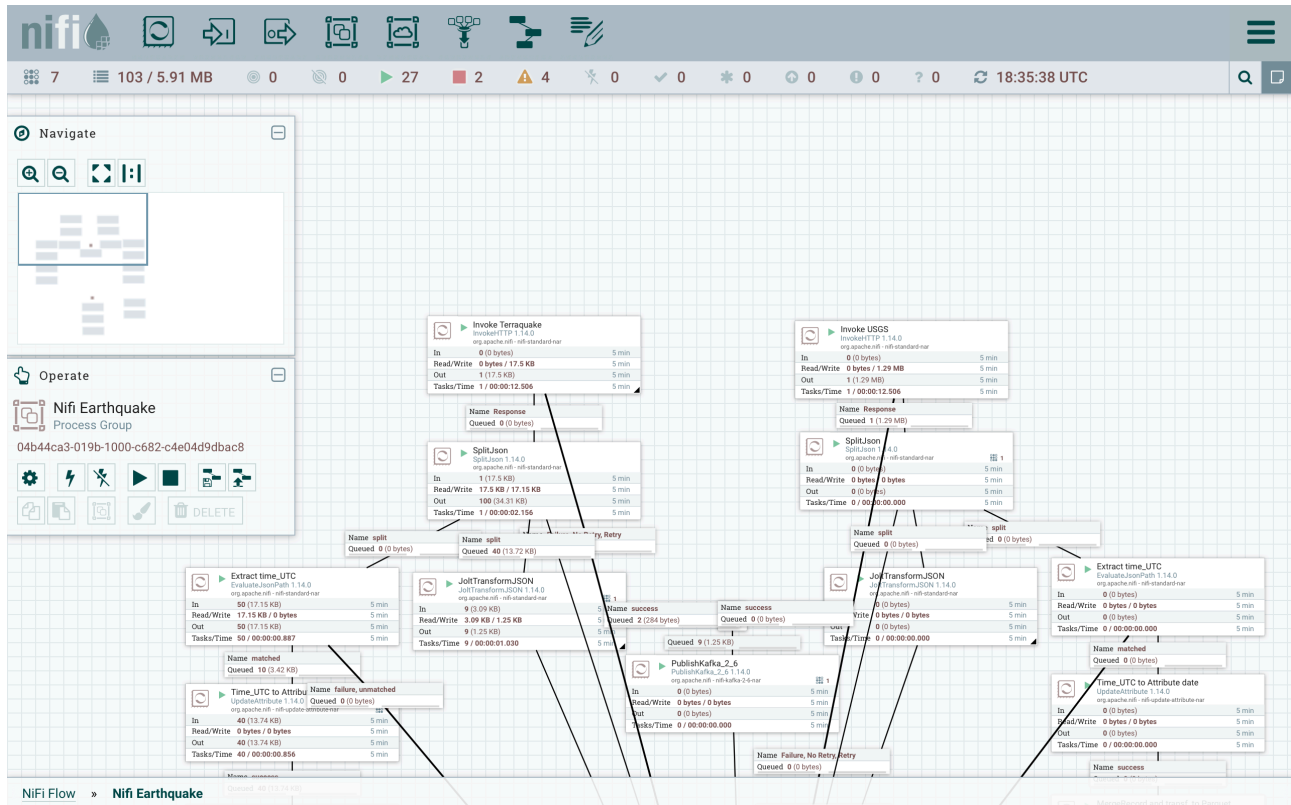# Earthquake MST2 – Testing & Evidence Document
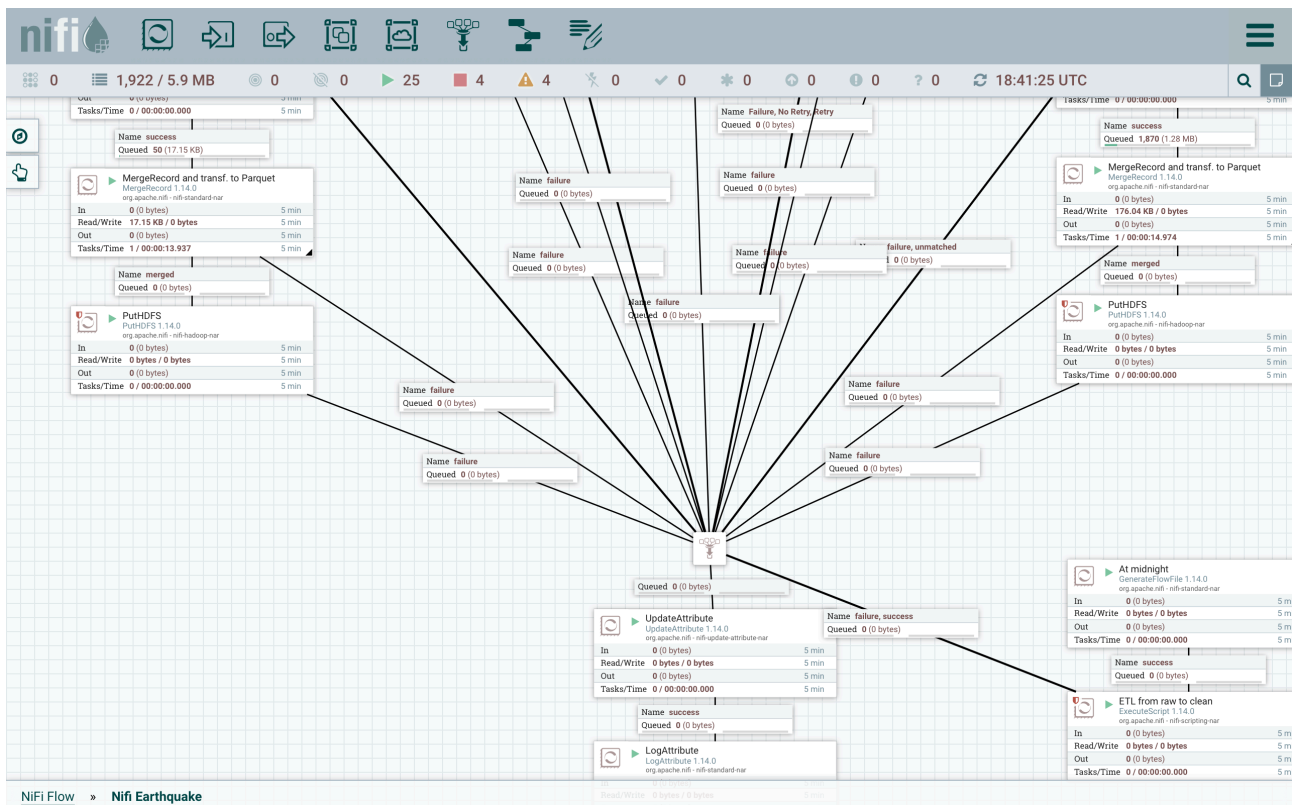
1. NiFi Flow Execution

