

# **TAHAPAN INSTALASI APACHE SPARK DAN DEMO CODING APACHE SPARK UNTUK WORD COUNT**

Disusun untuk memenuhi tugas Mata Kuliah Big Data

Dosen Pengampu :

Prof. Dr. Lala Septem Riza, M.T.



Disusun Oleh :

Kelompok 10

Abdullah Hafiz Furqon (2202729)

Muhammad Iqbal Fadhilah (2202292)

Legi Kuswandi (2203335)

Ilmu Komputer C2 2022

**FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS PENDIDIKAN INDONESIA**

## Tahapan Instalasi Apache Spark

1. Buka terminal, lalu akses/login sebagai root

```
ahkf12@fedora:~$ sudo su
[sudo] password for ahkf12:
```

2. Buat direktori baru, dan masuk ke dalam foldernya

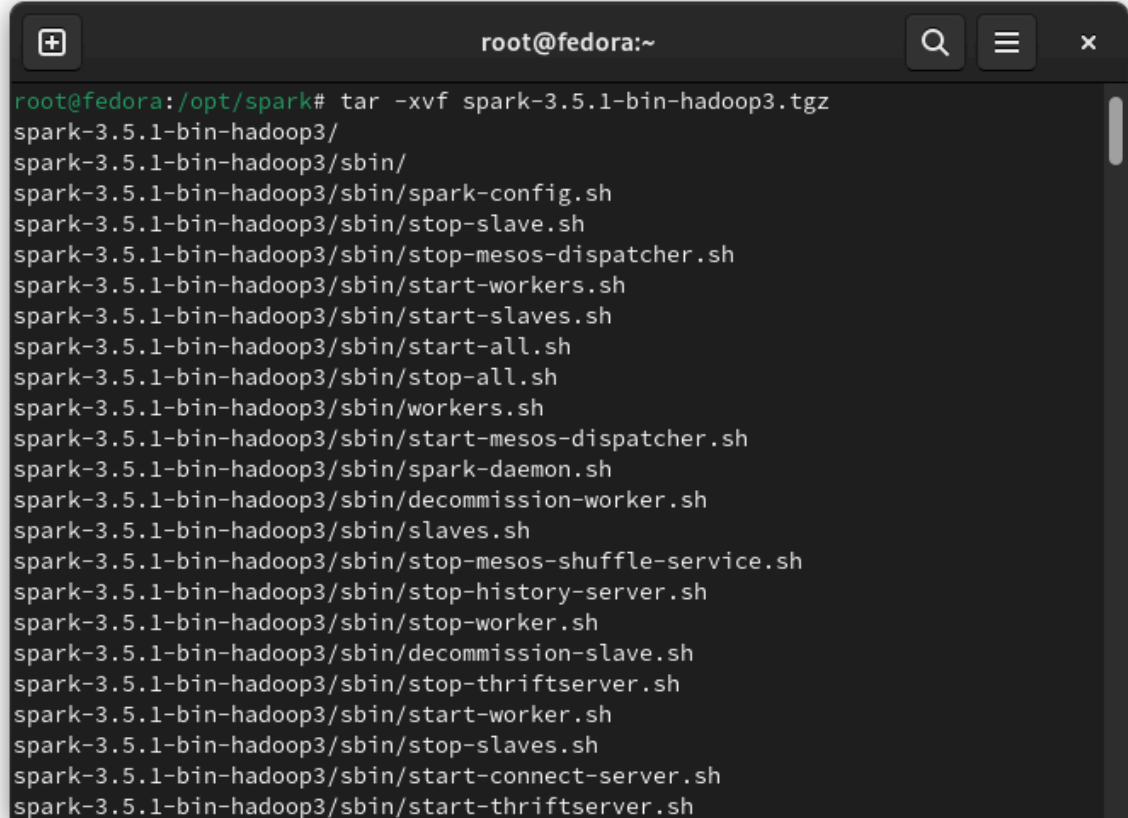
```
root@fedora:/home/ahkf12# mkdir -p /opt/spark
root@fedora:/home/ahkf12# cd /opt/spark
```

