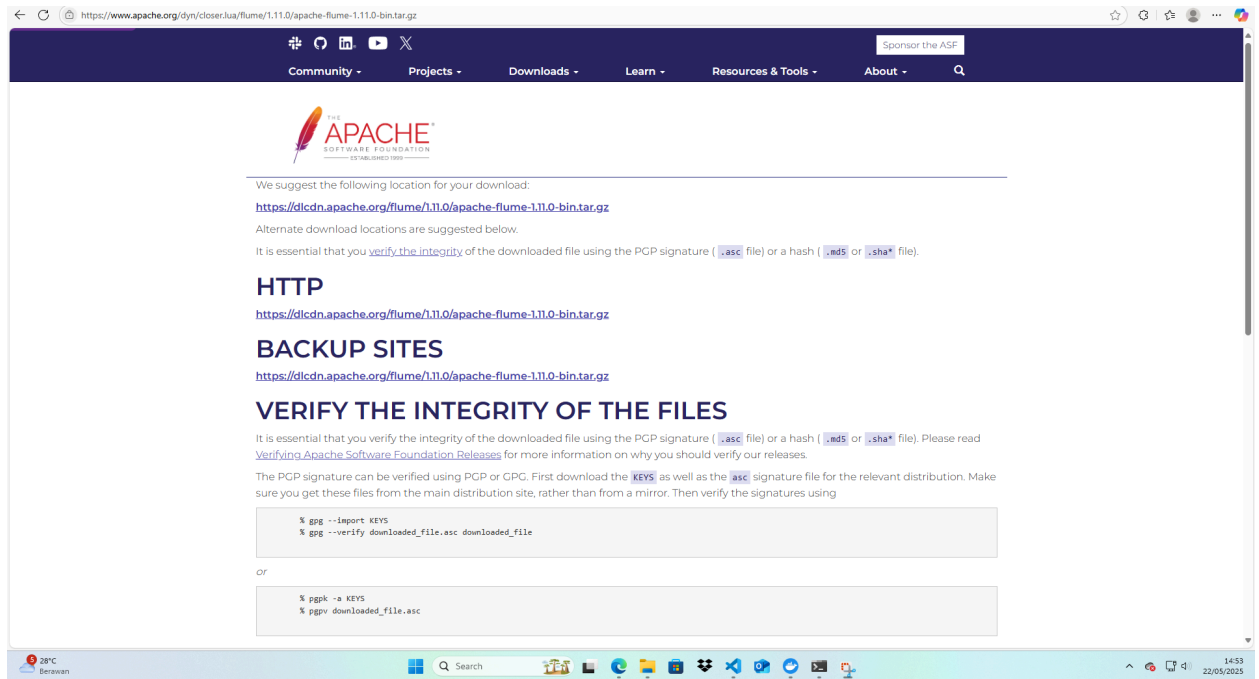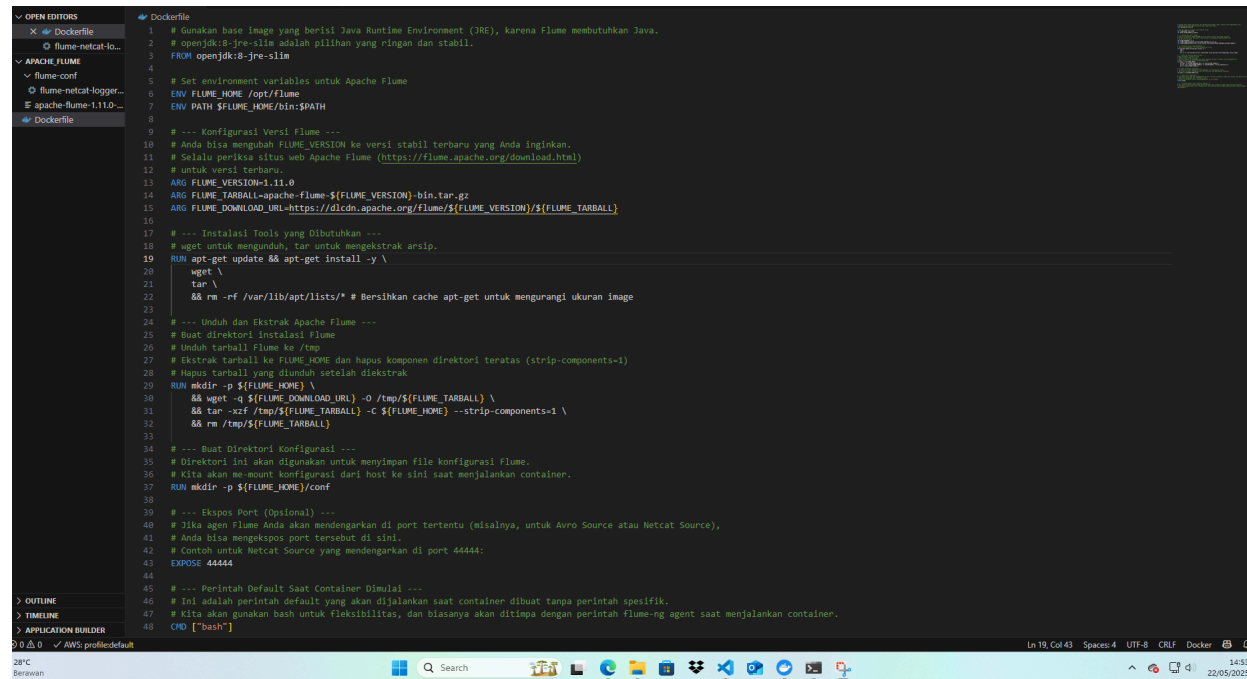# Instalasi dan konfigurasi apache flume

