Assignment 10

Multi-Node Cassandra Cluster with OpsCenter & IoT Use Case

**1. Setup: Multi-Node Cassandra Cluster on Single Windows Machine**

**Tools Used:**

* Docker (on Windows)
* Cassandra Official Docker Image
* Custom Docker Network (cassandra-net)

**Step-by-Step Commands:**

**Step 1: Create Docker Network**

**docker network create cassandra-net**

**Step 2: Start Seed Node**

**docker run -d --name cassandra-seed --hostname cassandra-seed --network cassandra-net -e CASSANDRA\_CLUSTER\_NAME=2025GRP01 -e CASSANDRA\_NUM\_TOKENS=256 -e CASSANDRA\_SEEDS=cassandra-seed cassandra**

**Step 3: Start Additional Nodes**

docker run -d --name cassandra-node1 --hostname cassandra-node1 --network cassandra-net -e CASSANDRA\_CLUSTER\_NAME=2025GRP01 -e CASSANDRA\_SEEDS=cassandra-seed cassandra

docker run -d --name cassandra-node2 --hostname cassandra-node2 --network cassandra-net -e CASSANDRA\_CLUSTER\_NAME=2025GRP01 -e CASSANDRA\_SEEDS=cassandra-seed Cassandra

**Step 4: Validate Cluster**

Check status from seed node:

docker exec -it cassandra-seed nodetool status

Expected output:

* All 3 nodes with status UN (Up/Normal)
* Cluster Name: 2025GRP01

**2. Setup on Physical Machines (Lab Environment)**

**Setup:**

* 3+ Physical Machines (on same network)
* Install Cassandra on each machine
* Configure each node’s cassandra.yaml:
  + cluster\_name: 2025GRP01
  + Set listen\_address, rpc\_address to machine's IP
  + Point seeds to seed node IP

cluster\_name: '2025GRP01'

listen\_address: 192.168.1.10

rpc\_address: 192.168.1.10

seed\_provider:

- seeds: "192.168.1.10"

cassandra -f

**3. Install & Configure DataStax OpsCenter (Docker)**

**Step 1: Run OpsCenter in Docker**

docker run -d --name opscenter --network cassandra-net -p 8888:8888 -e OPSCENTER\_ACCEPT\_LICENSE=true datastax/dse-opscenter

**Step 2: Access Dashboard**

Open browser:

<http://localhost:8888>

**Step 3: Add Cluster in OpsCenter**

* Click **"Add Cluster"**
* Choose **Manual Setup**
* Enter seed node hostname: cassandra-seed
* Finish setup

**4. Use Case: Weather Station IoT Temperature Data**

**Problem Statement:**

Each weather station (weatherStationID) records temperature every 5 minutes and pushes to nearest Cassandra node.

**Step 1: Open cqlsh**

docker exec -it cassandra-seed cqlsh

Step 2: Create Keyspace

CREATE KEYSPACE weather\_data WITH replication = {

'class': 'SimpleStrategy',

'replication\_factor': 2

};

**Step 3: Use Keyspace**

USE weather\_data;

Step 4: Create Table

CREATE TABLE temperature\_by\_station (

weatherStationID text,

date date,

time time,

temperature float,

PRIMARY KEY ((weatherStationID, date), time)

) WITH CLUSTERING ORDER BY (time DESC);

Step 5: Insert Sample Data

INSERT INTO temperature\_by\_station (weatherStationID, date, time, temperature)

VALUES ('WS1001', '2025-04-20', '12:00:00', 27.5);

INSERT INTO temperature\_by\_station (weatherStationID, date, time, temperature)

