LAPORAN PERTEMUAN 3

BIG DATA

MOCHAMMAD ZAKARO AL FAJRI

2241720175

TI – 3D

D-IV TEKNIK INFORMATIKA

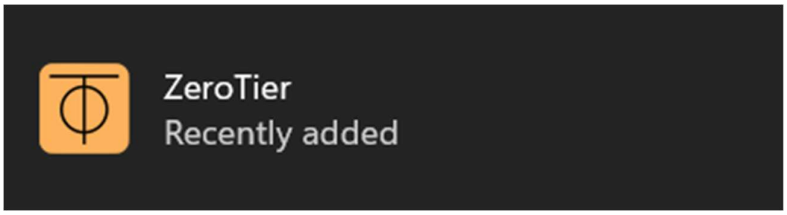
POLITEKNIK NEGERI MALANG

2025/2026

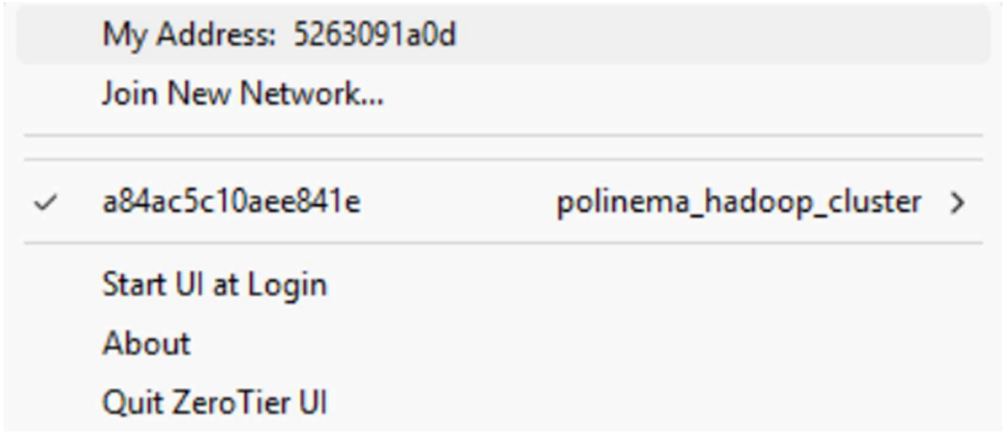
Daftar Isi

Setup HDFS

a. Langkah 1 : Pengunduhan ZeroTier VPN
   - Tujuan : Melakukan pengunduhan ZeroTier sebagai solusi VPN gratis untuk membuat koneksi aman antara komputer lokal dan NameNode.
   - Dengan VPN seperti ZeroTier, kita bisa mengakses NameNode Hadoop dari mana saja seolah-olah sedang berada di jaringan lokal. Ini mempermudah kita dalam mengelola dan menjalankan perintah Hadoop tanpa ribet, sekaligus menjaga koneksi tetap aman.
   - Bukti :



b. Langkah 2 : Konfigurasi VPN Client
   - Tujuan : Melakukan konfigurasi ZeroTier VPN agar dapat melakukan akses ke network dan cluster yang ditentukan yaitu polinema_hadoop_cluster
   - Hal yang perlu dilakukan yaitu memasukkan network id "**a84ac5c10aee841e**" di tempat yang ditentukan. Jika sudah kita hanya perlu melakukan join ke network yang ada.
   - Bukti :



c. Langkah 3 : Menguji VPN Client
   - Tujuan : Memastikan bahwa komputer telah berhasil terhubung ke jaringan VPN dan Menguji koneksi ke cluster Hadoop dengan melakukan ping ke IP yang diberikan.
   - Untuk menguji apakah Anda sudah terhubung dengan jaringan VPN cluster Hadoop ini, bukalah CMD atau terminal di computer Anda kemudian lakukan perintah PING pada IP **172.29.247.62**
   - Bukti :

```
C:\Users\KAKA>ping 172.29.247.62

Pinging 172.29.247.62 with 32 bytes of data:
Reply from 172.29.247.62: bytes=32 time=518ms TTL=64
Reply from 172.29.247.62: bytes=32 time=319ms TTL=64
Reply from 172.29.247.62: bytes=32 time=330ms TTL=64
Reply from 172.29.247.62: bytes=32 time=432ms TTL=64

Ping statistics for 172.29.247.62:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 319ms, Maximum = 518ms, Average = 399ms

C:\Users\KAKA>
```

