

LAPORAN PRAKTIKUM INFRASTRUKTUR BIG DATA
PERTEMUAN 4
HDFS



Oleh :

Nama : Jauhari Ahmad
No. Mhs : 205411167
Jurusan : Teknik Informatika / S1

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
AKAKOM
YOGYAKARTA
2020

PENDAHULUAN

A. TUJUAN

Mahasiswa menggunakan hdfs.

B. DASAR TEORI

Hadoop Distributed File System atau HDFS adalah sistem file terdistribusi berbasis Java yang memungkinkan untuk menyimpan data besar di beberapa node dalam cluster Hadoop. Jadi, jika menginstal Hadoop, maka akan mendapatkan HDFS sebagai sistem penyimpanan yang mendasari untuk menyimpan data di lingkungan terdistribusi.

Apache HDFS atau Hadoop Distributed File System adalah sistem file berstruktur blok di mana setiap file dibagi menjadi beberapa blok dengan ukuran yang telah ditentukan sebelumnya. Blok-blok ini disimpan di sekumpulan satu atau beberapa mesin. Arsitektur Apache Hadoop HDFS mengikuti Arsitektur Master / Slave, di mana sebuah cluster terdiri dari satu NameNode (Master node) dan semua node lainnya adalah DataNodes (node Slave). HDFS dapat digunakan di berbagai mesin yang mendukung Java. Meskipun seseorang dapat menjalankan beberapa DataNodes pada satu mesin, tetapi dalam dunia praktis, DataNode ini tersebar di berbagai mesin.

NameNode adalah node master dalam Arsitektur Apache Hadoop HDFS yang memelihara dan mengelola blok yang ada di DataNodes (node slave). NameNode adalah server yang memiliki ketersediaan sangat tinggi yang bertugas mengelola File System Namespace dan mengontrol akses ke file oleh klien. Arsitektur HDFS dibuat sedemikian rupa sehingga data pengguna tidak pernah berada di NameNode. Data berada di DataNodes saja. DataNodes adalah node slave di HDFS. Tidak seperti NameNode, DataNode adalah komoditas perangkat keras, yaitu sistem yang tidak mahal yang tidak berkualitas

tinggi atau ketersediaan tinggi. DataNode adalah server blok yang menyimpan data di file lokal ext3 atau ext4

PEMBAHASAN PRAKTIK

1. Menjalankan sistem operasi Linux Ubuntu yang sudah terinstal hadoop dan menjalankan Hadoop.

```
hadoop@jauhmad-VirtualBox: ~
jauhmad@jauhmad-VirtualBox:~$ su - hadoop
Password:
hadoop@jauhmad-VirtualBox:~$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [jauhmad-VirtualBox]
hadoop@jauhmad-VirtualBox:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@jauhmad-VirtualBox:~$
```

2. Donwload file profile.txt dan disimpan dalam /home/hadoop/
3. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls /
hadoop@jauhmad-VirtualBox:~$
```

Perintah ini digunakan untuk menampilkan daftar file dari HDFS, hasilnya kosong karena tidak ada file pada /

4. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -put profile.txt /
2020-09-28 19:41:37,441 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false
, remoteHostTrusted = false
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls /
Found 1 items
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
hadoop@jauhmad-VirtualBox:~$
```

Perintah ini digunakan untuk meng-copy file dari sistem file /home/hadoop/profile.txt ke / pada sistem file HDFS

5. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -cat /profile.txt
2020-10-05 07:19:02,526 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
Profil STMIK AKAKOM

Dalam abad ke-20 ini dunia banyak diwarnai dengan berbagai kemajuan secara menakjubkan yang telah diciptakan oleh umat manusia, baik di bidang ilmu maupun teknologi. Menanggapi kondisi zaman seperti tersebut di atas, pada tanggal 30 Juni 1979 didirikan sebuah yayasan dengan nama Yayasan Pendidikan Widy a Bakti, yang bertujuan mengembangkan dan menyebarkan informatika dan teknologi komputer di kalangan masyarakat Indonesia melalui usaha pendidikan yang sistematis dan ilmiah. Yayasan tersebut mengelola sebuah akademi yang bernama Akademi Aplikasi Komputer, disingkat AKAKOM.