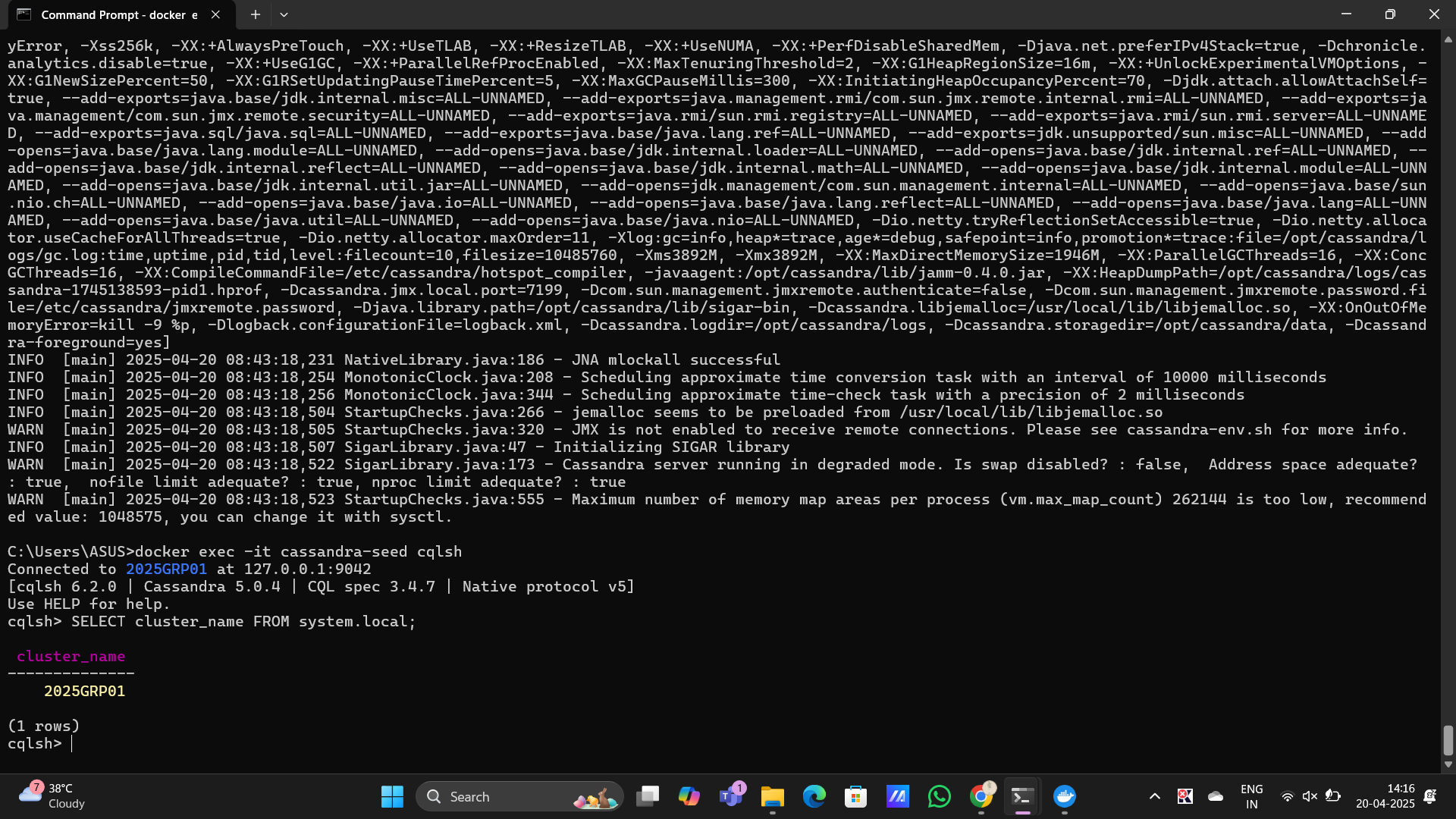
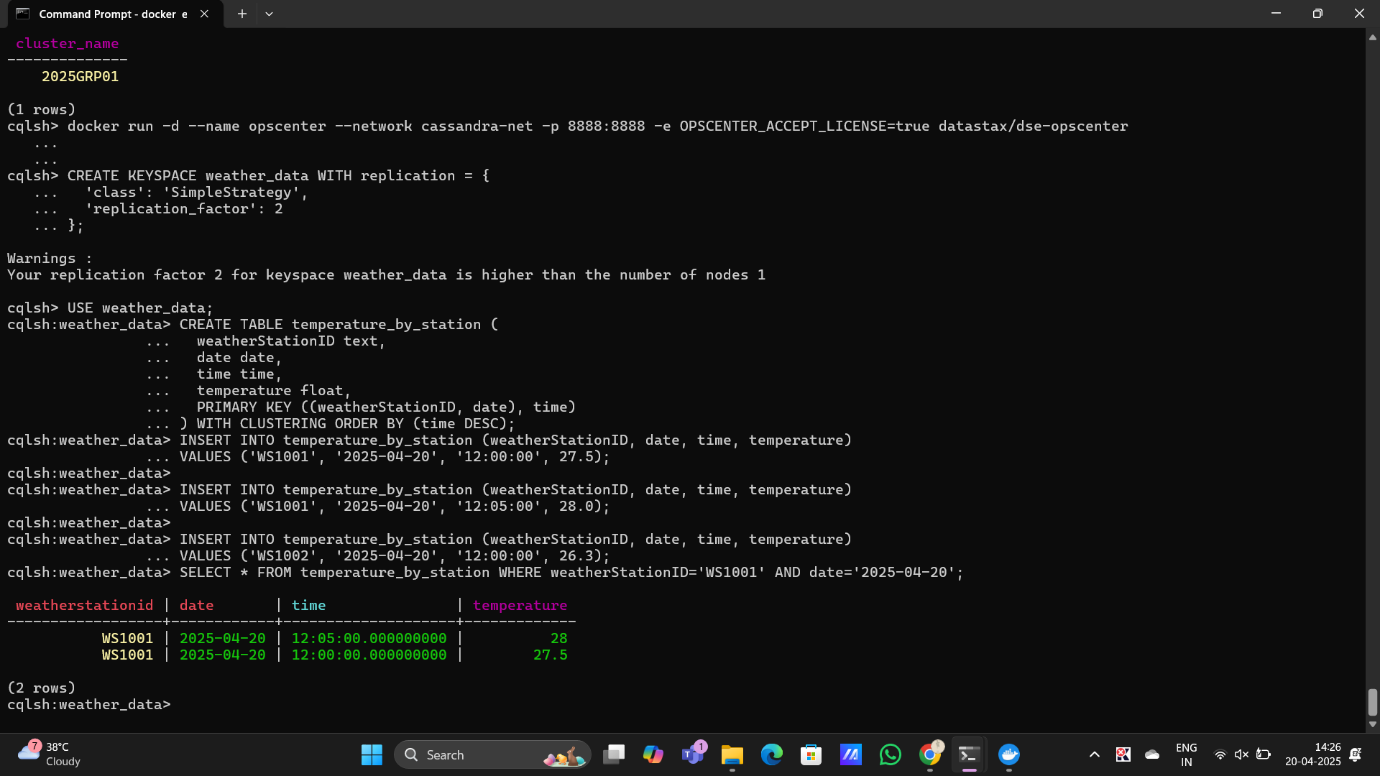
