximately 231 MB. They were broken down into blocks of 64 MB. We noticed that each of them got replicated on 4 different VMs. Steps to add the files are:

1. Copying the file from hadoopuser to hadoop file system.
   bin/hadoop fs -copyFromLocal /File1/

2. Copying the files just after starting may throw out an error saying that safemode is on. Just wait for safemode to switch off after which you can proceed adding the files. See the figure for example.

error.png error.png

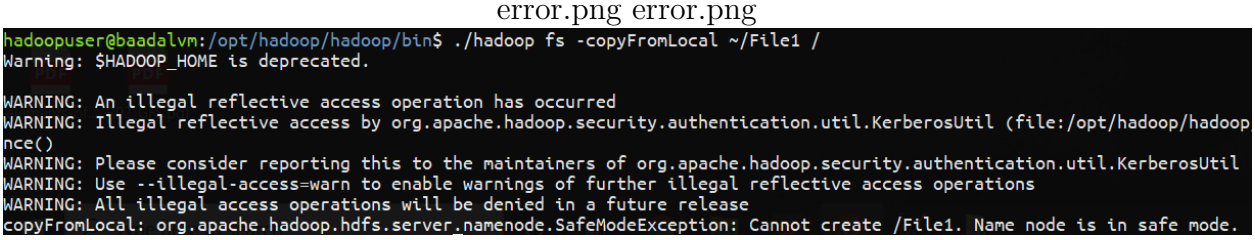

Figure 1: Safemode Error

3. See the description of each file with the distribution of their blocks over different VMs. We kept the replication factor of 3 for this assignment.

Block ID	Block distribution File 1 before shutdown		
4664252444576818739	10.17.6.25:50010	10.17.3.64:50010	10.17.50.25:50010
-140921271508634830	10.17.6.25:50010	10.17.50.156:50010	10.17.50.25:50010
-954538023745594084	10.17.3.64:50010	10.17.50.156:50010	10.17.50.25:50010
3311543059387042653	10.17.6.25:50010	10.17.50.156:50010	10.17.50.25:50010

Block ID	Block distribution File 2 before shutdown		
-9108694018629044460	10.17.6.25:50010	10.17.3.64:50010	10.17.50.25:50010
2617830728431356278	10.17.6.25:50010	10.17.50.156:50010	10.17.50.25:50010
3778736967815178453	10.17.3.64:50010	10.17.50.156:50010	10.17.50.25:50010
4436040171665550495	10.17.6.25:50010	10.17.50.156:50010	10.17.50.25:50010

Block ID	Block distribution File 3 before shutdown		
604131027942071764	10.17.6.25:50010	10.17.50.25:50010	10.17.3.64:50010
6844246988529318064	10.17.6.25:50010	10.17.50.156:50010	10.17.50.25:50010
-8508448695289813430	10.17.50.156:50010	10.17.50.25:50010	10.17.3.64:50010
8965565501969458440	10.17.6.25:50010	10.17.50.156:50010	10.17.50.25:50010

Block ID	Block distribution File 4 before shutdown		
5117826362547550415	10.17.6.25:50010	10.17.50.25:50010	10.17.50.156:50010
-7012850736704884080	10.17.6.25:50010	10.17.50.25:50010	10.17.3.64:50010
3227726318878224303	10.17.3.64:50010	10.17.50.25:50010	10.17.50.156:50010
-196786929323127591	10.17.6.25:50010	10.17.50.25:50010	10.17.50.156:50010

3.2 After ShutDown

We shutdown a datanode that is not master. We have shutdown the one with IP 10.17.6.25 (avaljot) to see the changes in the files.

1. We noticed that the files were replicated back to have a replication factor of 3. The process of replication took about 650 seconds to replicate the files.

2. Following was the cluster state after we shutdown 1 datanode.

Live Datanodes : 3

Node	Last Contact	Admin State	Configured Capacity (GB)	Used (GB)	Non DFS Used (GB)	Remaining (GB)	Used (%)	Used (%)	Remaining (%)	Blocks
baadalvm	1	In Service	70.37	0.91	8.14	61.32	1.29		87.14	18
baadalvm	0	In Service	70.37	0.55	7.06	62.75	0.79		89.17	11
baadalvm	0	In Service	70.37	0.64	7.08	62.64	0.91		89.03	12

Figure 2: Health Status after Shutdown

3. Following are the tables corresponding to the state of File blocks after replication.

Block ID	Block distribution File 1 after shutdown		
4664252444576818739	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010
-140921271508634830	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
-954538023745594084	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
3311543059387042653	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010

Block ID	Block distribution File 2 after shutdown		
-9108694018629044460	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010
2617830728431356278	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
3778736967815178453	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
4436040171665550495	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010

Block ID	Block distribution File 3 after shutdown		
604131027942071764	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
6844246988529318064	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
-8508448695289813430	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010
8965565501969458440	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010

Block ID	Block distribution File 4 after shutdown		
-7012850736704884080	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010
5117826362547550415	10.17.50.25:50010	10.17.50.156:50010	10.17.3.64:50010
3227726318878224303	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010
-196786929323127591	10.17.50.25:50010	10.17.3.64:50010	10.17.50.156:50010

4. Notice that the files are replicated back to 3 copies.

3.3 After Restoration

We restored the datanode with IP 10.17.6.25 (avaljot) to see the changes again in the files.

1. We noticed that the files were first replicated back to a factor of more than 3 but soon in sometime they reduced to the factor of 3. Probably HDFS first tried to load balance and somehow got the replicatio factor increased in between.

2. Following was the cluster state after we restore the datanode.

Live Datanodes : 4

Node	Last Contact	Admin State	Configured Capacity (GB)	Used (GB)	Non DFS Used (GB)	Remaining (GB)	Used (%)	Used (%)	Remaining (%)	Blocks
baadalvm	0	In Service	70.37	0.29	8.37	61.71	0.41		87.69	5
baadalvm	1	In Service	70.37	0.62	7.06	62.68	0.88		89.08	12
baadalvm	1	In Service	70.37	0.91	7.06	62.4	1.29		88.67	17
baadalvm	1	In Service	70.37	0.91	7.08	62.38	1.29		88.65	17

Figure 3: Health Status after restoration

3. Following are the tables corresponding to the state of File blocks after replication.

Block ID	Block distribution File 1 after restoration		
4664252444576818739	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010
-140921271508634830	10.17.50.156:50010	10.17.3.64:50010	10.17.3.64:50010
-954538023745594084	10.17.50.156:50010	10.17.6.25:50010	10.17.50.25:50010
3311543059387042653	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010

Block ID	Block distribution File 2 after restoration		
-9108694018629044460	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010
2617830728431356278	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010
3778736967815178453	10.17.50.156:50010	10.17.3.64:50010	10.17.50.25:50010
4436040171665550495	10.17.50.156:50010	10.17.3.64:50010	10.17.50.25:50010

Block ID	Block distribution File 3 after restoration		
604131027942071764	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010
6844246988529318064	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010
-8508448695289813430	10.17.50.156:50010	10.17.3.64:50010	10.17.6.25:50010
8965565501969458440	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010

Block ID	Block distribution File 4 after restoration		
-7012850736704884080	10.17.50.156:50010	10.17.3.64:50010	10.17.50.25:50010
5117826362547550415	10.17.50.156:50010	10.17.6.25:50010	10.17.3.64:50010
3227726318878224303	10.17.50.156:50010	10.17.3.64:50010	10.17.50.25:50010
-196786929323127591	10.17.50.156:50010	10.17.3.64:50010	10.17.6.25:50010

4. Notice that the files are replicated back to 3 copies.