1. Download apache flume.tar
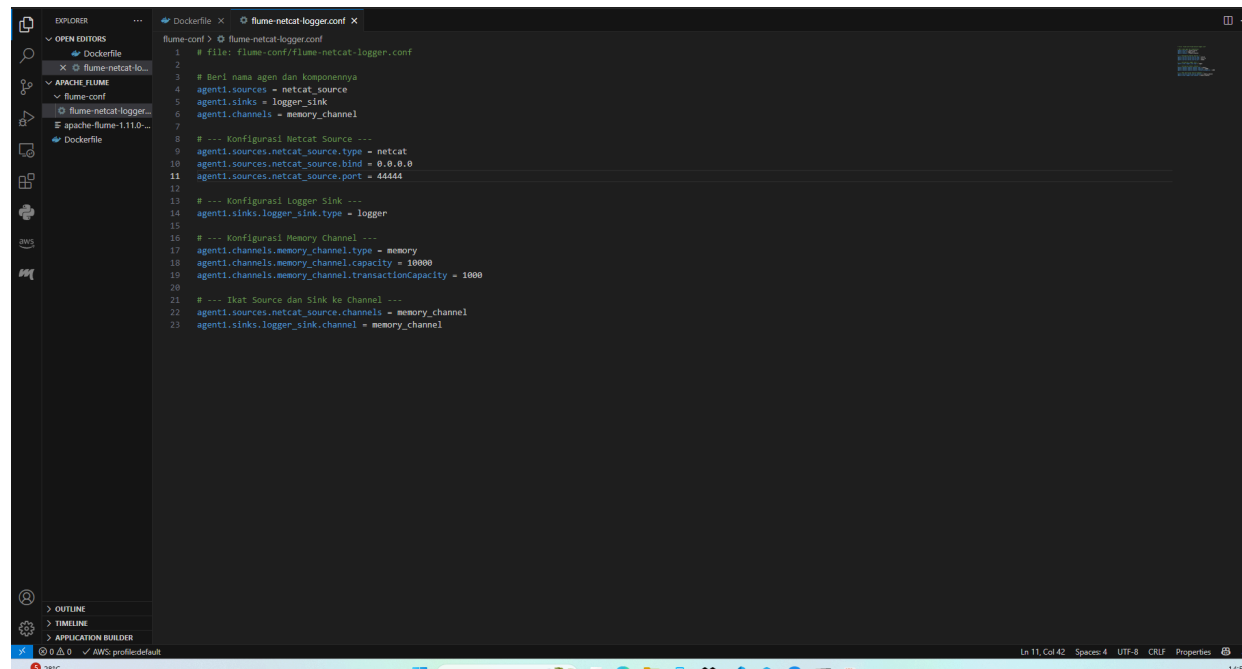

2. Buat file konfigurasi untuk apache flume untuk dijalankan di docker
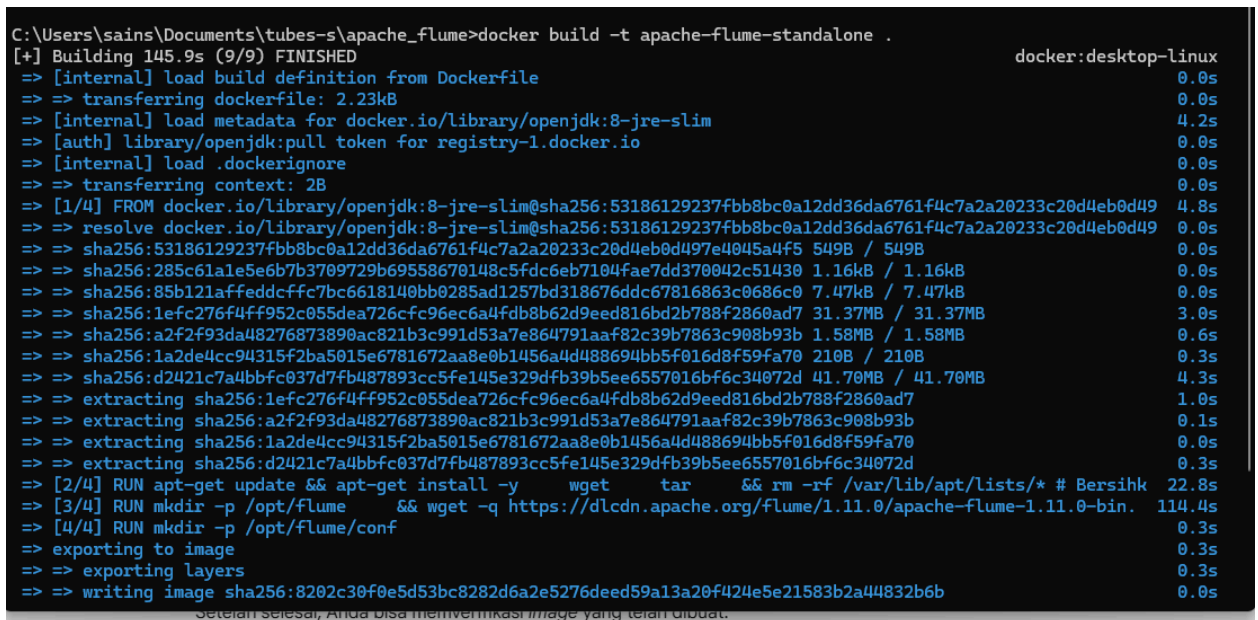   a. Buat file Dockerfile

b. Buat file konfigurasi flume

3. Build docker image untuk apache flume

4. Jalankan container


```
C:\Users\sains\Documents\tubes-s\apache_flume>docker run -d ^
More?    -p 44444:44444 ^
More?    --name flume_agent_01 ^
More?    -v "C:/Users/sains/Documents/tubes-s/apache_flume/flume-conf/flume-netcat-logger.conf":/opt/flume/conf/flume.con
f ^
More?    apache-flume-standalone ^
More?    /opt/flume/bin/flume-ng agent ^
More?    --conf /opt/flume/conf ^
More?    --conf-file /opt/flume/conf/flume.conf ^
More?    --name agent1 ^
More?    -Dflume.root.logger=INFO,console
9eab3fdbf8eb513edfa6671ca9f43fc97439f3cef812b49af86f509d6016d6d6

C:\Users\sains\Documents\tubes-s\apache_flume>docker ps
CONTAINER ID   IMAGE          COMMAND          CREATED       STATUS        PORTS
```