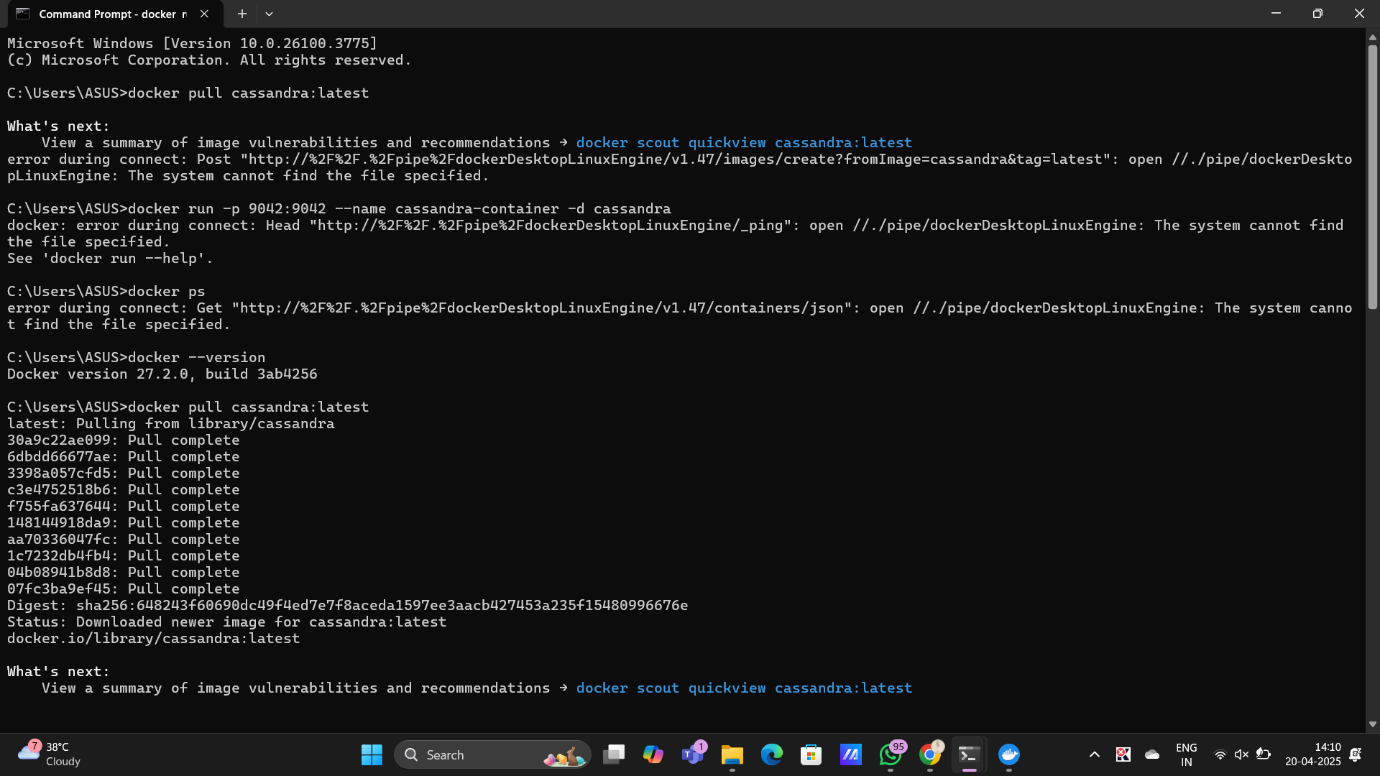
