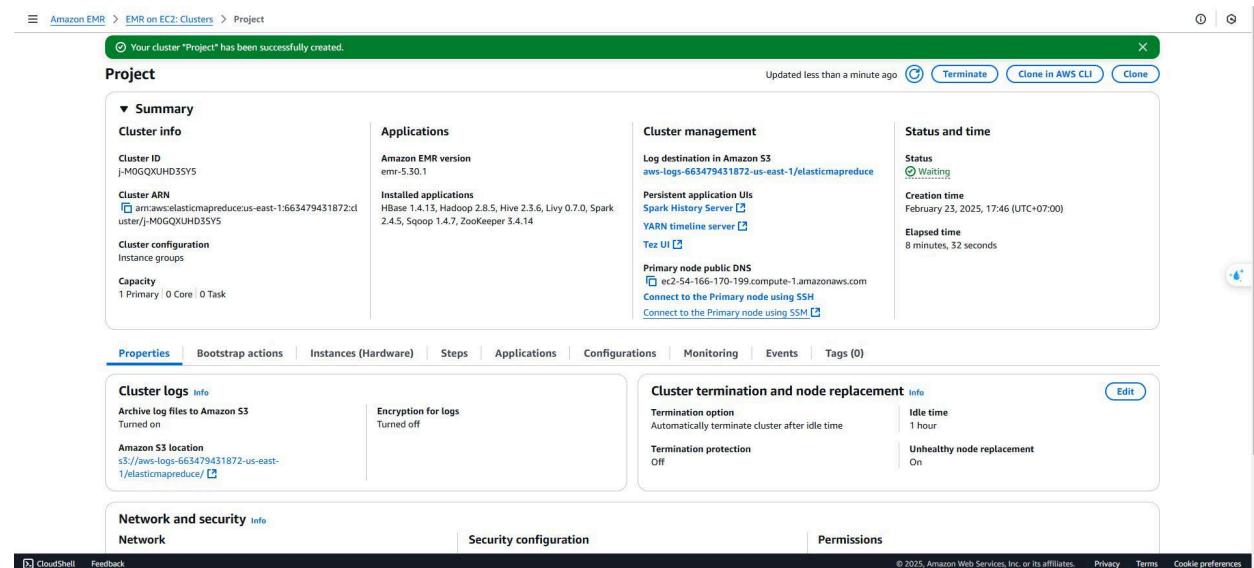
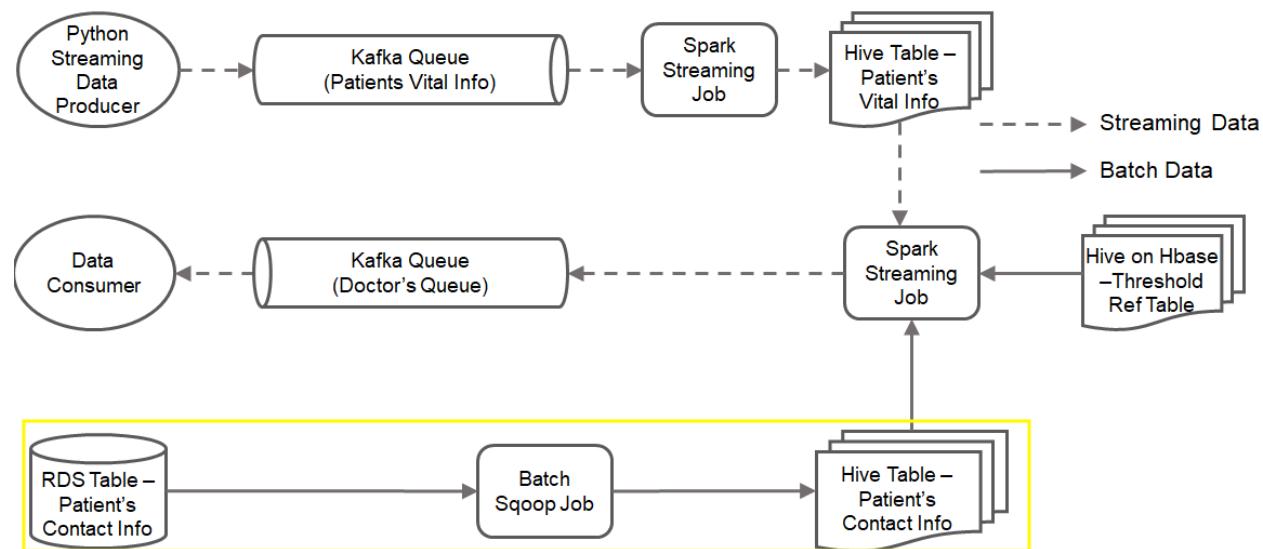


EMR config:



PART 1: We will import the patient contact information from the RDS using Sqoop and load it into a Hive table



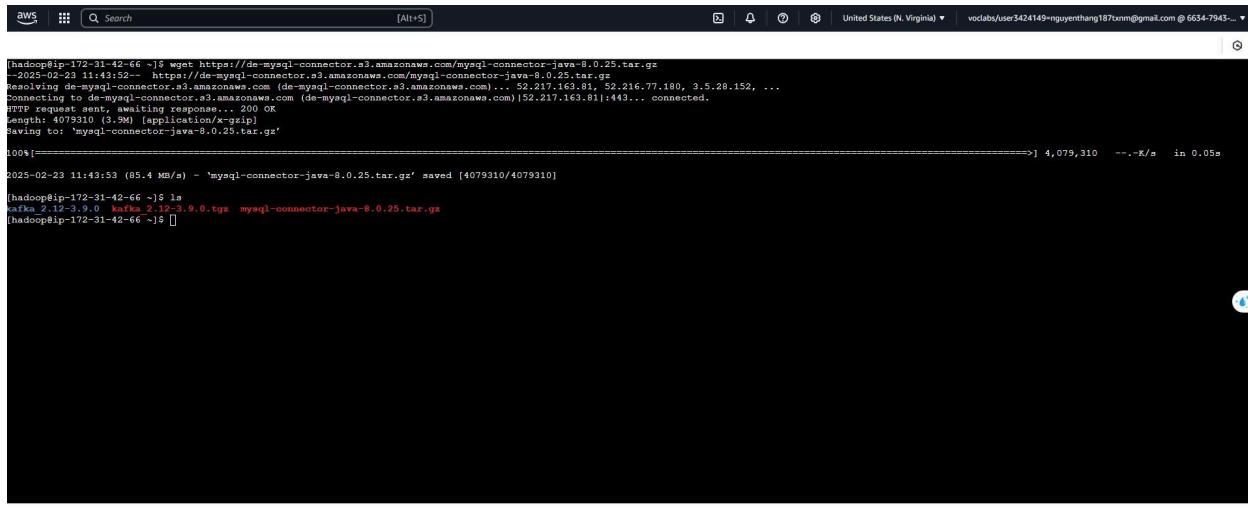
1. Import Patient Contact to HDFS using Sqoop

2. Create Patient Contact table to store it

Sqoop Setup

Script to install mySQL Connector jar file:

`wget https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz`



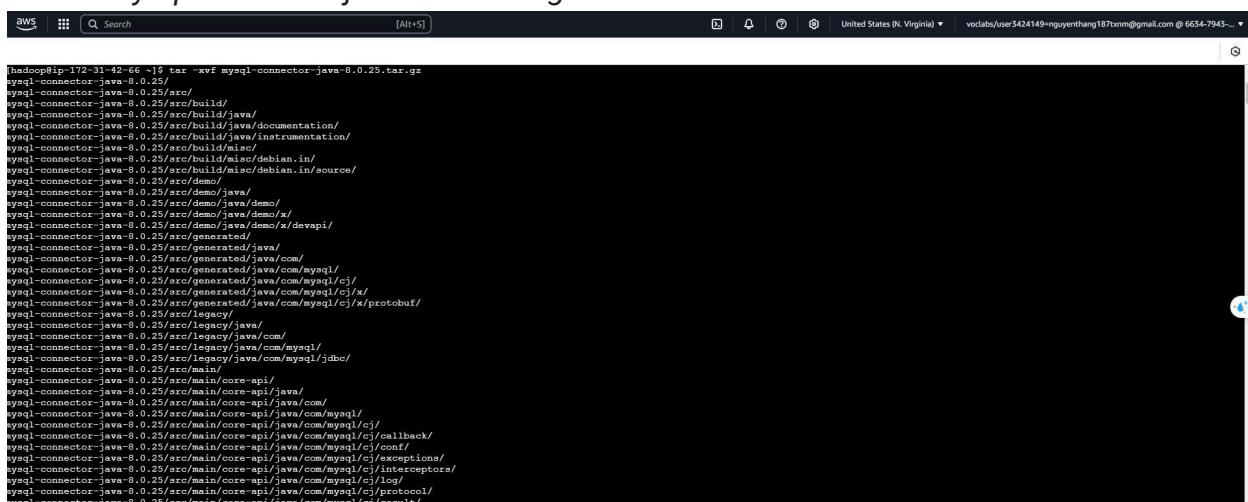
```
[aws ~ ⚡ Search [Alt+S] United States (N. Virginia) v ocabs/user5424149=nguyenthang187xnm@gmail.com @ 6634-7945... ▾
[hadoop@ip-172-31-42-66 ~]$ wget https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz
--2025-02-23 11:43:52 --  https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz
Resolving de-mysql-connector.s3.amazonaws.com (de-mysql-connector.s3.amazonaws.com)... 52.217.163.81, 52.216.77.180, 3.5.28.152, ...
Connecting to de-mysql-connector.s3.amazonaws.com (de-mysql-connector.s3.amazonaws.com)|52.217.163.81|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 4079310 (3.9M) [application/x-gzip]
Saving to: 'mysql-connector-java-8.0.25.tar.gz'

100%[==================================================>] 9,079,310  --.-K/s   in 0.05s
[hadoop@ip-172-31-42-66 ~]$ ls
catkin_2_12-3.9.0_kafka_2.12-3.9.0.tgz  mysql-connector-java-8.0.25.tar.gz
[hadoop@ip-172-31-42-66 ~]$ ]
```

i-0857d5ce17830a01f
 PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

Script to extract [mysql-connector-java-8.0.25.tar.gz](#)