Terhitung mulai 1 Maret 1983, Akademi Aplikasi Komputer (AKAKOM), diubah menjadi Akademi Komputer dan Informatika AKAKOM. Selanjutnya terhitung mulai tanggal 2 Mei 1985, nama Akademi Komputer dan Informatika AKAKOM diubah dan dibakukan menjadi Akademi Manajemen Informatika dan Komputer (AMIK) AKAKOM.

Agar lembaga tersebut mampu menghasilkan tenaga-tenaga profesional maupun akademik yang lebih berbobot dalam bidang informatika dan komputer, maka sejak tanggal 8 Juni 1992, berdasarkan Surat Keputusan Direktorat Jendral Pendidikan Tinggi Departemen Pendidikan dan Kebudayaan RI Nomor 262/DIKTI/kep/199 2, AMIK AKAKOM diubah bentuknya menjadi Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) AKAKOM yang membuka program sarjana dan program diploma.

Saat ini STMIK AKAKOM Yogyakarta mempunyai 2 program sarjana (Prodi. Teknik Informatika dan Prodi. Sistem Informasi) dan 3 program diploma (Prodi. Manajemen Informatika, Prodi. Teknik Komputer, dan Prodi. Komputerisasi Akuntansi) dengan masing-masing status akreditasi B.
```

Perintah ini digunakan untuk menampilkan isi teks dari file profile.txt

6. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -cp /profile.txt /copy-profile.txt
2020-10-05 07:22:21,017 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2020-10-05 07:22:21,537 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
```

Perintah ini digunakan untuk meng-copy file profile.txt ke file copy-profile.txt pada HDFS

Printah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls /
Found 2 items
-rw-r--r-- 1 hadoop supergroup 1637 2020-10-05 07:22 /copy-profile.txt
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
```

Perintah ini digunakan untuk list file pada direktori / di sistem HDFS

Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -cat /copy-profile.txt
2020-10-05 07:27:17,651 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
Profil STMIK AKAKOM

Dalam abad ke-20 ini dunia banyak diwarnai dengan berbagai kemajuan secara menakjubkan yang telah diciptakan oleh umat manusia, baik di bidang ilmu maupun teknologi. Menanggapi kondisi zaman seperti tersebut di atas, pada tanggal 30 Juni 1979 didirikan sebuah yayasan dengan nama Yayasan Pendidikan Widya Bakti, yang bertujuan mengembangkan dan menyebarluaskan informatika dan teknologi komputer di kalangan masyarakat Indonesia melalui usaha pendidikan yang sistematis dan ilmiah. Yayasan tersebut mengelola sebuah akademi yang bernama Akademi Aplikasi Komputer, disingkat AKAKOM.

Terhitung mulai 1 Maret 1983, Akademi Aplikasi Komputer (AKAKOM), diubah menjadi Akademi Komputer dan Informatika AKAKOM. Selanjutnya terhitung mulai tanggal 2 Mei 1985, nama Akademi Komputer dan Informatika AKAKOM diubah dan dibakukan menjadi Akademi Manajemen Informatika dan Komputer (AMIK) AKAKOM.

Ager lembaga tersebut mampu menghasilkan tenaga-tenaga profesional maupun akademik yang lebih berbobot dalam bidang informatika dan komputer, maka sejak tanggal 8 Juni 1992, berdasarkan Surat Keputusan Direktorat Jendral Pendidikan Tinggi Departemen Pendidikan dan Kebudayaan RI Nomor 262/DIKTI/Kep/1992, AMIK AKAKOM diubah bentuknya menjadi Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) AKAKOM yang membuka program sarjana dan program diploma.

Saat ini STMIK AKAKOM Yogyakarta mempunyai 2 program sarjana (Prodi. Teknik Informatika dan Prodi. Sistem Informasi) dan 3 program diploma (Prodi. Manajemen Informatika, Prodi. Teknik Komputer, dan Prodi. Komputerisasi Akuntansi) dengan masing-masing status akreditasi B.
```