5. Masukkan data ke folder direktori apache flume

Name	Date modified	Type	Size
📁 data-input	22/05/2025 15:13	File folder	
📁 flume-conf	22/05/2025 15:13	File folder	
🗜 apache-flume-1.11.0-bin.tar	22/05/2025 14:49	Compressed Archi...	85.333 KB
📄 Dockerfile	22/05/2025 14:49	File	3 KB

Name	Date modified	Type	Size
📊 data_gempa	22/05/2025 14:43	Microsoft Excel C...	116 KB

6. Buat file yg berisi konfigurasi untuk memasukkan data csv ke apache flume yang akan diteruskan ke hdfs

7. Stop dan hapus container sebelumnya agar tidak ada konflik
   Selanjutnya jalankan lagi container yg baru yg sudah ada file datanya

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker stop flume_agent_01
flume_agent_01

C:\Users\sains\Documents\tubes-s\apache_flume>docker rm flume_agent_01
flume_agent_01

C:\Users\sains\Documents\tubes-s\apache_flume>docker run -d ^
More?    --name flume_csv_agent ^
More?    -v "C:/Users/sains/Documents/tubes-s/apache_flume/flume-conf/flume-csv-hdfs.conf":/opt/flume/conf/flume.conf ^
More?    -v "C:/Users/sains/Documents/tubes-s/apache_flume/data-input":/flume-data-input ^
More?    apache-flume-standalone ^
More?    /opt/flume/bin/flume-ng agent ^
More?    --conf /opt/flume/conf ^
More?    --conf-file /opt/flume/conf/flume.conf ^
More?    --name a1 ^
More?    -Dflume.root.logger=INFO,console
c3ce85943a59ba45a408a560c043a721ecbfd5059b8477804953d741c47c5c89

C:\Users\sains\Documents\tubes-s\apache_flume>
```

```
C:\Users\sains\Documents\tubes-s\apache_flume>docker ps
CONTAINER ID    IMAGE                   COMMAND              CREATED          STATUS             PORTS

                        NAMES
c3ce85943a59    apache-flume-standalone    "/opt/flume/bin/flum…"   About a minute ago   Up About a minute   44444/tcp

                        flume_csv_agent
e519c20be8f5    bigdata-hadoop              "/bin/bash -c '/boot…"   About an hour ago    Up About an hour    0.0.0.0:3306-
>3306/tcp, 0.0.0.0:8000->8000/tcp, 0.0.0.0:8088->8088/tcp, 0.0.0.0:9000->9000/tcp, 0.0.0.0:9866->9866/tcp, 0.0.0.0:9870-
>9870/tcp, 0.0.0.0:10000-10002->10000-10002/tcp, 0.0.0.0:50030->50030/tcp, 0.0.0.0:32768->1004/tcp, 0.0.0.0:32769->1006/
tcp, 0.0.0.0:32770->1180/tcp, 0.0.0.0:32771->8020/tcp, 0.0.0.0:32772->8030/tcp, 0.0.0.0:32773->8031/tcp, 0.0.0.0:32774->
8032/tcp, 0.0.0.0:32775->8033/tcp, 0.0.0.0:32776->8040/tcp, 0.0.0.0:32777->8041/tcp, 0.0.0.0:32778->8042/tcp, 0.0.0.0:32
779->9864/tcp, 0.0.0.0:32780->9867/tcp, 0.0.0.0:32781->10020/tcp, 0.0.0.0:32782->19888/tcp, 0.0.0.0:32783->50070/tcp, 0.
0.0.0:32784->50470/tcp    bigdata-hadoop

C:\Users\sains\Documents\tubes-s\apache_flume>
```

8. Menjalankan perintah untuk meneruskan data dari flume ke hdfs

9. Melakukan verifikasi apakah file sudah masuk ke hdfs atau belum

10. Datanya sudah berhasil masuk ke hdfs dengan format CSV