`tar -xvf mysql-connector-java-8.0.25.tar.gz`

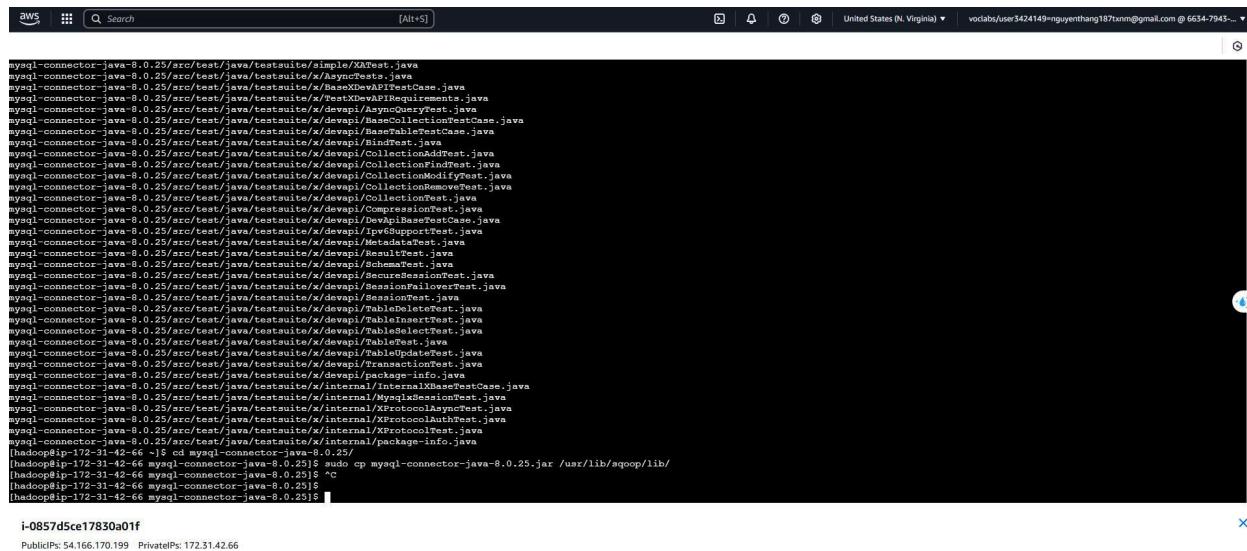


```
[aws ~ ⚡ Search [Alt+S] United States (N. Virginia) v ocabs/user5424149=nguyenthang187xnm@gmail.com @ 6634-7945... ▾
[hadoop@ip-172-31-42-66 ~]$ tar -xvf mysql-connector-java-8.0.25.tar.gz
mysql-connector-java-8.0.25
mysql-connector-java-8.0.25/src/
mysql-connector-java-8.0.25/src/build/
mysql-connector-java-8.0.25/src/build/java/
mysql-connector-java-8.0.25/src/build/java/documentation/
mysql-connector-java-8.0.25/src/build/java/instrumentation/
mysql-connector-java-8.0.25/src/build/java/jdbc/
mysql-connector-java-8.0.25/src/build/macos/debian.in/
mysql-connector-java-8.0.25/src/build/macos/debian.in/source/
mysql-connector-java-8.0.25/src/demo/
mysql-connector-java-8.0.25/src/demo/java/
mysql-connector-java-8.0.25/src/demo/java/demo/
mysql-connector-java-8.0.25/src/demo/java/demo/x/
mysql-connector-java-8.0.25/src/demo/java/demo/x/devapi/
mysql-connector-java-8.0.25/src/generated/
mysql-connector-java-8.0.25/src/generated/java/
mysql-connector-java-8.0.25/src/generated/java/com/
mysql-connector-java-8.0.25/src/generated/java/com/mysql/
mysql-connector-java-8.0.25/src/generated/java/com/mysql/c/
mysql-connector-java-8.0.25/src/generated/java/com/mysql/c/x/
mysql-connector-java-8.0.25/src/generated/java/com/mysql/c/x/protobuf/
mysql-connector-java-8.0.25/src/legacy/
mysql-connector-java-8.0.25/src/legacy/java/
mysql-connector-java-8.0.25/src/legacy/java/com/
mysql-connector-java-8.0.25/src/main/
mysql-connector-java-8.0.25/src/main/core-api/
mysql-connector-java-8.0.25/src/main/core-api/java/
mysql-connector-java-8.0.25/src/main/core-api/java/com/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/callback/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/conf/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/exceptions/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/interceptors/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/log/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/protocol/
mysql-connector-java-8.0.25/src/main/core-api/java/com/mysql/cj/result/]
```

i-0857d5ce17830a01f
 PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

Script to copy to sqoop:

```
cd mysql-connector-java-8.0.25/
sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/
```



```
aws CloudShell Search [Alt+S] United States (N. Virginia) vocabs/user342419=nguyenthang187@nmn@gmail.com @ 6634-7945... ▾

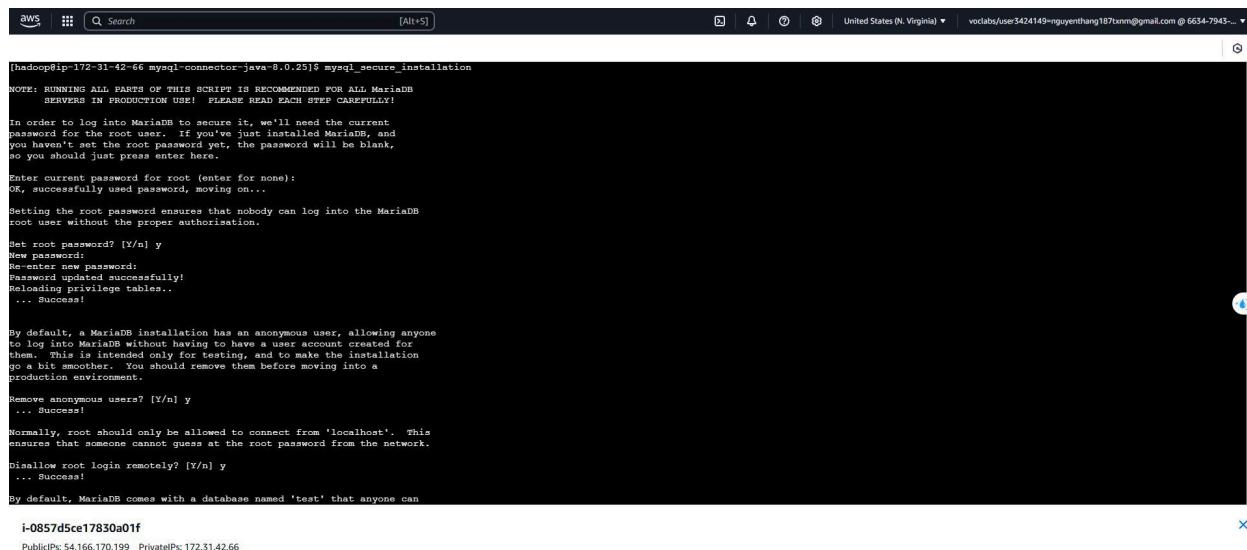
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/XATest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/asyncTests.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/collectionTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/testXDevProtocolTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/AsyncQueryTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/BaseCollectionTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/BaseSessionTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/BindTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionAddTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionFindTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionModifyTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionRemoveTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CompressionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/DeVTableBaseTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/Ipv6SupportTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/LoadBalanceTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/LogInTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SchemaTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SecureSessionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SessionOverTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableBatchTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableInsertTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableSelectTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableUpdateTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TransactionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/package-info.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/InternalXBaseTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolAnywhereTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolAsynchTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolAuthTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/package-info.java
[hadoop@ip-172-31-42-66 ~]$ sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/
[hadoop@ip-172-31-42-66 ~]$ ^C
[hadoop@ip-172-31-42-66 ~]$ mysql-connector-java-8.0.25
[hadoop@ip-172-31-42-66 ~]$ ^C
[hadoop@ip-172-31-42-66 ~]$ mysql-connector-java-8.0.25
[hadoop@ip-172-31-42-66 ~]$ ^C
[hadoop@ip-172-31-42-66 ~]$ ^C
```

i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to install mySQL: *mysql_secure_installation*



```
[hadoop@ip-172-31-42-66 ~]$ mysql-connector-java-8.0.25$ mysql_secure_installation
NOTE: RUNNING ALL PARTS OF THIS SCRIPT IS RECOMMENDED FOR ALL MariaDB
SERVERS IN PRODUCTION USE! PLEASE READ EACH STEP CAREFULLY!

In order to log into MariaDB to secure it, we'll need the current
password for the root user. If you've just installed MariaDB, and
you haven't set the root password yet, the password will be blank,
so you should just press enter here.

Enter current password for root (enter for none):
OK, successfully used password, moving on...

Setting the root password ensures that nobody can log into the MariaDB
root user without the proper authorisation.

Set root password? [Y/n] y
New password:
Re-enter new password:
Password updated successfully!
Reloading privilege tables...
... Success!

By default, a MariaDB installation has an anonymous user, allowing anyone
to log into MariaDB without having to have a user account created for
them. This is intended only for testing, and to make the installation
go a bit smoother. You should remove them before moving into a
production environment.

Remove anonymous users? [Y/n] y
... success!

Normally, root should only be allowed to connect from 'localhost'. This
ensures that someone cannot guess at the root password from the network.

Disallow root login remotely? [Y/n] y
... Success!

By default, MariaDB comes with a database named 'test' that anyone can
```