Mengakses Cluster Hadoop

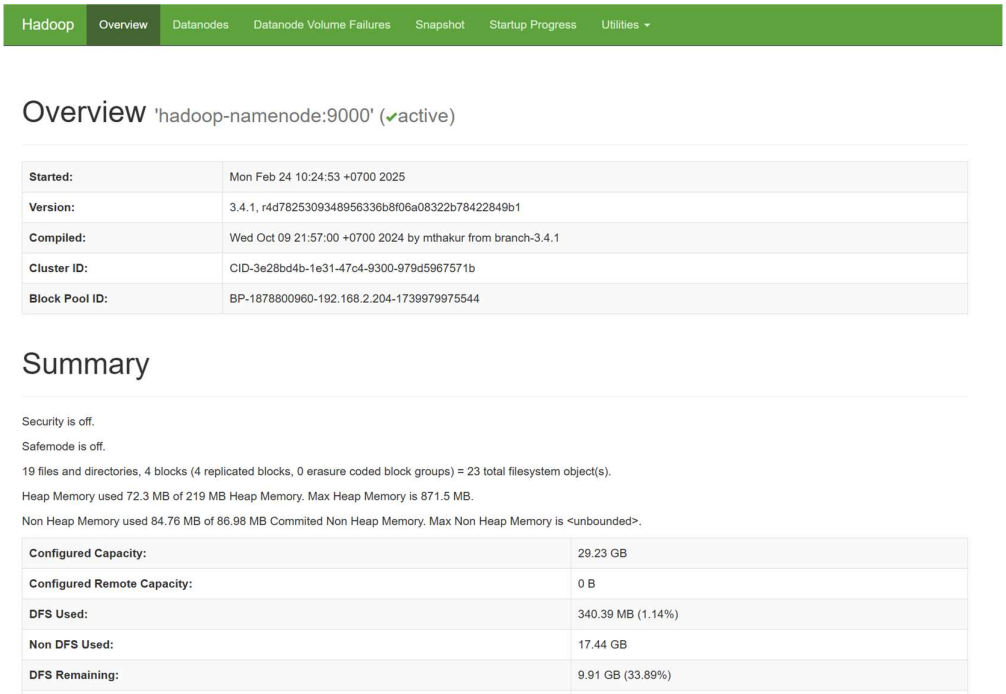
a. Langkah 1 : Masuk pada cluster Hadoop
   - Tujuan : Mengakses cluster Hadoop dengan menghubungkan komputer ke NameNode melalui SSH.
   - Pada Langkah ini, kita perlu masuk ke cluster Hadoop melalui terminal / command prompt. Masukkan ssh **hadoopuser@172.29.253.67** dengan password : **Hadoop.** Nantinya kita akan terhubung ke cluster Hadoop
   - Bukti :

```
PS C:\Users\KAKA> ssh hadoopuser@172.29.247.62
The authenticity of host '172.29.247.62 (172.29.247.62)' can't be established.
ED25519 key fingerprint is SHA256:BjHVtsH6B+jEGxXCmJXyg4N7pbcsIvsSS+Eih0qkHOE.
This key is not known by any other names.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '172.29.247.62' (ED25519) to the list of known hosts.
hadoopuser@172.29.247.62's password:
Permission denied, please try again.
hadoopuser@172.29.247.62's password:
Welcome to Ubuntu 24.04.2 LTS (GNU/Linux 6.8.0-53-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

 System information as of Tue Feb 25 07:41:15 AM UTC 2025

  System load:  0.11              Processes:              129
  Usage of /:   62.6% of 9.74GB   Users logged in:        1
  Memory usage: 31%               IPv4 address for enp0s3: 192.168.2.148
  Swap usage:   0%

 * Strictly confined Kubernetes makes edge and IoT secure. Learn how MicroK8s
   just raised the bar for easy, resilient and secure K8s cluster deployment.

   https://ubuntu.com/engage/secure-kubernetes-at-the-edge

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.
```