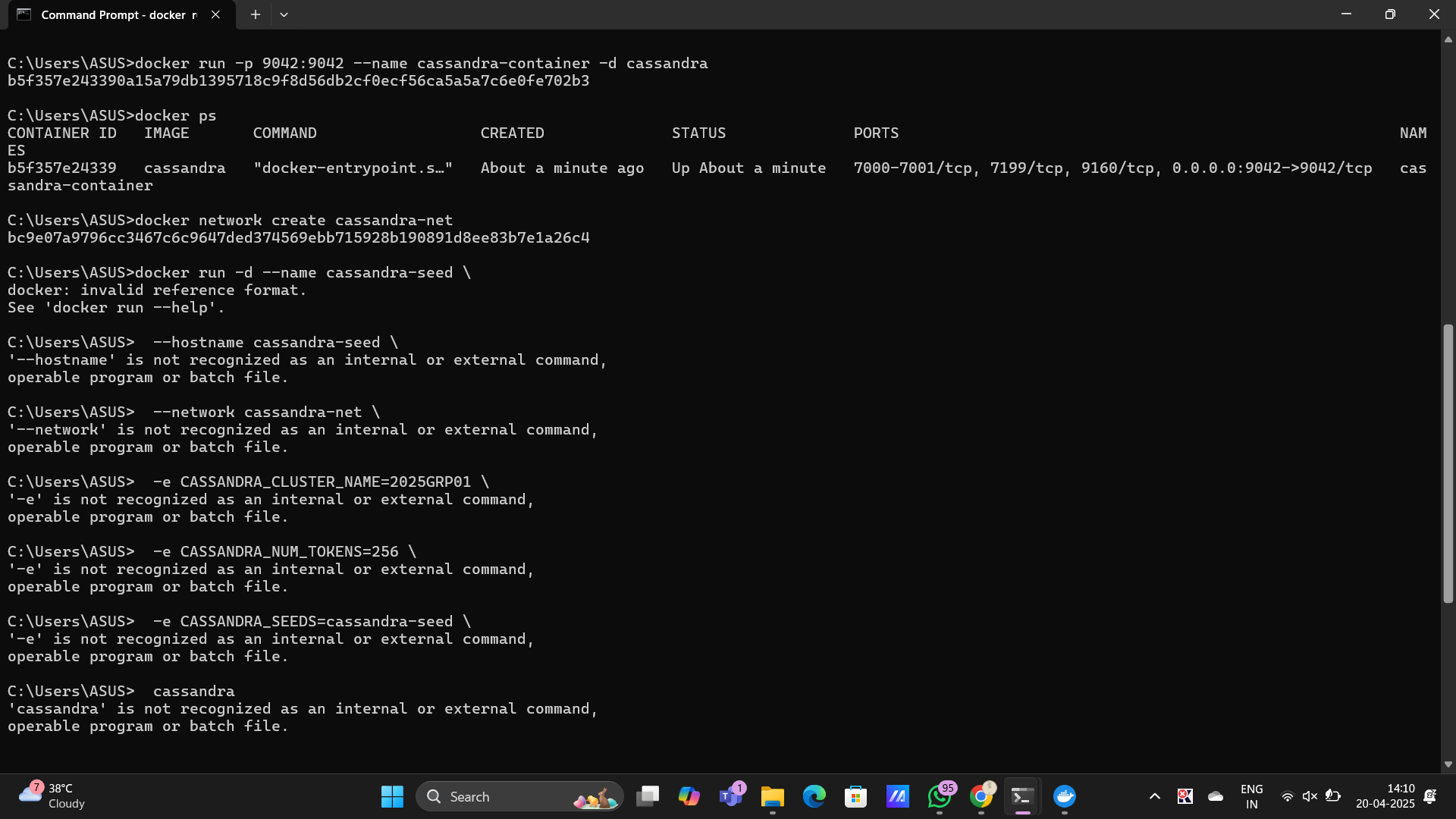
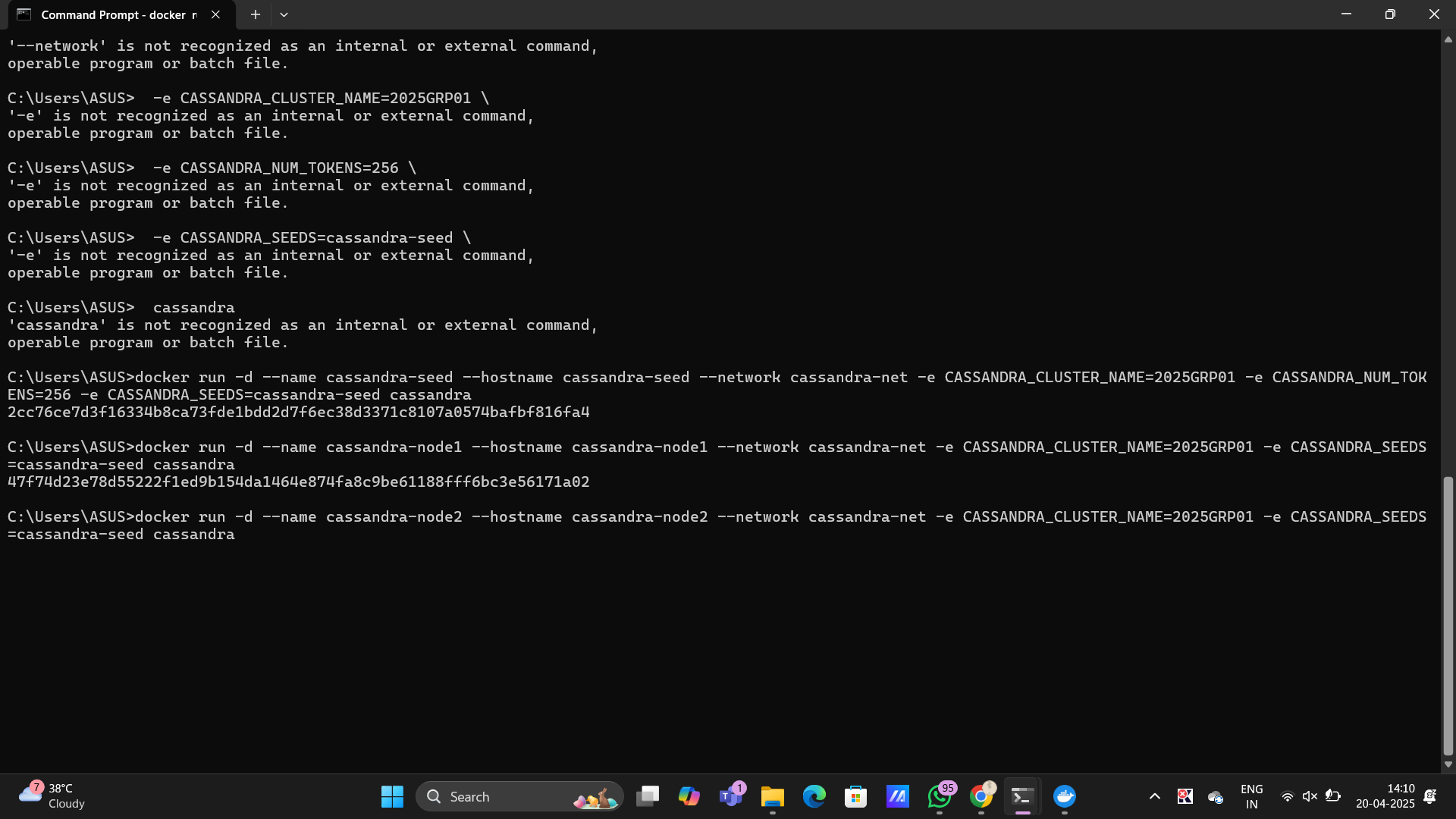
VALUES ('WS1001', '2025-04-20', '12:05:00', 28.0);