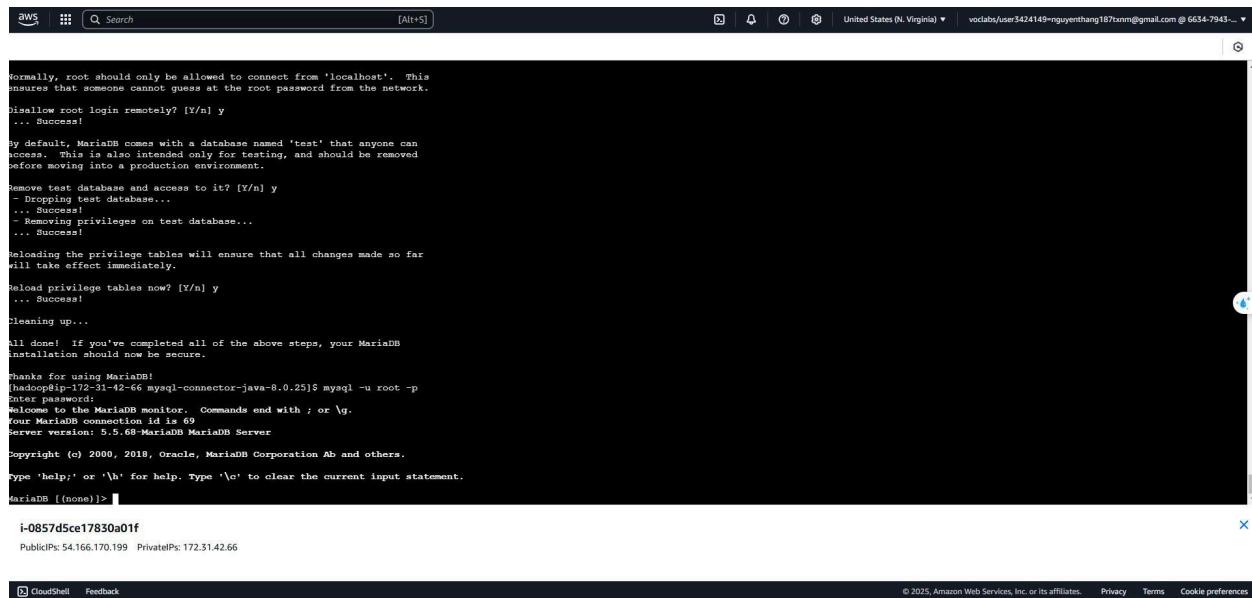
i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to login mysql with root:

```
mysql -u root -p
```



```
aws | [Alt+S] | United States (N. Virginia) | vclabs/user3424149-nnguyenthang187xnm@gmail.com @ 6634-7943... | Search | [Alt+S]

Normally, root should only be allowed to connect from 'localhost'. This
ensures that someone cannot guess at the root password from the network.

Disallow root login remotely? [Y/n] y
... Success!

By default, MariaDB comes with a database named 'test' that anyone can
access. This is also intended only for testing, and should be removed
before moving into a production environment.

Remove test database and access to it? [Y/n] y
- Dropping test database...
... Success!
- Removing privileges on test database...
... Success!

Reloading the privilege tables will ensure that all changes made so far
will take effect immediately.

Reload privilege tables now? [Y/n] y
... Success!

Cleaning up...

All done! If you've completed all of the above steps, your MariaDB
installation should now be secure.

Thanks for using MariaDB!
[hadoopip-172-31-42-66 mysql-connector-java-8.0.25]$ mysql -u root -p
Enter password:
Welcome to the MariaDB monitor. Commands end with ; or \g.
Your MariaDB connection id is 69
Server version: 5.5.68-MariaDB MariaDB Server

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> 
```

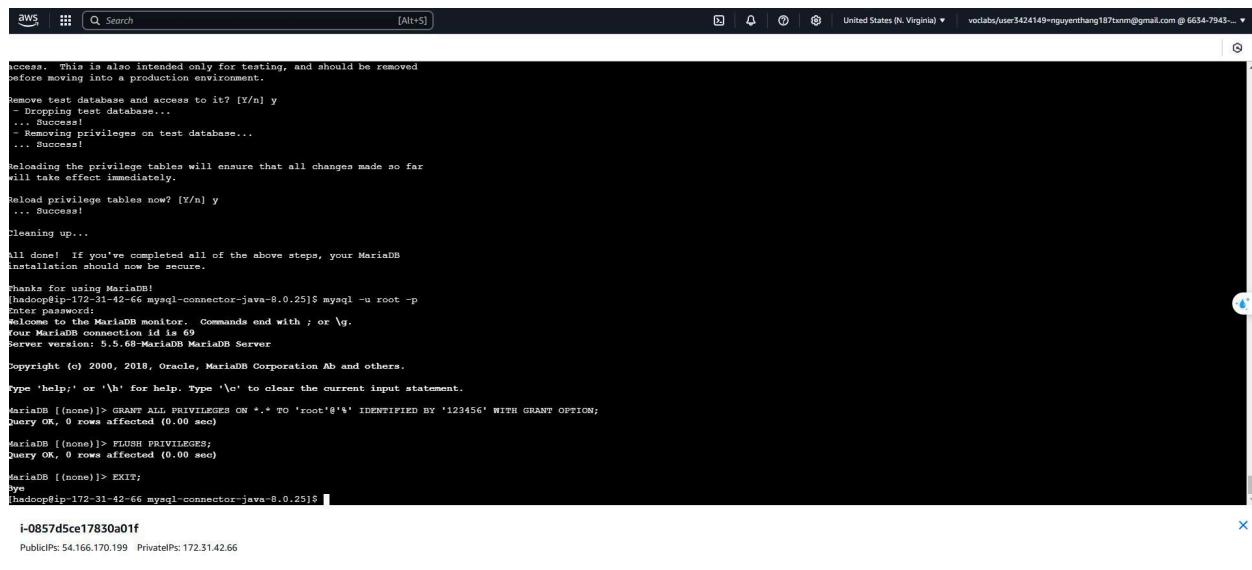
i-0857d5ce17830a01f

PublicIP: 54.166.170.199 PrivateIP: 172.31.42.66

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to grant full permission to root:

```
GRANT ALL PRIVILEGES ON *.* TO 'root'@'%' IDENTIFIED BY '123456' WITH GRANT
OPTION;
FLUSH PRIVILEGES;
EXIT;
```



```
aws | [Alt+S] | United States (N. Virginia) | vclabs/user3424149-nnguyenthang187xnm@gmail.com @ 6634-7943... | Search | [Alt+S]

Access. This is also intended only for testing, and should be removed
before moving into a production environment.

Remove test database and access to it? [Y/n] y
- Dropping test database...
... Success!
- Removing privileges on test database...
... Success!

Reloading the privilege tables will ensure that all changes made so far
will take effect immediately.

Reload privilege tables now? [Y/n] y
... Success!

Cleaning up...

All done! If you've completed all of the above steps, your MariaDB
installation should now be secure.

Thanks for using MariaDB!
[hadoopip-172-31-42-66 mysql-connector-java-8.0.25]$ mysql -u root -p
Enter password:
Welcome to the MariaDB monitor. Commands end with ; or \g.
Your MariaDB connection id is 69
Server version: 5.5.68-MariaDB MariaDB Server

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> GRANT ALL PRIVILEGES ON *.* TO 'root'@'%' IDENTIFIED BY '123456' WITH GRANT OPTION;
Query OK, 0 rows affected (0.00 sec)

MariaDB [(none)]> FLUSH PRIVILEGES;
Query OK, 0 rows affected (0.00 sec)

MariaDB [(none)]> EXIT;
Bye
[hadoopip-172-31-42-66 mysql-connector-java-8.0.25]$ 
```

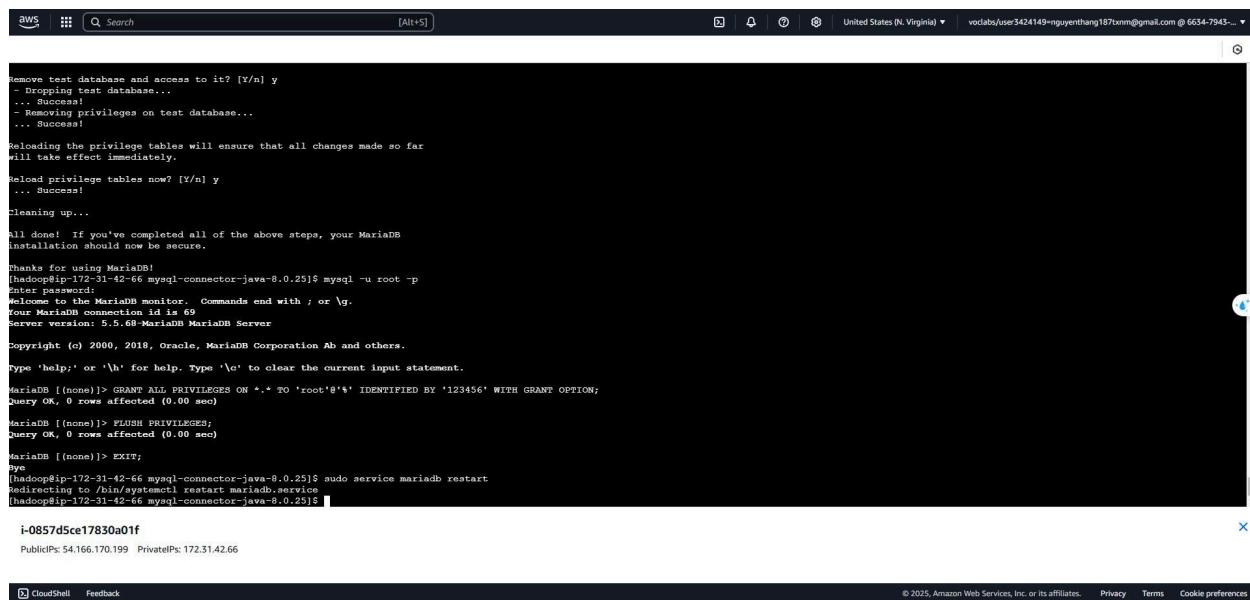
i-0857d5ce17830a01f

PublicIP: 54.166.170.199 PrivateIP: 172.31.42.66

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to restart mySQL server:

sudo service mariadb restart



```

aws [Alt+S] Search United States (N. Virginia) vclabs/user3424149=nguyenthang187xnm@gmail.com @ 6634-7943... ▾

Remove test database and access to it? [Y/n] y
- Dropping test database...
... Success!
- Removing privileges on test database...
... Success!

Reloading the privilege tables will ensure that all changes made so far
will take effect immediately.

Reload privilege tables now? [Y/n] y
... Success!

Cleaning up...

All done! If you've completed all of the above steps, your MariaDB
installation should now be secure.

Thanks for using MariaDB!
[hadoop@ip-172-31-42-66 mysql-connector-java-8.0.25]$ mysql -u root -p
Enter password:
Welcome to the MariaDB monitor. Commands end with ; or \g.
Your MariaDB connection id is 69
Server version: 5.5.68-MariaDB MariaDB Server

Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> GRANT ALL PRIVILEGES ON *.* TO 'root'@'%' IDENTIFIED BY '123456' WITH GRANT OPTION;
Query OK, 0 rows affected (0.00 sec)

MariaDB [(none)]> FLUSH PRIVILEGES;
Query OK, 0 rows affected (0.00 sec)

MariaDB [(none)]> EXIT;
Bye
[hadoop@ip-172-31-42-66 mysql-connector-java-8.0.25]$ sudo service mariadb restart
Redirecting to /bin/systemctl restart mariadb.service
[hadoop@ip-172-31-42-66 mysql-connector-java-8.0.25]$ 

```

i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Data Ingestion:

Script to import data to HDFS:

sqoop import --connect jdbc:mysql://upgraddetest.cyaielc9bmnf.us-east-1.rds.amazonaws.com/testdatabase --table patients_information --username student --password STUDENT123 --target-dir /user/livy/patient_contact_info -m 1

aws CloudShell Search [Alt+S]

```
[hadoop@ip-172-31-42-66 ~]$ sqoop import --connect jdbc:mysql://upgradtest.cyaiecl9bmnf.us-east-1.rds.amazonaws.com/testdatabase --table patients_information --username student --password STUDENT123 --target-dir /user/livy/patient_contact_info -m 1
Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please add $ACUMULO_HOME to the root of your Accumulo installation.
15/02/23 12:03:39 INFO Configuration: Sqoop version: 1.4.7
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
25/02/23 12:03:40 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
25/02/23 12:03:40 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
25/02/23 12:03:40 INFO tool.CodeGenTool: Beginning code generation
Loading class 'com.mysql.jdbc.Driver'. This is deprecated. The new driver class is 'com.mysql.cj.jdbc.Driver'. The driver is automatically registered via the SPI and manual loading of the driver class is generally unnecessary.
25/02/23 12:03:41 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
25/02/23 12:03:41 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
Note: /tmp/sqoop-hadoop/compile/799a4cba5d252e90e5a1c4d6bd406fb7/patients_information.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
25/02/23 12:03:43 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-hadoop/compile/799a4cba5d252e90e5a1c4d6bd406fb7/patients_information.jar
25/02/23 12:03:43 INFO manager.MySQLManager: It looks like you are importing from mysql.
25/02/23 12:03:43 INFO manager.MySQLManager: This is convenient if you don't want to write a select query.
25/02/23 12:03:43 WARN manager.MySQLManager: option to execute a MySQL-specific fast path.
25/02/23 12:03:43 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
25/02/23 12:03:43 INFO mapreduce.ImportJobBase: Beginning import of patients_information
25/02/23 12:03:43 INFO mapreduce.Job: Configuration parameters: mapred.job_mapper=mapred.map.tasks=1
25/02/23 12:03:44 INFO Configuration: mapred.job.mapper=mapred.map.tasks=1 is deprecated. Instead, use mapreduce.job.maps
25/02/23 12:03:44 INFO client.RMProxy: Connecting to ResourceManager at ip-172-31-42-66.ec2.internal/172.31.42.66:8032
25/02/23 12:03:49 INFO db.DBInputFormat: Using read committed transaction
25/02/23 12:03:49 INFO mapreduce.JobSubmitter: number of splits:1
25/02/23 12:03:49 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1740308050138_0009
25/02/23 12:03:49 INFO Configuration: mapred.job.tracker=http://ip-172-31-42-66.ec2.internal:20888
25/02/23 12:03:49 INFO mapreduce.Job: The url to track the job: http://ip-172-31-42-66.ec2.internal:20888/proxy/application_1740308050138_0009/
25/02/23 12:03:49 INFO mapreduce.Job: Job: running job: job_1740308050138_0009
25/02/23 12:03:56 INFO mapreduce.Job: Job job_1740308050138_0009 running in uber mode : false
25/02/23 12:04:01 INFO mapreduce.Job: map 100% reduce 0%
25/02/23 12:04:01 INFO mapreduce.Job: Job job_1740308050138_0009 completed successfully
```

i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

aws CloudShell Search [Alt+S]

```
25/02/23 12:04:01 INFO mapreduce.Job: Job job_1740308050138_0009 completed successfully
25/02/23 12:04:01 INFO mapreduce.Job: Counters: 30
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=189902
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=0
  HDFS: Number of bytes written=230
  HDFS: Number of read operations=4
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Other local map tasks=1
  Total time spent by all maps in occupied slots (ms)=8577
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=2855
  Total time spent by all reduce tasks (ms)=2855
  Total map memory consumed (bytes)=2959
  Total map bytes+milliseconds taken by all map tasks=8782848
Map-Reduce Framework
  Map input records=5
  Map output records=5
  Map output bytes=5
  Input split bytes=67
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time (seconds)=0.64
  OSGI time (ms)=1520
  Physical memory (bytes) snapshot=256475136
  Virtual memory (bytes) snapshot=4614033408
  Total committed heap usage (bytes)=247463936
  File Input Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=230
25/02/23 12:04:01 INFO mapreduce.ImportJobBase: Transferred 230 bytes in 17.3897 seconds (13.2262 bytes/sec)
25/02/23 12:04:01 INFO mapreduce.ImportJobBase: Retrieved 5 records.
[hadoop@ip-172-31-42-66 ~]$
```

i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to view HDFS file:

`hadoop fs -ls /user/livy/patient_contact_info`

aws | [Alt+S] | Search | United States (N. Virginia) | vclabs/user3424149+nguyenthang187@nm@gmail.com @ 6634-7943... ▾

```

FILE: Number of bytes written=189902
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes written=0
HDFS: Number of bytes written=230
HDFS: Number of read operations=4
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=8577
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=8577
Total worker-milliseconds taken by all map tasks=2059
Total megabyte-milliseconds taken by all map tasks=8782848
Map-Reduce Framework
Map input records=5
Map output records=5
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=64
CPU time spent (ms)=1520
Physical memory (bytes) snapshot=256475136
Virtual memory (bytes) snapshot=4614033408
Total committed heap usage (bytes)=247463936
File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=230
05/02/23 12:04:01 INFO mapreduce.ImportJobBase: Transferred 230 bytes in 17.3897 seconds (13.2262 bytes/sec)
05/02/23 12:04:01 INFO mapreduce.ImportJobBase: Retrieved 5 records
[hadoopcpip-172-31-42-66 :1$ hadoop fs -ls /user/livy/patient_contact_info
Found 2 items
-rw-r--r-- 1 hadoop livy 0 2025-02-23 12:04 /user/livy/patient_contact_info/_SUCCESS
-rw-r--r-- 1 hadoop livy 230 2025-02-23 12:04 /user/livy/patient_contact_info/part-m-00000
[hadoopcpip-172-31-42-66 :1$ ]

```

i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to view content:

```
hadoop fs -cat /user/livy/patient_contact_info/part-m-00000
```

aws | [Alt+S] | Search | United States (N. Virginia) | vclabs/user3424149+nguyenthang187@nm@gmail.com @ 6634-7943... ▾

```

HDFS: Number of read operations=4
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=8577
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=2859
Total worker-milliseconds taken by all map tasks=2859
Total megabyte-milliseconds taken by all map tasks=8782848
Map-Reduce Framework
Map input records=5
Map output records=5
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=64
CPU time spent (ms)=1520
Physical memory (bytes) snapshot=256475136
Virtual memory (bytes) snapshot=4614033408
Total committed heap usage (bytes)=247463936
File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=230
05/02/23 12:04:01 INFO mapreduce.ImportJobBase: transferred 230 bytes in 17.3897 seconds (13.2262 bytes/sec)
05/02/23 12:04:01 INFO mapreduce.ImportJobBase: Retrieved 5 records
[hadoopcpip-172-31-42-66 :1$ hadoop fs -ls /user/livy/patient_contact_info
Found 2 items
-rw-r--r-- 1 hadoop livy 0 2025-02-23 12:04 /user/livy/patient_contact_info/_SUCCESS
-rw-r--r-- 1 hadoop livy 230 2025-02-23 12:04 /user/livy/patient_contact_info/part-m-00000
[hadoopcpip-172-31-42-66 :1$ hadoop fs -cat /user/livy/patient_contact_info/part-m-00000
[{"id": "1", "name": "Alex S. Khan", "age": 25, "city": "New York", "country": "USA", "address": "8982739282, 1, null}, {"id": "2", "name": "Karan C.", "age": 28, "city": "Delhi", "country": "India", "address": "8923739282, 2, 45, null}, {"id": "3", "name": "Dara M.", "age": 30, "city": "Mumbai", "country": "India", "address": "2182739282, 4, 67, null}, {"id": "4", "name": "Pam ABC", "age": 22, "city": "London", "country": "UK", "address": "4982739282, 5, 72, null}
[hadoopcpip-172-31-42-66 :1$ ]