b. Langkah 2 : Cek Status Namenode
   - Tujuan : Memastikan bahwa koneksi SSH ke NameNode berhasil dan mengecek apakah NameNode sedang berjalan dengan benar.
   - Disini kita hanya perlu melakukan satu perintah yaitu jps. Perintah jps digunakan untuk menampilkan daftar proses Java yang sedang berjalan di NameNode, termasuk layanan Hadoop seperti NameNode, DataNode, atau ResourceManager.
   - Bukti :

```
hadoopuser@hadoop-namenode:~$ hadoopuser@hadoop-namenode
hadoopuser@hadoop-namenode: command not found
hadoopuser@hadoop-namenode:~$ jps
22193 Jps
2067 ResourceManager
1655 NameNode
1887 SecondaryNameNode
hadoopuser@hadoop-namenode:~$
```

c. Langkah 3 : Cek status Datanodes
   - Tujuan : Memastikan DataNode Berfungsi dengan Baik dan Melakuka pemantauan singkat terkait performa dari cluster
   - DataNode adalah komponen dalam Hadoop Distributed File System (HDFS) yang bertugas menyimpan data dalam bentuk blok dan menangani operasi baca/tulis dari klien. Kita bisa melihat status data nodes saat ini melalui GUI Web yang dapat diakses melalui alamat: **http://172.29.247.62:9870/dfshealth.html**
   - Bukti :

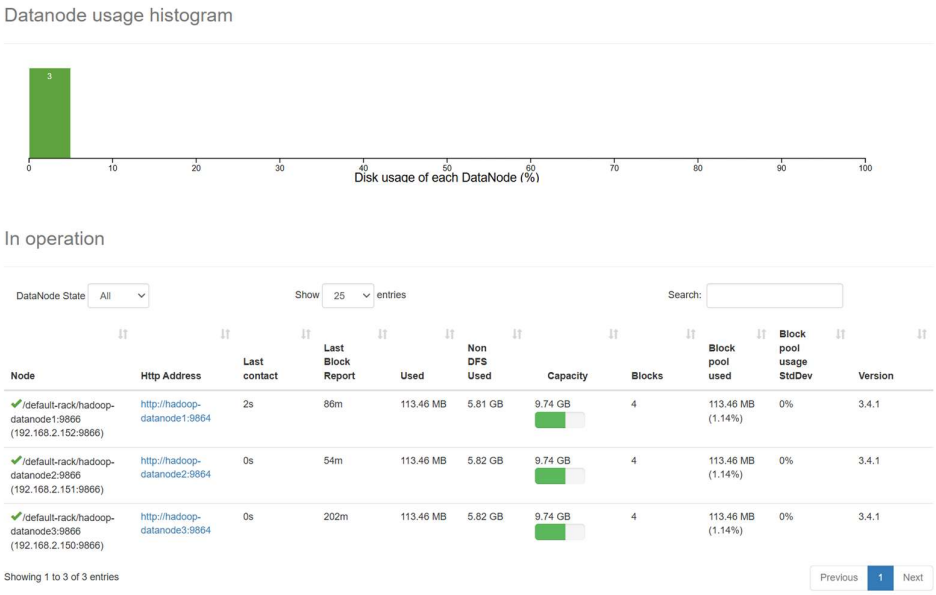


Penjelasan :

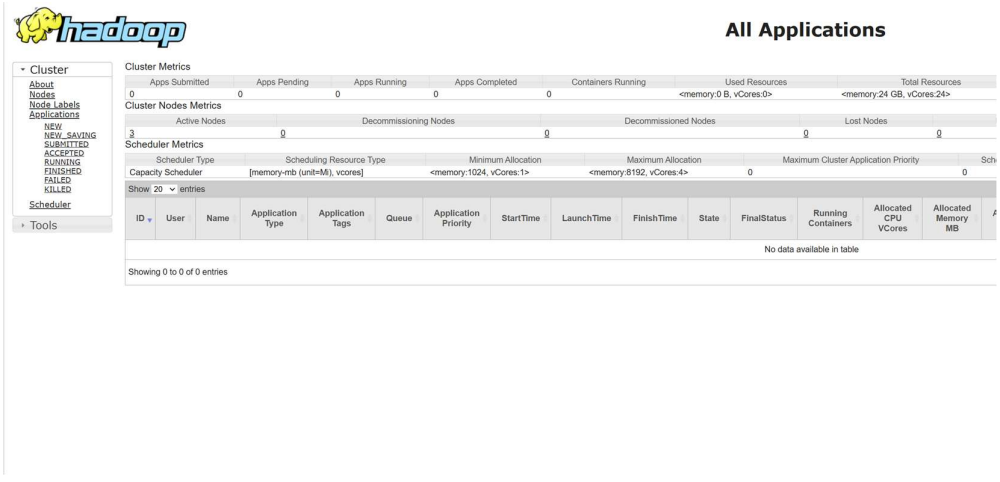
a) Bagian Overview
   - Menampilkan status NameNode, versi Hadoop yang digunakan, , serta informasi unik seperti Cluster ID dan Block Pool ID. Status (active) menunjukkan bahwa NameNode sedang berjalan dengan normal.