Objective: Verify that the NiFi flow executes successfully without errors.
Expected Result:
All processors are running
No errors observed

Result: PASS

2. Raw Data Ingestion

HDFS Path:
/user/testuser/earthquake_data/terraquake_raw
/user/testuser/earthquake_data/USGS_raw

Objective: Validate that raw data is successfully ingested into HDFS.
Validation Checks:
Files exist in HDFS
Evidence (Screenshot):

```
[vagrant@node1:~$ hdfs dfs -ls /user/testuser/earthquake_data/USGS_raw
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.7.6/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/apache-tez-0.9.1-bin/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Found 4 items
drwxr-xr-x   - root supergroup          0 2025-12-17 18:24 /user/testuser/earthquake_data/USGS_raw/dt=2025-12-16
drwxr-xr-x   - root supergroup          0 2025-12-17 18:24 /user/testuser/earthquake_data/USGS_raw/dt=2025-12-17
drwxr-xr-x   - root supergroup          0 2025-12-22 18:45 /user/testuser/earthquake_data/USGS_raw/dt=2025-12-21
drwxr-xr-x   - root supergroup          0 2025-12-22 18:45 /user/testuser/earthquake_data/USGS_raw/dt=2025-12-22
[vagrant@node1:~$ hdfs dfs -ls /user/testuser/earthquake_data/terraquake_raw
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.7.6/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/apache-tez-0.9.1-bin/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Found 2 items
drwxr-xr-x   - root supergroup          0 2025-12-17 18:18 /user/testuser/earthquake_data/terraquake_raw/dt=2025-12-17
drwxr-xr-x   - root supergroup          0 2025-12-22 18:45 /user/testuser/earthquake_data/terraquake_raw/dt=2025-12-22
```