```

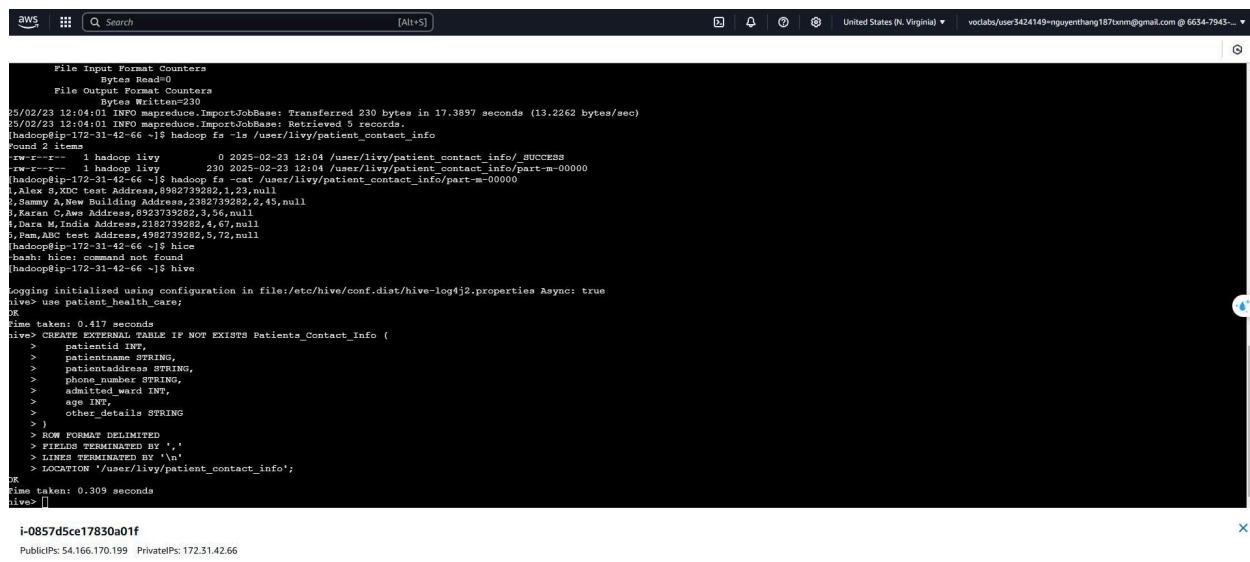
i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to open hive and table Patients_Contact_Info:

```
hive
use patient_health_care;
CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Contact_Info (
    patientid INT,
    patientname STRING,
    patientaddress STRING,
    phone_number STRING,
    admitted_ward INT,
    age INT,
    other_details STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
LOCATION '/user/livy/patient_contact_info';
```



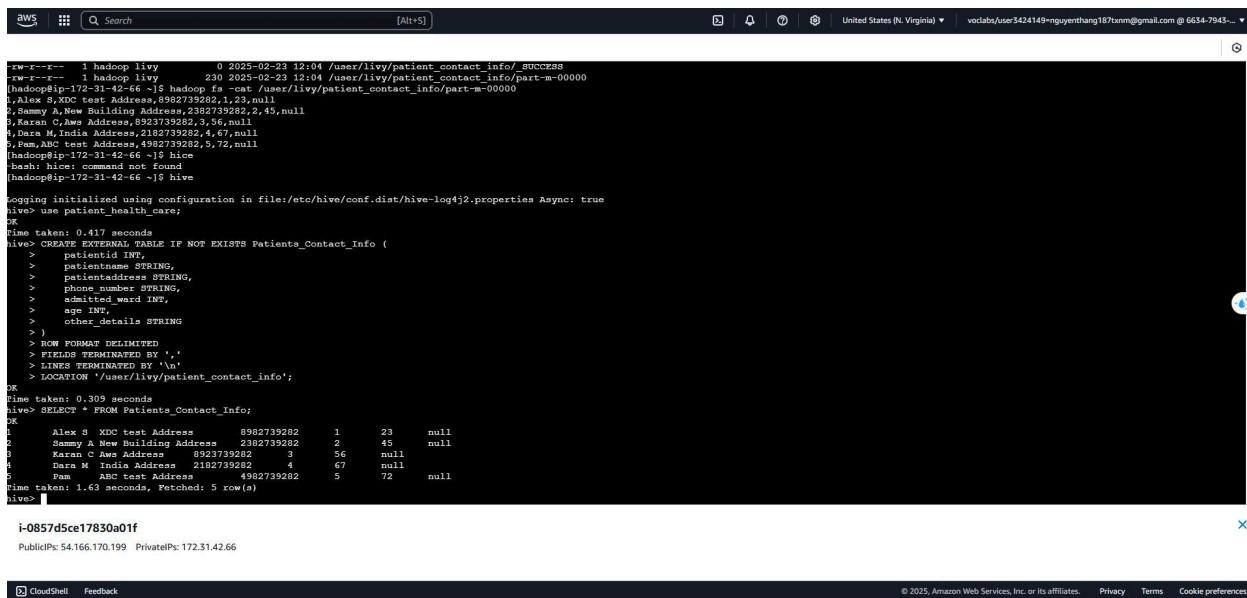
```
File Input Format Counters
  Bytes Read:0
  File Output Format Counters
  Bytes Written:230
09/02/23 12:04:01 INFO mapreduce.ImportJobBase: Transferred 230 bytes in 17.3897 seconds (13.2262 bytes/sec)
09/02/23 12:04:01 INFO mapreduce.ImportJobBase: Retrieved 5 records
[hadoop@ip-172-31-42-66 ~]$ hadoop fs -ls /user/livy/patient_contact_info
Found 2 items
-rw-r--r-- 1 hadoop livy 0 2025-02-23 12:04 /user/livy/patient_contact_info/_SUCCESS
-rw-r--r-- 1 hadoop livy 230 2025-02-23 12:04 /user/livy/patient_contact_info/part-m-00000
[hadoop@ip-172-31-42-66 ~]$ hadoop fs -cat /user/livy/patient_contact_info/part-m-00000
1.Sammy A,New Building Address,2382739282,2,45,null
2,Karen C,Aws Address,8923739282,3,56,null
3,Alex S,XDC test Address,8982739282,1,23,null
4,David M,India Address,2182739282,4,67,null
5,John M,US test Address,8923739282,5,72,null
[hadoop@ip-172-31-42-66 ~]$ hive
[hadoop@ip-172-31-42-66 ~]$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log$2.properties Async: true
hive> use patient_health_care;
OK
Time taken: 0.417 seconds
hive> CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Contact_Info (
    >     patientid INT,
    >     patientname STRING,
    >     patientaddress STRING,
    >     phone_number STRING,
    >     admitted_ward INT,
    >     age INT,
    >     other_details STRING
    > )
    > ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ','
    > LINES TERMINATED BY '\n'
    > LOCATION '/user/livy/patient_contact_info';
OK
Time taken: 0.309 seconds
hive> [REDACTED]
```