Perintah ini digunakan untuk menampilkan teks file copy-profile.txt

7. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -mkdir /modul4
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls /
Found 3 items
-rw-r--r-- 1 hadoop supergroup 1637 2020-10-05 07:22 /copy-profile.txt
drwxr-xr-x - hadoop supergroup 0 2020-10-05 07:28 /modul4
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
```

Perintah ini digunakan untuk membuat direktori “modul4” pada direktori / dan menampilkan list file dan direktori pada /

8. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -mv /copy-profile.txt /modul4/copy-profile.txt
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls /
Found 2 items
drwxr-xr-x - hadoop supergroup 0 2020-10-05 07:31 /modul4
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
```

Perintah ini digunakan untuk memindah file copy-profile.txt dari direktori / ke direktori /modul4, sehingga file copy-profile.txt sudah hilang dari direktori / karena sudah dipindah

9. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls /modul4/
Found 1 items
-rw-r--r-- 1 hadoop supergroup 1637 2020-10-05 07:22 /modul4/copy-profile.txt
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls -R /
drwxr-xr-x - hadoop supergroup 0 2020-10-05 07:31 /modul4
-rw-r--r-- 1 hadoop supergroup 1637 2020-10-05 07:22 /modul4/copy-profile.txt
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
```

Perintah pertama digunakan untuk menampilkan list file pada direktori /modul4 dan berisi file copy-profile.txt, kemudian pada perintah kedua digunakan untuk menampilkan list file sampai ke sub-sub direktori sekaligus pada direktori /

10. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -mv /modul4/copy-profile.txt /modul4/profile-2.txt
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls -R /
drwxr-xr-x - hadoop supergroup 0 2020-10-05 07:40 /modul4
-rw-r--r-- 1 hadoop supergroup 1637 2020-10-05 07:22 /modul4/profile-2.txt
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
```

Perintah pertama digunakan untuk memindah sekaligus mengubah nama file, dari /modul4/copy-profile.txt ke /modul4/profile-2.txt

Sehingga setelah di cek pada perintah kedua, file yang eksis adalah /modul4/profile-2.txt karena sudah di *rename* file nya.

11. Perintah:

```
hadoop@jauhmad-VirtualBox:~$ ls
dfsdata  examples.desktop  hadoop-3.2.1  hadoop-3.2.1.tar.gz  mapper.py  profile.txt  reducer.py  tmpdata
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -get /modul4/profile-2.txt
2020-10-05 07:47:28,649 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
hadoop@jauhmad-VirtualBox:~$ ls
dfsdata  examples.desktop  hadoop-3.2.1  hadoop-3.2.1.tar.gz  mapper.py  profile-2.txt  profile.txt  reducer.py  tmpdata
```

Perintah ini digunakan untuk download file /modul4/profile-2.txt yang ada di HDFS ke direktori lokal sistem

12. Perintah

```
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls -R /
drwxr-xr-x - hadoop supergroup 0 2020-10-05 07:40 /modul4
-rw-r--r-- 1 hadoop supergroup 1637 2020-10-05 07:22 /modul4/profile-2.txt
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -rm /modul4/profile-2.txt
Deleted /modul4/profile-2.txt
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -rm -r /modul4
Deleted /modul4
hadoop@jauhmad-VirtualBox:~$ hdfs dfs -ls -R /
-rw-r--r-- 1 hadoop supergroup 1637 2020-09-28 19:41 /profile.txt
```

Perintah ini digunakan untuk menghapus file /modul4/profile-2.txt, kemudian menghapus direktori /modul4 dan isinya, sehingga setelah dicek pada perintah ke 4 direktori /modul4 dan isinya sudah hilang.

KESMIPULAN

1. Dalam HDFS terdapat perintah-perintah dasar operasi file (copy, pindah, rename, download)
2. Untuk mengawali perintah HDFS didahului perintah `hdfs dfs` diikuti perintah yang dimaksud
