

**LAPORAN PRAKTIKUM INFRASTRUKTUR BIG DATA
PERTEMUAN 5
HBASE**



Oleh :

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

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
AKAKOM
YOGYAKARTA
2020**

PENDAHULUAN

A. TUJUAN

Mahasiswa melakukan instalasi HBase.

B. DASAR TEORI

HBase adalah sistem database terdistribusi open-source berorientasi kolom dalam lingkungan Hadoop. Pada awalnya HBase adalah Google Big Table, kemudian dinamai ulang sebagai HBase dan HBase ditulis di bahasa Java. Apache HBase diperlukan untuk aplikasi Big Data secara real time. HBase dapat menyimpan data dalam jumlah besar dari terabyte hingga petabyte. Tabel yang ada di HBase terdiri dari milyaran baris yang memiliki jutaan kolom. HBase dibuat untuk operasi latensi rendah, yang memiliki beberapa fitur spesifik dibandingkan dengan model relasional tradisional.

HBase adalah database berorientasi kolom dan data disimpan dalam tabel. Tabel diurutkan berdasarkan RowId. Seperti yang ditunjukkan di bawah ini, HBase memiliki RowId, yang merupakan kumpulan dari beberapa keluarga kolom yang ada dalam tabel.

Keluarga kolom yang ada dalam skema adalah pasangan nilai kunci. Jika diamati secara rinci setiap keluarga kolom memiliki banyak kolom. Nilai kolom disimpan ke dalam memori disk. Setiap sel tabel memiliki Metadata sendiri seperti stempel waktu dan informasi lainnya.

PEMBAHASAN PRAKTIKUM

1. Beralih ke user hadoop

```
jauhmad@jauhmad-VirtualBox:~$ su - hadoop
Password:
hadoop@jauhmad-VirtualBox:~$ cd /home/hadoop/
hadoop@jauhmad-VirtualBox:~$
```

2. Unduh source Hbase

```
hadoop@jauhmad-VirtualBox:~$ wget https://downloads.apache.org/hbase/stable/hbase-2.2.5-bin.tar.gz
--2020-10-05 20:08:53-- https://downloads.apache.org/hbase/stable/hbase-2.2.5-bin.tar.gz
Resolving downloads.apache.org (downloads.apache.org)... 88.99.95.219, 2a01:4f8:10a:201a::2
Connecting to downloads.apache.org (downloads.apache.org)|88.99.95.219|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 220221311 (210M) [application/x-gzip]
Saving to: 'hbase-2.2.5-bin.tar.gz'

hbase-2.2.5-bin.tar.gz      23%[=====>] 49,99M  139KB/s  eta 38m 18s^
hbase-2.2.5-bin.tar.gz    100%[=====>] 210,02M  110KB/s  in 50m 54s

2020-10-05 20:59:49 (70,4 KB/s) - 'hbase-2.2.5-bin.tar.gz' saved [220221311/220221311]
```

3. Ekstrak dan salin menjadi Hbase

```
hadoop@jauhmad-VirtualBox:~$ tar -zxvf hbase-2.2.5-bin.tar.gz
hbase-2.2.5/LICENSE.txt
hbase-2.2.5/NOTICE.txt
hbase-2.2.5/LLEGAL
hbase-2.2.5/CHANGES.md
hbase-2.2.5/README.txt
hbase-2.2.5/RELEASENOTES.md
hbase-2.2.5/conf/
hbase-2.2.5/conf/hadoop-metrics2-hbase.properties
hbase-2.2.5/conf/hbase-env.cmd
hbase-2.2.5/conf/hbase-env.sh
```

```
hadoop@jauhmad-VirtualBox:~$ cp hbase-2.2.5 HBase
cp: omitting directory 'hbase-2.2.5'
```