i-0857d5ce17830a01f

PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

Script to view data:

```
SELECT * FROM Patients_Contact_Info;
```



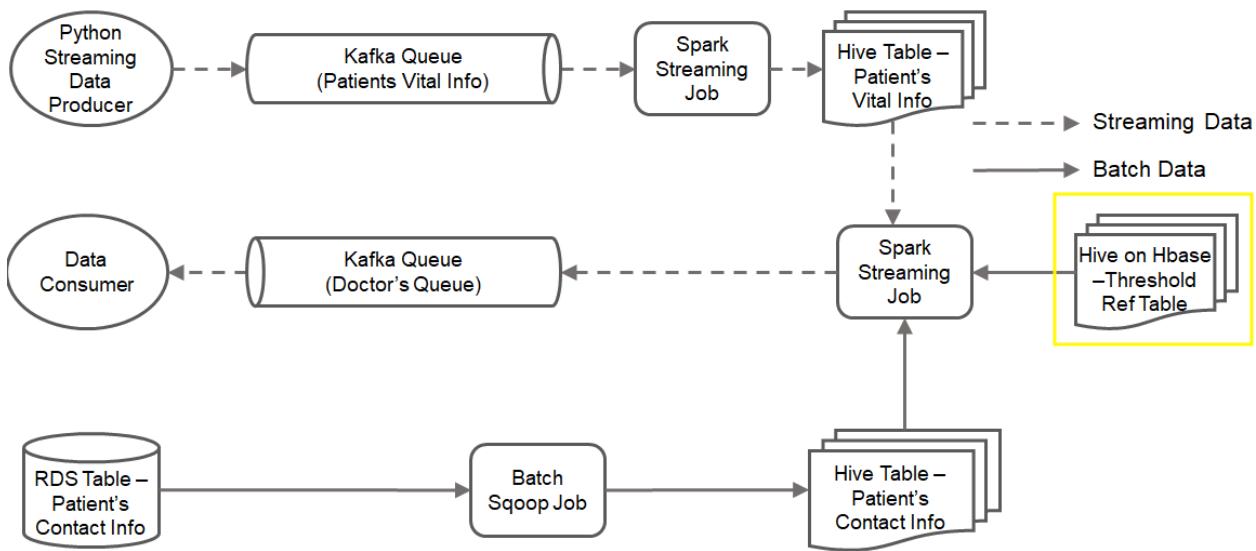
```

aws [Alt+S] Search United States (N. Virginia) vodlabs/user3424149=nguyenthang187xnm@gmail.com @ 6634-7945... ▾

-rw-r--r-- 1 hadoop livy      0 2025-02-23 12:04 /user/livy/patient_contact_info/_SUCCESS
-rw-r--r-- 1 hadoop livy     230 2025-02-23 12:04 /user/livy/patient_contact_info/part-m-00000
[hadoop@ip-172-31-42-66 ~]$ hadoop fs -cat /user/livy/patient_contact_info/part-m-00000
1,Alex S,XDC test Address,8982739282,1,23,null
2,Sammy A New Building Address,2302739282,2,45,null
3,Karan C Aws Address,8923739282,3,56,null
4,Dara M India Address,2102739282,4,67,null
5,Pam ABC test Address,4982739282,5,72,null
[hadoop@ip-172-31-42-66 ~]$ hive
hive> use patient_health_care;
hive> CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Contact_Info (
    >     patientid INT,
    >     patientname STRING,
    >     patientaddress STRING,
    >     phonenumber STRING,
    >     admittedward INT,
    >     age INT,
    >     other_details STRING
    > ) ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ','
    > LINES TERMINATED BY '\n'
    > LOCATION '/user/livy/patient_contact_info';
OK
Time taken: 0.417 seconds
hive> SELECT * FROM Patients_Contact_Info;
OK
1 Alex S XDC test Address 8982739282 1 23 null
2 Sammy A New Building Address 2302739282 2 45 null
3 Karan C Aws Address 8923739282 3 56 null
4 Dara M India Address 2102739282 4 67 null
5 Pam ABC test Address 4982739282 5 72 null
Time taken: 1.63 seconds, Fetched: 5 row(s)
hive> i-0857d5ce17830a01f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

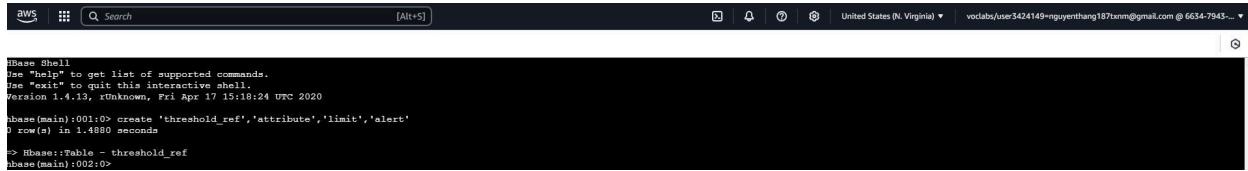
PART 2: Create an HBase table to store threshold reference information and define a Hive external table on top of it



1. Create HBase table *threshold_ref* with 3 columns: attribute, limit, alert
2. Insert 12 records into *threshold_ref*
3. Setup Hive, HBase integration
4. Create Hive table *Threshold_Reference_Table*

Script to create Hbase table:

```
hbase shell
create 'threshold_ref','attribute','limit','alert'
```



A screenshot of a terminal window titled "AWS" with a search bar and an "Alt+S" key binding. The window shows the HBase Shell interface. The output of the command "create 'threshold_ref','attribute','limit','alert'" is displayed, indicating a successful creation of the table "threshold_ref". The terminal also shows the HBase version information and some statistics.

```
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.13, r0Unknown, Fri Apr 17 15:18:24 UTC 2020
hbase(main):001:0> create 'threshold_ref','attribute','limit','alert'
0 row(s) in 1.4880 seconds
-> Hbase::Table - threshold_ref
hbase(main):002:0>
```

Script to insert 12 records into Hbase table:

```
put 'threshold_ref', '1', 'attribute:attribute', 'heartBeat';
put 'threshold_ref', '1', 'limit:low_age_limit', '0';
put 'threshold_ref', '1', 'limit:high_age_limit', '40';
put 'threshold_ref', '1', 'limit:low_value', '50';
put 'threshold_ref', '1', 'limit:high_value', '100';
put 'threshold_ref', '1', 'alert:alert_flag', '1';
put 'threshold_ref', '1', 'alert:alert_message', 'Abnormal Heart Rate Detected';

put 'threshold_ref', '2', 'attribute:attribute', 'heartBeat';
put 'threshold_ref', '2', 'limit:low_age_limit', '0';
put 'threshold_ref', '2', 'limit:high_age_limit', '40';
put 'threshold_ref', '2', 'limit:low_value', '70';
put 'threshold_ref', '2', 'limit:high_value', '78';
put 'threshold_ref', '2', 'alert:alert_flag', '0';
put 'threshold_ref', '2', 'alert:alert_message', 'Normal';

put 'threshold_ref', '3', 'attribute:attribute', 'heartBeat';
put 'threshold_ref', '3', 'limit:low_age_limit', '0';
put 'threshold_ref', '3', 'limit:high_age_limit', '40';
put 'threshold_ref', '3', 'limit:low_value', '79';
put 'threshold_ref', '3', 'limit:high_value', '9999';
put 'threshold_ref', '3', 'alert:alert_flag', '1';
put 'threshold_ref', '3', 'alert:alert_message', 'Higher Heart Rate than Normal';

put 'threshold_ref', '4', 'attribute:attribute', 'bp';
put 'threshold_ref', '4', 'limit:low_age_limit', '0';
put 'threshold_ref', '4', 'limit:high_age_limit', '40';
put 'threshold_ref', '4', 'limit:low_value', '0';
put 'threshold_ref', '4', 'limit:high_value', '160';
put 'threshold_ref', '4', 'alert:alert_flag', '1';
put 'threshold_ref', '4', 'alert:alert_message', 'Low BP than Normal';

put 'threshold_ref', '5', 'attribute:attribute', 'bp';
put 'threshold_ref', '5', 'limit:low_age_limit', '0';
put 'threshold_ref', '5', 'limit:high_age_limit', '40';
put 'threshold_ref', '5', 'limit:low_value', '161';
put 'threshold_ref', '5', 'limit:high_value', '222';
put 'threshold_ref', '5', 'alert:alert_flag', '0';
put 'threshold_ref', '5', 'alert:alert_message', 'Normal';

put 'threshold_ref', '6', 'attribute:attribute', 'bp';
put 'threshold_ref', '6', 'limit:low_age_limit', '0';
put 'threshold_ref', '6', 'limit:high_age_limit', '40';
```

```
put 'threshold_ref', '6', 'limit:low_value', '221';
put 'threshold_ref', '6', 'limit:high_value', '9999';
put 'threshold_ref', '6', 'alert:alert_flag', '1';
put 'threshold_ref', '6', 'alert:alert_message', 'Higher BP than Normal';

put 'threshold_ref', '7', 'attribute:attribute', 'heartBeat';
put 'threshold_ref', '7', 'limit:low_age_limit', '41';
put 'threshold_ref', '7', 'limit:high_age_limit', '100';
put 'threshold_ref', '7', 'limit:low_value', '0';
put 'threshold_ref', '7', 'limit:high_value', '65';
put 'threshold_ref', '7', 'alert:alert_flag', '1';
put 'threshold_ref', '7', 'alert:alert_message', 'Low Heart Rate than Normal';

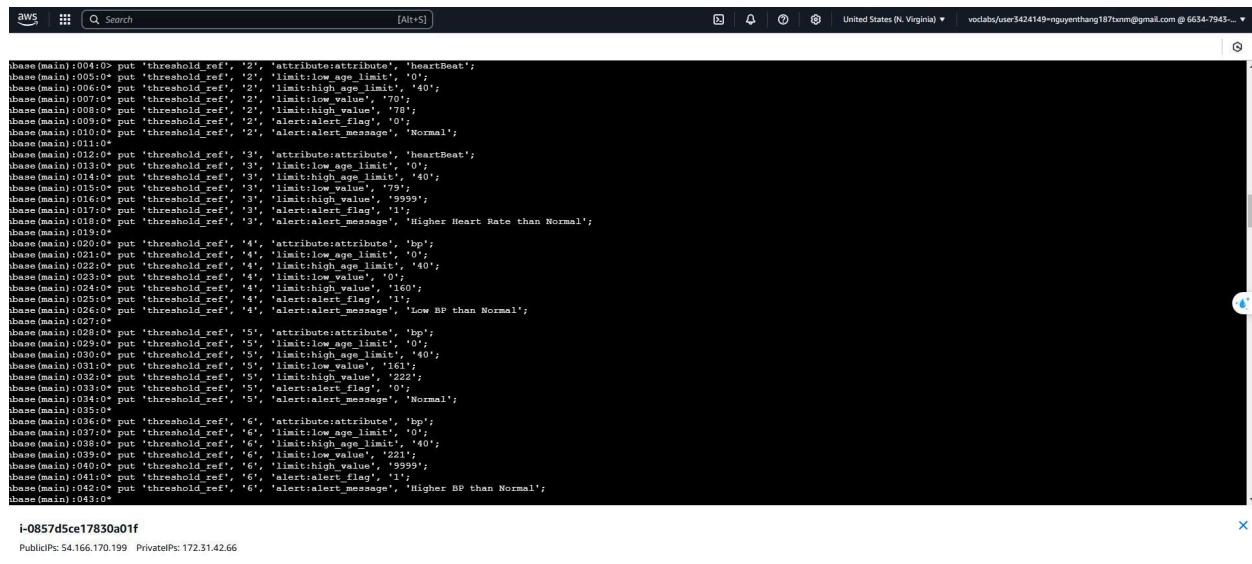
put 'threshold_ref', '8', 'attribute:attribute', 'heartBeat';
put 'threshold_ref', '8', 'limit:low_age_limit', '41';
put 'threshold_ref', '8', 'limit:high_age_limit', '100';
put 'threshold_ref', '8', 'limit:low_value', '66';
put 'threshold_ref', '8', 'limit:high_value', '73';
put 'threshold_ref', '8', 'alert:alert_flag', '0';
put 'threshold_ref', '8', 'alert:alert_message', 'Normal';

put 'threshold_ref', '9', 'attribute:attribute', 'heartBeat';
put 'threshold_ref', '9', 'limit:low_age_limit', '41';
put 'threshold_ref', '9', 'limit:high_age_limit', '100';
put 'threshold_ref', '9', 'limit:low_value', '74';
put 'threshold_ref', '9', 'limit:high_value', '9999';
put 'threshold_ref', '9', 'alert:alert_flag', '1';
put 'threshold_ref', '9', 'alert:alert_message', 'Higher Heart Rate than Normal';

put 'threshold_ref', '10', 'attribute:attribute', 'bp';
put 'threshold_ref', '10', 'limit:low_age_limit', '41';
put 'threshold_ref', '10', 'limit:high_age_limit', '100';
put 'threshold_ref', '10', 'limit:low_value', '0';
put 'threshold_ref', '10', 'limit:high_value', '150';
put 'threshold_ref', '10', 'alert:alert_flag', '1';
put 'threshold_ref', '10', 'alert:alert_message', 'Low BP than Normal';

put 'threshold_ref', '11', 'attribute:attribute', 'bp';
put 'threshold_ref', '11', 'limit:low_age_limit', '41';
put 'threshold_ref', '11', 'limit:high_age_limit', '100';
put 'threshold_ref', '11', 'limit:low_value', '151';
put 'threshold_ref', '11', 'limit:high_value', '180';
put 'threshold_ref', '11', 'alert:alert_flag', '0';
put 'threshold_ref', '11', 'alert:alert_message', 'Normal';
```