3. Download apache spark, melalui syntax wget (link download apache spark)

```
root@fedora:/opt/spark# wget https://dlcdn.apache.org/spark/spark-3.5.1/spark-3.5.1-bin-hadoop3.tgz
--2024-05-12 21:02:45-- https://dlcdn.apache.org/spark/spark-3.5.1/spark-3.5.1-bin-hadoop3.tgz
Resolving dlcdn.apache.org (dlcdn.apache.org)... 151.101.2.132, 2a04:4e42::644
Connecting to dlcdn.apache.org (dlcdn.apache.org)[151.101.2.132]:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 400446614 (382M) [application/x-gzip]
Saving to: 'spark-3.5.1-bin-hadoop3.tgz'

spark-3.5.1-bin-had  1%[          ]  4.10M  --.-KB/s  in 17m 43s
```

4. Extract file apache spark yang didownload tadi

A terminal window titled 'root@fedora:~' with search, menu, and close icons in the title bar. The terminal shows the command 'tar -xvf spark-3.5.1-bin-hadoop3.tgz' and its output, which lists the contents of the extracted archive. The contents include a directory 'spark-3.5.1-bin-hadoop3/' and a subdirectory 'spark-3.5.1-bin-hadoop3/sbin/' containing various shell scripts for managing the Spark daemon and workers.

```
root@fedora:/opt/spark# tar -xvf spark-3.5.1-bin-hadoop3.tgz
spark-3.5.1-bin-hadoop3/
spark-3.5.1-bin-hadoop3/sbin/
spark-3.5.1-bin-hadoop3/sbin/spark-config.sh
spark-3.5.1-bin-hadoop3/sbin/stop-slave.sh
spark-3.5.1-bin-hadoop3/sbin/stop-mesos-dispatcher.sh
spark-3.5.1-bin-hadoop3/sbin/start-workers.sh
spark-3.5.1-bin-hadoop3/sbin/start-slaves.sh
spark-3.5.1-bin-hadoop3/sbin/start-all.sh
spark-3.5.1-bin-hadoop3/sbin/stop-all.sh
spark-3.5.1-bin-hadoop3/sbin/workers.sh
spark-3.5.1-bin-hadoop3/sbin/start-mesos-dispatcher.sh
spark-3.5.1-bin-hadoop3/sbin/spark-daemon.sh
spark-3.5.1-bin-hadoop3/sbin/decommission-worker.sh
spark-3.5.1-bin-hadoop3/sbin/slaves.sh
spark-3.5.1-bin-hadoop3/sbin/stop-mesos-shuffle-service.sh
spark-3.5.1-bin-hadoop3/sbin/stop-history-server.sh
spark-3.5.1-bin-hadoop3/sbin/stop-worker.sh
spark-3.5.1-bin-hadoop3/sbin/decommission-slave.sh
spark-3.5.1-bin-hadoop3/sbin/stop-thriftserver.sh
spark-3.5.1-bin-hadoop3/sbin/start-worker.sh
spark-3.5.1-bin-hadoop3/sbin/stop-slaves.sh
spark-3.5.1-bin-hadoop3/sbin/start-connect-server.sh
spark-3.5.1-bin-hadoop3/sbin/start-thriftserver.sh
```

5. Konfigurasi pada file `.bashrc`, dan masukkan konfigurasi seperti di bawah ini

```

GNU nano 7.2                                .bashrc

# User specific aliases and functions

alias rm='rm -i'
alias cp='cp -i'
alias mv='mv -i'

export HADOOP_HOME=/home/hadoop/hadoop
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

SPARK_HOME=/opt/spark/spark-3.5.1-bin-hadoop3
export PATH=$PATH:$SPARK_HOME/bin:$SPARK_HOME/bin

```

6. Jangan lupa lakukan perintah `source .bashrc`, dan coba jalankan spark nya dengan `spark-shell`

```
root@fedora:~  
root@fedora:~# vi .bashrc  
root@fedora:~# nano .bashrc  
root@fedora:~# spark-shell  
Setting default log level to "WARN".  
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).  
24/05/13 06:27:21 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
Spark context Web UI available at http://fedora:4040  
Spark context available as 'sc' (master = local[*], app id = local-1715556441812).  
Spark session available as 'spark'.  
Welcome to
```

```
      ____  
     /___/\_-- ___ _---/_\_/__  
    _\ \/_ - \/_ - `/_ __/_'\_  
   /___/_.-_\_\_,_-/_/_/_\_\\ version 3.5.1  
    /\_/  
  
Using Scala version 2.12.18 (OpenJDK 64-Bit Server VM, Java 11.0.22)  
Type in expressions to have them evaluated.  
Type :help for more information.
```

```
scala>
```