INSERT INTO temperature\_by\_station (weatherStationID, date, time, temperature)

VALUES ('WS1002', '2025-04-20', '12:00:00', 26.3);

**Step 6: Query the Data**

SELECT \* FROM temperature\_by\_station WHERE weatherStationID='WS1001' AND date='2025-04-20';



\*docker network create cassandra-net

\*seed node=>

docker run -d \

--name cassandra-seed \

--hostname cassandra-seed \

--network cassandra-net \

-e CASSANDRA\_CLUSTER\_NAME=2025GRP01 \

-e CASSANDRA\_NUM\_TOKENS=256 \

-e CASSANDRA\_SEEDS=cassandra-seed \

Cassandra

\***Step 3: Start 2 More Cassandra Nodes**

docker run -d \

--name cassandra-node1 \

--hostname cassandra-node1 \

--network cassandra-net \

-e CASSANDRA\_CLUSTER\_NAME=2025GRP01 \

-e CASSANDRA\_SEEDS=cassandra-seed \

Cassandra

\*4. 2nd node =>

docker run -d \

--name cassandra-node2 \

--hostname cassandra-node2 \

--network cassandra-net \

-e CASSANDRA\_CLUSTER\_NAME=2025GRP01 \

-e CASSANDRA\_SEEDS=cassandra-seed \

cassandra

\*step 4: Validate Cluster

docker exec -it cassandra-seed nodetool status

\*step 5:

\*docker exec -it cassandra-seed cqlsh

\*CREATE KEYSPACE weather\_data WITH replication = {

'class': 'SimpleStrategy',

'replication\_factor': 2

};

\*USE weather\_data;

\*CREATE TABLE temperature\_by\_station (

weatherStationID text,

date date,

time time,

temperature float,

PRIMARY KEY ((weatherStationID, date), time)

) WITH CLUSTERING ORDER BY (time DESC);

\*INSERT INTO temperature\_by\_station (weatherStationID, date, time, temperature)

VALUES ('WS1001', '2025-04-20', '12:00:00', 27.5);

INSERT INTO temperature\_by\_station (weatherStationID, date, time, temperature)

VALUES ('WS1001', '2025-04-20', '12:05:00', 28.0);

INSERT INTO temperature\_by\_station (weatherStationID, date, time, temperature)

VALUES ('WS1002', '2025-04-20', '12:00:00', 26.3);

\*SELECT \* FROM temperature\_by\_station WHERE weatherStationID='WS1001' AND date='2025-04-20';