b) Bagian Summary
   - Menunjukkan penggunaan kapasitas penyimpanan di HDFS (Hadoop Distributed File System), termasuk total kapasitas, ruang yang digunakan, dan sisa ruang yang tersedia.

- Klik menu Datanodes untuk melihat informasi data node yang ada pada cluster.



d. Langkah 4 : Melihat Status Pemrosesan MapReduce
   - Tujuan : Memantau Status Pemrosesan MapReduce
   - Untuk mengakses MapReduce, kita hanya perlu membuka Halaman web **http://<IP_NameNode>:8088/cluster** Yang merupakan antarmuka Resource Manager. Antar muka ini berisikan informasi memungkinkan kita untuk melihat daftar pekerjaan yang sedang diproses oleh MapReduce, status masing-masing job, serta pemakaian sumber daya dalam cluster.
   - Bukti :



Latihan

Cobalah untuk:

a. Terhubung ke cluster Hadoop.
   Bukti :

```
PS C:\Users\KAKA> ssh hadoopuser@172.29.247.62
The authenticity of host '172.29.247.62 (172.29.247.62)' can't be established.
ED25519 key fingerprint is SHA256:BjHVtsH6B+jEGxXCmJXyg4N7pbcsIvsSS+Eih0qkHOE.
This key is not known by any other names.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added '172.29.247.62' (ED25519) to the list of known hosts.
hadoopuser@172.29.247.62's password:
Permission denied, please try again.
hadoopuser@172.29.247.62's password:
Welcome to Ubuntu 24.04.2 LTS (GNU/Linux 6.8.0-53-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

 System information as of Tue Feb 25 07:41:15 AM UTC 2025

  System load:  0.11              Processes:               129
  Usage of /:   62.6% of 9.74GB   Users logged in:         1
  Memory usage: 31%               IPv4 address for enp0s3: 192.168.2.148
  Swap usage:   0%

 * Strictly confined Kubernetes makes edge and IoT secure. Learn how MicroK8s
   just raised the bar for easy, resilient and secure K8s cluster deployment.

   https://ubuntu.com/engage/secure-kubernetes-at-the-edge

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.
```

b. Jalankan 3 perintah dasar yaitu
   - Melihat Versi Hadoop

```
hadoopuser@hadoop-namenode:~$ hadoop version
Hadoop 3.4.1
Source code repository https://github.com/apache/hadoop.git -r 4d7825309348956336b8f06a08322b78422849b1
Compiled by mthakur on 2024-10-09T14:57Z
Compiled on platform linux-x86_64
Compiled with protoc 3.23.4
From source with checksum 7292fe9dba5e2e44e3a9f763fce3e680
This command was run using /usr/local/hadoop/share/hadoop/common/hadoop-common-3.4.1.jar
```

   - Membuat Folder

```
hadoopuser@hadoop-namenode:~$ hadoop fs -mkdir /10_MochammadZakaroAlFajri
```

   - Melihat Folder-folder yang Ada

```
hadoopuser@hadoop-namenode:~$ hdfs dfs -ls /
Found 8 items
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-25 07:33 /05_FebbyMathelda
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-25 08:04 /10_MochammadZakaroAlFajri
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-25 07:53 /18_ShasiaSasaSalsabyla
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-25 07:40 /21_TriyanaDewiFatmawati
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-24 04:08 /TI-3E
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-25 07:39 /TriyanaDewiFatmawati
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-25 04:29 /hsycoba
drwxr-xr-x   - hadoopuser supergroup          0 2025-02-19 16:50 /yunhasnawa
```

   .