Result: PASS

3. Kafka Event Registration

Objective: Verify that earthquake events are successfully registered in Kafka.
Validation Checks:
Kafka topic exists

Events are being produced/consumed

```
vagrant@node1:~$ /usr/local/kafka/bin/kafka-console-consumer.sh --topic unified_quakes --bootstrap-server localhost:9092
{"id":44835232,"magnitude":1.3,"mag_type":"ML","place":"8 km NE Assisi (PG)","time_utc":"2025-12-22T13:47:49.590000","source":"terraquake"}
{"id":44835162,"magnitude":1.6,"mag_type":"ML","place":"Costa Siciliana centro-settentrionale (Palermo)","time_utc":"2025-12-22T13:38:03.530000","source":"terraquake"}
{"id":44834542,"magnitude":0.9,"mag_type":"ML","place":"1 km W Campotosto (AQ)","time_utc":"2025-12-22T11:53:12.290000","source":"terraquake"}
{"id":44834102,"magnitude":6.4,"mag_type":"Mwp","place":"Papua New Guinea [Land]","time_utc":"2025-12-22T10:31:28.277000","source":"terraquake"}
{"id":44834002,"magnitude":2.1,"mag_type":"ML","place":"3 km E Mangone (CS)","time_utc":"2025-12-22T10:23:20.800000","source":"terraquake"}
{"id":44833672,"magnitude":1.1,"mag_type":"ML","place":"2 km NE Cascia (PG)","time_utc":"2025-12-22T09:21:01.750000","source":"terraquake"}
{"id":44833612,"magnitude":0.5,"mag_type":"ML","place":"8 km N Locana (TO)","time_utc":"2025-12-22T08:57:27.230000","source":"terraquake"}
{"id":44833132,"magnitude":1.3,"mag_type":"ML","place":"2 km S Ussita (MC)","time_utc":"2025-12-22T07:04:07.470000","source":"terraquake"}
{"id":44832862,"magnitude":2.2,"mag_type":"ML","place":"Golfo di Salerno (Salerno)","time_utc":"2025-12-22T04:37:54.680000","source":"terraquake"}
{"id":44832832,"magnitude":1.3,"mag_type":"ML","place":"6 km SE Cittareale (RI)","time_utc":"2025-12-22T03:25:06.520000","source":"terraquake"}
{"id":44832782,"magnitude":1.2,"mag_type":"ML","place":"2 km NW Norcia (PG)","time_utc":"2025-12-22T02:45:57.760000","source":"terraquake"}
{"id":44832722,"magnitude":1.2,"mag_type":"ML","place":"4 km E Cellara (CS)","time_utc":"2025-12-22T02:07:04.400000","source":"terraquake"}
{"id":44832682,"magnitude":0.8,"mag_type":"ML","place":"4 km E Mangone (CS)","time_utc":"2025-12-22T01:21:09.270000","source":"terraquake"}
{"id":44832662,"magnitude":0.4,"mag_type":"ML","place":"5 km NW Pietralunga (PG)","time_utc":"2025-12-22T01:05:09.400000","source":"terraquake"}
{"id":44832612,"magnitude":1.2,"mag_type":"ML","place":"4 km E Mangone (CS)","time_utc":"2025-12-22T00:51:41.120000","source":"terraquake"}
{"id":44832572,"magnitude":1.5,"mag_type":"ML","place":"2 km SW Mottafollone (CS)","time_utc":"2025-12-22T00:40:56.690000","source":"terraquake"}
{"id":44832542,"magnitude":0.9,"mag_type":"ML","place":"5 km NE Cittareale (RI)","time_utc":"2025-12-22T00:40:26.610000","source":"terraquake"}
{"id":44832502,"magnitude":1.2,"mag_type":"ML","place":"6 km E Gubbio (PG)","time_utc":"2025-12-21T23:23:57.410000","source":"terraquake"}
{"id":44832482,"magnitude":1,"mag_type":"ML","place":"4 km E Cellara (CS)","time_utc":"2025-12-21T23:22:31.520000","source":"terraquake"}
{"id":44832432,"magnitude":1.4,"mag_type":"ML","place":"7 km SE Bronte (CT)","time_utc":"2025-12-21T23:07:00.290000","source":"terraquake"}
{"id":44832382,"magnitude":1.2,"mag_type":"ML","place":"7 km SE Bronte (CT)","time_utc":"2025-12-21T22:26:01.930000","source":"terraquake"}
{"id":44832312,"magnitude":1.2,"mag_type":"ML","place":"4 km SE Pietralunga (PG)","time_utc":"2025-12-21T22:19:06.430000","source":"terraquake"}
{"id":44832282,"magnitude":1,"mag_type":"ML","place":"3 km SE Campello sul Clitunno (PG)","time_utc":"2025-12-21T22:08:30.140000","source":"terraquake"}
```

