## Week 9 Assignment: Exploring Apache NiFi

### Objective:

Familiarize yourself with the capabilities of Apache NiFi by designing and implementing dataflows.

#### **1. Environment Initialization**

* Navigate to the NiFi directory:

cd nifi

* Start the NiFi Docker container:

docker-compose up -d

* Access the NiFi User Interface using the provided link:

[Access NiFi UI](https://localhost:8443/nifi)

**Note:** Use the username and password obtained in Week 1 to access the interface.

**Deliverable:** Screenshot confirming successful access to the NiFi UI.

#### **2. Creating a Processor Group**

**Exercise 1:** Drag and drop the “Processor Group” icon onto the canvas. Name this processor group “My First NiFi Flow”.

**Deliverable:** Screenshot of the NiFi canvas showing the Processor Group.

**Exercise 2:** Create a Parameter Context for your new processor group. Define a parameter (e.g., File\_Size) that will determine the size of the files generated in the next step.

**Deliverable:** Screenshot of the defined parameter within the Parameter Context.

#### **3. Designing a Simple Flow**

**Exercise 3:** Enter the “My First NiFi Flow” processor group by double-clicking it.

* Add the GenerateFlowFile processor to the canvas.
* Configure its properties:
  + Set File Size using the parameter (File\_Size) you defined earlier.
  + Adjust the Scheduling tab to run the processor every 5 seconds.
* Next, add the LogAttribute processor to the canvas.
* Connect GenerateFlowFile to LogAttribute. The relationship should be “success”.
* Start both processors and observe the flow of files.

**Deliverable:** Screenshot of the simple flow (GenerateFlowFile to LogAttribute).

#### **4. Setting Up Kafka**

* Navigate to the Kafka directory and initiate the Kafka Docker container, as instructed in the Kafka assignment:

cd kafka  
docker-compose up -d

**Exercise 4:** Create a topic named nifi-syslog.

/opt/kafka\_2.13-2.8.1/bin/kafka-topics.sh --create --topic nifi-syslog --bootstrap-server localhost:9092

**Deliverable:** Screenshot confirming the successful creation of the nifi-syslog topic.

#### **5. NiFi Advanced Flow with Kafka**

**Exercise 5:** Back in NiFi, on the NiFi canvas:

* Import the provided JSON file. This will load a pre-built dataflow onto your canvas.
* This flow will:
  + Generate log data.
  + Filter the logs with SQL.
  + Convert logs from the syslog format to JSON.
  + Publish the processed data to Kafka.
* Ensure that the Kafka producer processor in this flow is correctly configured to write data to the nifi-syslog topic.
* Before starting your simple flow, activate the advanced flow that sends data to Kafka to see the data appear in real-time on your Kafka consumer.

**Deliverable:** Screenshot of the advanced NiFi flow processing and sending data to Kafka.

#### **6. Consuming Data from Kafka**

**Exercise 6:** Use the Kafka consumer commands, as learned in the Kafka assignment, to consume data from the nifi-syslog topic.

/opt/kafka\_2.13-2.8.1/bin/kafka-console-consumer.sh --topic nifi-syslog --from-beginning --bootstrap-server localhost:9092

**Deliverable:** Screenshot of the Kafka consumer terminal showing the consumed data.