4. Edit file /home/hadoop/HBase/conf/hbase-env.sh

Tambah environment JAVA_HOME

```
GNU nano 2.5.3 File: /home/hadoop/HBase/conf/hbase-env.sh

# Extra ssh options. Empty by default.
# export HBASE_SSH_OPTS="-o ConnectTimeout=1 -o SendEnv=HBASE_CONF_DIR"

# Where log files are stored. $HBASE_HOME/logs by default.
# export HBASE_LOG_DIR=${HBASE_HOME}/Logs

# Enable remote JDWP debugging of major HBase processes. Meant for Core Developers
# export HBASE_MASTER_OPTS="$HBASE_MASTER_OPTS -Xdebug -Xrunjdpw:transport=dt_socket,server=y,transport=dt_socket,server=y"
# export HBASE_REGIONSERVER_OPTS="$HBASE_REGIONSERVER_OPTS -Xdebug -Xrunjdpw:transport=dt_socket,server=y,transport=dt_socket,server=y"
# export HBASE_THRIFT_OPTS="$HBASE_THRIFT_OPTS -Xdebug -Xrunjdpw:transport=dt_socket,server=y,transport=dt_socket,server=y"
# export HBASE_ZOOKEEPER_OPTS="$HBASE_ZOOKEEPER_OPTS -Xdebug -Xrunjdpw:transport=dt_socket,server=y,transport=dt_socket,server=y"
# export HBASE_REST_OPTS="$HBASE_REST_OPTS -Xdebug -Xrunjdpw:transport=dt_socket,server=y,transport=dt_socket,server=y"

# A string representing this instance of hbase. $USER by default.
# export HBASE_IDENT_STRING=$USER

# The scheduling priority for daemon processes. See 'man nice'.
# export HBASE_NICENESS=10

# The directory where pid files are stored. /tmp by default.
# export HBASE_PID_DIR=/var/hadoop/pids

# Seconds to sleep between slave commands. Unset by default. This
# can be useful in large clusters, where, e.g., slave rsyncs can
# otherwise arrive faster than the master can service them.
# export HBASE_SLAVE_SLEEP=0.1

# Tell HBase whether it should manage it's own instance of ZooKeeper or not.
# export HBASE_MANAGES_ZK=true

# The default log rolling policy is RFA, where the log file is rolled as per the size def
# RFA appender. Please refer to the log4j.properties file to see more details on this app
# In case one needs to do log rolling on a date change, one should set the environment pr
# HBASE_ROOT_LOGGER to "<DESIRED_LOG_LEVEL>,DRFA".
# For example:
# HBASE_ROOT_LOGGER=INFO,DRFA
# The reason for changing default to RFA is to avoid the boundary case of filling out dis
# DRFA doesn't put any cap on the log size. Please refer to HBase-5655 for more context.

# Tell HBase whether it should include Hadoop's lib when start up,
# the default value is false, means that includes Hadoop's lib.
# export HBASE_DISABLE_HADOOP_CLASSPATH_LOOKUP="true"
export JAVA_HOME=/opt/jdk1.8.0_261
```

5. Edit file /home/hadoop/Hbase/conf/hbase-site.xml

Konfigurasi diubah menjadi sbb:

```
GNU nano 2.5.3 File: /home/hadoop/HBase/conf/hbase-site.xml

<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
/*
 * Licensed to the Apache Software Foundation (ASF) under one
 * or more contributor license agreements. See the NOTICE file
 * distributed with this work for additional information
 * regarding copyright ownership. The ASF licenses this file
 * to you under the Apache License, Version 2.0 (the
 * "License"); you may not use this file except in compliance
 * with the License. You may obtain a copy of the License at
 *
 * http://www.apache.org/licenses/LICENSE-2.0
 *
 * Unless required by applicable law or agreed to in writing, software
 * distributed under the License is distributed on an "AS IS" BASIS,
 * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
 * See the License for the specific language governing permissions and
 * limitations under the License.
 */
-->
<configuration>
<property>
<name>hbase.rootdir</name>
<value>file:/home/hadoop/HBase/HFiles</value>
</property>
<property>
<name>hbase.zookeeper.property.dataDir</name>
<value>/home/hadoop/zookeeper</value>
</property>
</configuration>
```

6. Membuat folder Hfiles dan zookeeper di /home/hadoop

```
hadoop@jauhmad-VirtualBox:~$ mkdir -p /home/hadoop/HBase/Hfiles
hadoop@jauhmad-VirtualBox:~$ mkdir -p /home/hadoop/zookeeper
hadoop@jauhmad-VirtualBox:~$ ls
dfsdata  examples.desktop  hadoop-3.2.1  HBase  hbase-2.2.5-bin.tar.gz  mapper.py  profile-2.txt  profile.txt  reducer.py  tmpdata  zookeeper
```

7. Edit file .bashrc

Tambahkan konfigurasi sbb:

```
#Hadoop Related Options
export HADOOP_HOME=/home/hadoop/hadoop-3.2.1
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"
#export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/"

export HBASE_HOME=/home/hadoop/HBase
export PATH=$PATH:$HBASE_HOME/bin
```

8. Jalankan perintah pada .bashrc untuk mengatur ulang environment instalasi Hadoop

9. Menjalankan server Hadoop dengan Hbase

```
jauhmad@jauhmad-VirtualBox:~$ su - hadoop
Password:
hadoop@jauhmad-VirtualBox:~$ hdfs namenode -format
2020-10-12 07:25:47,424 INFO namenode.NameNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = jauhmad-VirtualBox/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 3.2.1
```

```
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
```

KESMIPULAN

Hbase adalah database terdistribusi yang berorientasi pada kolom yang berjalan di atas HDFS