Result: PASS

4. Hive Raw Tables Validation

Objective: Validate Hive tables created on top of HDFS data.
Validation Checks:
Table column follow correct naming
Hive returns example data

```
hive> describe terraquake_raw
    . . > ;
OK
type                   string
properties             struct<eventId:bigint,originId:bigint,time:string,author:string,magType:string,mag:double,magAuthor:string,type:string,place:string,version:int,geojson_creationTime:string>
geometry               struct<type:string,coordinates:array<double>>
dt                     string

# Partition Information
# col_name              data_type           comment

dt                     string
9 rows selected (0.582 seconds)
hive> describe usgs_raw;
OK
id                     string
properties             struct<mag:double,place:string,time:bigint,title:string>
geometry               struct<type:string,coordinates:array<double>>
dt                     string

# Partition Information
# col_name              data_type           comment

dt                     string
9 rows selected (0.696 seconds)
hive>
```

```
hive> SELECT * FROM terraquake_raw LIMIT 1;
OK
Feature {"eventId":44815002,"originId":141835741,"time":"2025-12-17T17:25:39.110000","author":"SURVEY-INGV","magType":"ML","mag":0.9,"magAuthor":"--","type":"earthquake","place":"4 km E Cagli (PU)","version":100,"geojson_creationTime":"2025-12-17T18:08:38"} {"type":"Point","coordinates":[12.6988,43.5598,18.0]} 2025-12-17
1 row selected (6.406 seconds)
hive> SELECT * FROM usgs_raw LIMIT 1;
OK
uu80005863 {"mag":0.81,"place":"13 km N of Summit Park, Utah","time":1765929135360,"title":"M 0.8 - 13 km N of Summit Park, Utah"} {"type":"Point","coordinates":[-111.582333333333,40.8688333333333,11.39]} 2025-12-16
1 row selected (1.629 seconds)
hive>
```

Result: PASS

5. Clean Data Availability

Objective: Verify that cleaned data is generated after the cleaning job.
Validation Checks:
Clean data exists in HDFS

```
vagrant@node1:~$ hdfs dfs -ls /user/testuser/earthquake_data/clean
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.7.6/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/apache-tez-0.9.1-bin/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Found 4 items
drwxr-xr-x   - vagrant supergroup          0 2025-12-22 20:00 /user/testuser/earthquake_data/clean/dt=2025-12-16
drwxr-xr-x   - vagrant supergroup          0 2025-12-22 20:00 /user/testuser/earthquake_data/clean/dt=2025-12-17
drwxr-xr-x   - vagrant supergroup          0 2025-12-22 20:00 /user/testuser/earthquake_data/clean/dt=2025-12-21
drwxr-xr-x   - vagrant supergroup          0 2025-12-22 20:00 /user/testuser/earthquake_data/clean/dt=2025-12-22
```

Result: PASS

6. Hive Clean Table Validation

Objective: Validate Hive tables created on top of HDFS data.

Validation Checks:

Table column follow correct naming

Queries return data

```
hive> DESCRIBE earthquake_clean;
OK
id string
magnitude double
mag_type string
place string
time_utc bigint
longitude double
latitude double
depth_km double
source_system string
dt string

# Partition Information
# col_name              data_type              comment

dt string
15 rows selected (0.449 seconds)
hive> SELECT * FROM earthquake_clean LIMIT 1;
OK
nc75280916 1.49 ml 9 km WNW of Templeton, CA 1765912519940 -120.800666809082 35.5833320617676 5.19999980926514 USGS 2025-12-16
1 row selected (2.381 seconds)
hive> |
```

Result: PASS