```
put 'threshold_ref', '12', 'attribute:attribute', 'bp';
put 'threshold_ref', '12', 'limit:low_age_limit', '41';
put 'threshold_ref', '12', 'limit:high_age_limit', '100';
put 'threshold_ref', '12', 'limit:low_value', '181';
put 'threshold_ref', '12', 'limit:high_value', '9999';
put 'threshold_ref', '12', 'alert:alert_flag', '1';
put 'threshold_ref', '12', 'alert:alert_message', 'Higher BP than Normal';
```



```
aws [Alt+S] Search United States (N. Virginia) vclabs/user3424149-nghuyenthang187xnm@gmail.com @ 6634-794... ▾
base(main):004:0> put 'threshold_ref', '2', 'attribute:attribute', 'heartBeat';
base(main):005:0> put 'threshold_ref', '2', 'limit:low_age_limit', '0';
base(main):006:0> put 'threshold_ref', '2', 'limit:high_age_limit', '40';
base(main):007:0> put 'threshold_ref', '2', 'limit:low_value', '70';
base(main):008:0> put 'threshold_ref', '2', 'limit:high_value', '78';
base(main):009:0> put 'threshold_ref', '2', 'alert:alert_flag', '0';
base(main):010:0> put 'threshold_ref', '2', 'alert:alert_message', 'Normal';
base(main):011:0>
base(main):012:0> put 'threshold_ref', '3', 'attribute:attribute', 'heartBeat';
base(main):013:0> put 'threshold_ref', '3', 'limit:low_age_limit', '0';
base(main):014:0> put 'threshold_ref', '3', 'limit:high_age_limit', '40';
base(main):015:0> put 'threshold_ref', '3', 'limit:low_value', '79';
base(main):016:0> put 'threshold_ref', '3', 'limit:high_value', '9999';
base(main):017:0> put 'threshold_ref', '3', 'alert:alert_flag', '1';
base(main):018:0> put 'threshold_ref', '3', 'alert:alert_message', 'Higher Heart Rate than Normal';
base(main):019:0>
base(main):020:0> put 'threshold_ref', '4', 'attribute:attribute', 'bp';
base(main):021:0> put 'threshold_ref', '4', 'limit:low_age_limit', '0';
base(main):022:0> put 'threshold_ref', '4', 'limit:high_age_limit', '40';
base(main):023:0> put 'threshold_ref', '4', 'limit:low_value', '0';
base(main):024:0> put 'threshold_ref', '4', 'limit:high_value', '160';
base(main):025:0> put 'threshold_ref', '4', 'alert:alert_flag', '1';
base(main):026:0> put 'threshold_ref', '4', 'alert:alert_message', 'Low BP than Normal';
base(main):027:0>
base(main):028:0> put 'threshold_ref', '5', 'attribute:attribute', 'bp';
base(main):029:0> put 'threshold_ref', '5', 'limit:low_age_limit', '0';
base(main):030:0> put 'threshold_ref', '5', 'limit:high_age_limit', '40';
base(main):031:0> put 'threshold_ref', '5', 'limit:low_value', '161';
base(main):032:0> put 'threshold_ref', '5', 'limit:high_value', '222';
base(main):033:0> put 'threshold_ref', '5', 'alert:alert_flag', '0';
base(main):034:0> put 'threshold_ref', '5', 'alert:alert_message', 'Normal';
base(main):035:0>
base(main):036:0> put 'threshold_ref', '6', 'attribute:attribute', 'bp';
base(main):037:0> put 'threshold_ref', '6', 'limit:low_age_limit', '0';
base(main):038:0> put 'threshold_ref', '6', 'limit:high_age_limit', '40';
base(main):039:0> put 'threshold_ref', '6', 'limit:low_value', '221';
base(main):040:0> put 'threshold_ref', '6', 'limit:high_value', '9999';
base(main):041:0> put 'threshold_ref', '6', 'alert:alert_flag', '1';
base(main):042:0> put 'threshold_ref', '6', 'alert:alert_message', 'Higher BP than Normal';
base(main):043:0>
```

i-0857d5ce17830a01f

Public IPs: 54.166.170.199 Private IPs: 172.31.42.66

[CloudShell](#) [Feedback](#) © 2025, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

Script to view data:
scan 'threshold_ref'

aws |   Search [Alt+S] |    United States (N. Virginia)  vclabs/user3424149=nguyenthang187xnm@gmail.com @ 6634-7934... ▾ | 

```
hbase(main):001:0> scan "threshold_ref"
ROW                                     COLUMN+CELL
1                                         column=alert:alert_flag, timestamp=1740309420190, value=1
1                                         column=alert:alert_message, timestamp=1740309420194, value=Low Heart Rate than Normal
1                                         column=attribute:attribute, timestamp=1740309420194, value=heartBeat
1                                         column=limit:high age limit, timestamp=1740309420180, value=48
1                                         column=limit:high value, timestamp=1740309420188, value=69
1                                         column=limit:low age limit, timestamp=1740309420176, value=0
1                                         column=limit:low value, timestamp=1740309420185, value=41
1                                         column=alert:alert_flag, timestamp=1740309420187, value=1
1                                         column=alert:alert_message, timestamp=1740309420187, value=Low BP than Normal
1                                         column=attribute:attribute, timestamp=1740309420187, value=bp
1                                         column=limit:high age limit, timestamp=1740309420188, value=100
1                                         column=limit:high value, timestamp=1740309420188, value=150
1                                         column=limit:low age limit, timestamp=1740309420186, value=40
1                                         column=limit:low value, timestamp=1740309420186, value=41
1                                         column=alert:alert_flag, timestamp=1740309420186, value=0
1                                         column=alert:alert_message, timestamp=1740309420186, value=Normal
1                                         column=attribute:attribute, timestamp=1740309420186, value=bp
1                                         column=limit:high age limit, timestamp=1740309420187, value=300
1                                         column=limit:high value, timestamp=1740309420187, value=180
1                                         column=limit:low age limit, timestamp=1740309420185, value=41
1                                         column=limit:low value, timestamp=1740309420185, value=151
1                                         column=alert:alert_flag, timestamp=1740309420185, value=1
1                                         column=alert:alert_message, timestamp=1740309420185, value=Higher BP than Normal
1                                         column=attribute:attribute, timestamp=1740309420185, value=bp
1                                         column=limit:high age limit, timestamp=1740309420186, value=9999
1                                         column=limit:high value, timestamp=1740309420186, value=9999
1                                         column=limit:low age limit, timestamp=1740309420184, value=41
1                                         column=limit:low value, timestamp=1740309420184, value=18
2                                         column=alert:alert_flag, timestamp=174030944113, value=0
2                                         column=alert:alert_message, timestamp=174030944137, value=Normal
2                                         column=attribute:attribute, timestamp=174030944137, value=heartBeat
2                                         column=limit:high age limit, timestamp=174030944128, value=40
2                                         column=limit:high value, timestamp=174030944128, value=78
2                                         column=limit:low age limit, timestamp=174030944118, value=0
2                                         column=limit:low value, timestamp=174030944126, value=78
2                                         column=alert:alert_message, timestamp=174030944118, value=Higher Heart Rate than Normal
2                                         column=attribute:attribute, timestamp=174030944118, value=heartBeat

```

i-0857d5ce17830a01f 

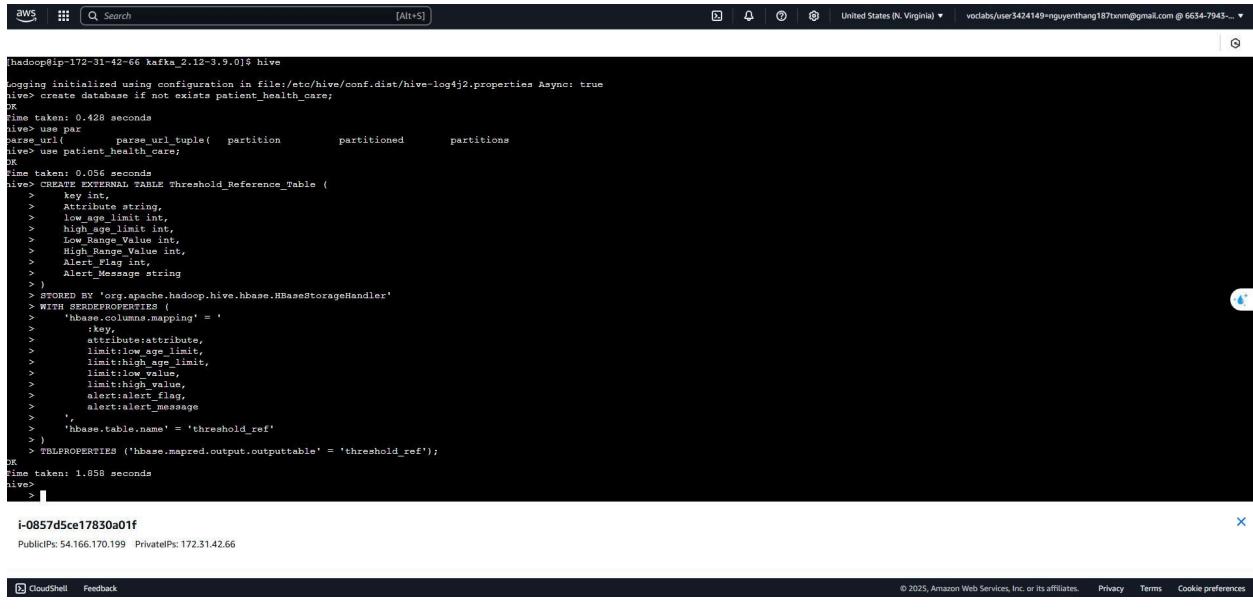
Public IPs: 54.166.170.199 Private IPs: 172.31.42.66

 CloudShell  Feedback © 2025, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

Script to create database and create table Threshold_Reference_Table:

```
hive
create database if not exists patient_health_care;
use patient_health_care;

CREATE EXTERNAL TABLE Threshold_Reference_Table (
    key int,
    Attribute string,
    low_age_limit int,
    high_age_limit int,
    Low_Range_Value int,
    High_Range_Value int,
    Alert_Flag int,
    Alert_Message string
)
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
WITH SERDEPROPERTIES (
    'hbase.columns.mapping' =
        ':key,
        attribute:attribute,
        limit:low_age_limit,
        limit:high_age_limit,
        limit:low_value,
        limit:high_value,
        alert:alert_flag,
        alert:alert_message
    ,
    'hbase.table.name' = 'threshold_ref'
)
TBLPROPERTIES ('hbase.mapred.output.outputtable' = 'threshold_ref');
```



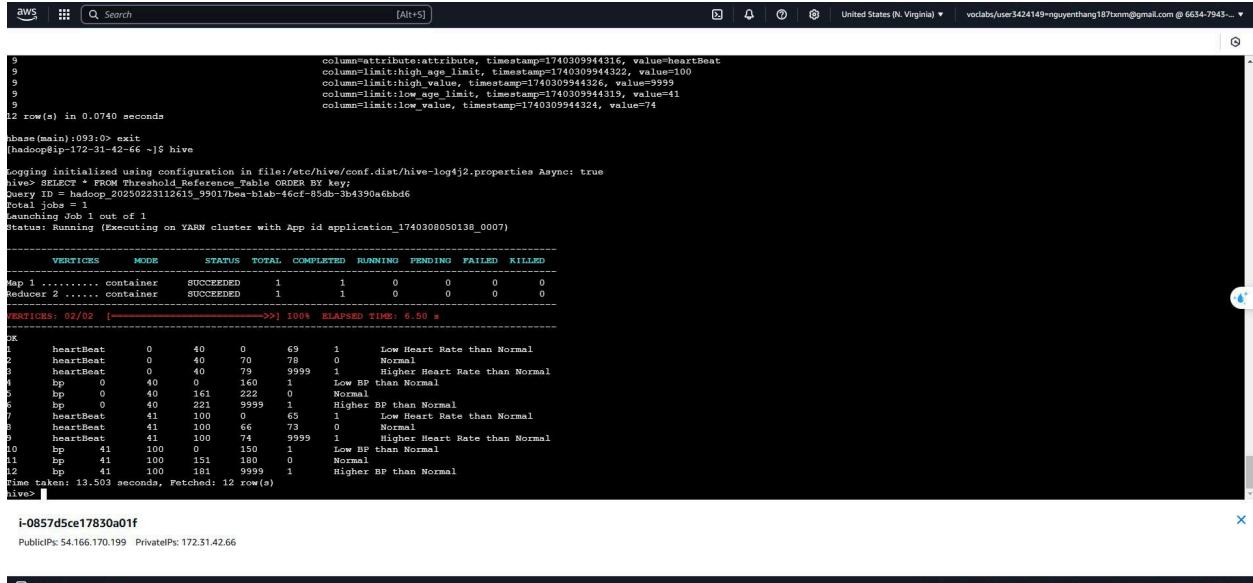
```
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j2.properties Async: true
hive> create database if not exists patient_health_care;
OK
Time taken: 0.428 seconds
hive> use pat
parse.url(
    parse.url_tuple(
        partition          partitioned      partitions
hive> use patient_health_care;
OK
Time taken: 0.056 seconds
hive> CREATE EXTERNAL TABLE Threshold_Reference_Table (
    key int,
    attribute:string,
    low_age_limit int,
    high_age_limit int,
    Low_Range_Value int,
    High_Range_Value int,
    Alert_Flag int,
    Alert_Message string
)
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
WITH SERDEPROPERTIES (
    'hbase.columns.mapping' = 'key,
    attribute:attribute,
    limit:low_age_limit,
    limit:high_age_limit,
    limit_low_value,
    limit_high_value,
    alert:alert_flag,
    alert:alert_message
',
    'hbase.table.name' = 'threshold_ref'
)
TBLPROPERTIES ('hbase.mapred.output.outputtable' = 'threshold_ref');
OK
Time taken: 1.858 seconds
hive> |
> |
i-0857d5ce17830a01f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Script to view data:

*SELECT * FROM Threshold_Reference_Table ORDER BY key;*



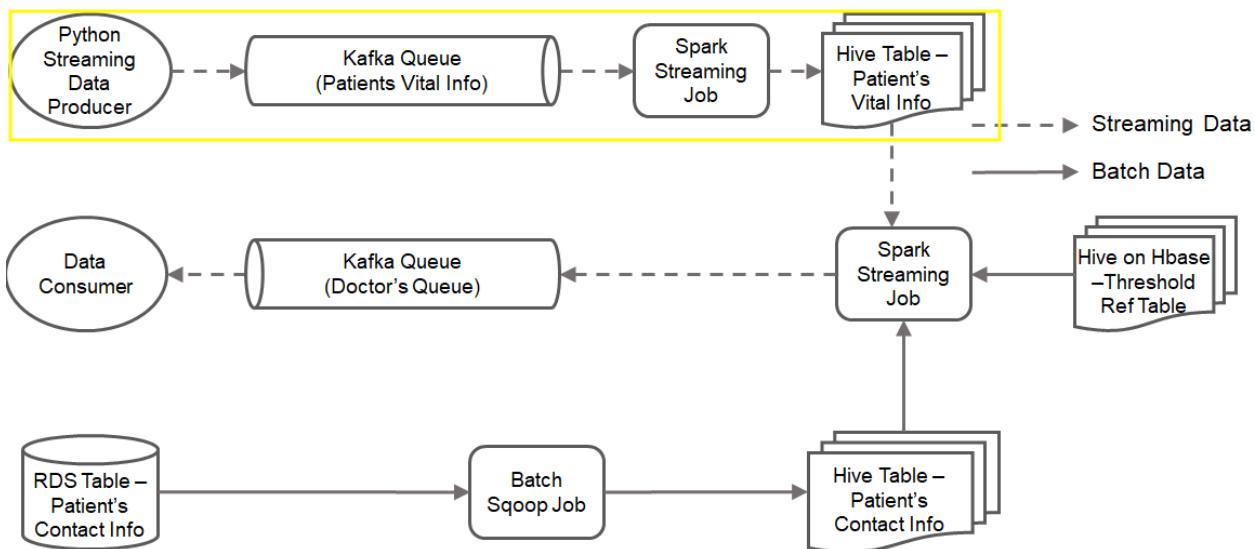
```
[aws] [Search] [Alt+S] United States (N. Virginia) vocabs/user3424149+nguyenthang187xnm@gmail.com @ 6634-7943... ▾
[hadoop@ip-172-31-42-66 ~]$ hive
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j2.properties Async: true
hive> SELECT * FROM Threshold_Reference_Table ORDER BY key;
Query ID = hadoop_20250223112615_99017bea-blab-46cf-85db-3b4390a6bb46
Total jobs = 1
Launching job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1740308050138_0007)

----- VERTICES  MODE STATUS TOTAL COMPLETED RUNNING PENDING FAILED KILLED -----
map 1 ..... container  SUCCEEDED 1 1 0 0 0 0
Reducer 2 ..... container  SUCCEEDED 1 1 0 0 0 0
----- VERTICES: 02/02 [---->>>] 100% ELAPSED TIME: 6.50 s
-----
OK
1 heartBeat 0 40 0 69 1 Low Heart Rate than Normal
2 heartBeat 0 40 70 78 0 Normal
3 heartBeat 0 40 75 9999 1 Higher Heart Rate than Normal
4 bp 0 40 0 160 7 Low BP than Normal
5 bp 0 40 161 222 0 Normal
6 bp 0 40 221 9999 1 Higher BP than Normal
7 heartBeat 41 100 0 65 1 Low Heart Rate than Normal
8 heartBeat 41 100 66 73 0 Normal
9 heartBeat 41 100 74 9999 1 Higher Heart Rate than Normal
10 bp 41 100 0 150 1 Low BP than Normal
11 bp 41 100 151 180 0 Normal
12 bp 41 100 181 9999 1 Higher BP than Normal
Time taken: 13.503 seconds, Fetched: 12 row(s)
hive> |
> |
i-0857d5ce17830a01f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

PART 3: Ingesting streaming data (patient's vital information) and storing it in a table.



1. Develop a Kafka Producer application in Python to simulate streaming data by reading information from RDS every second. This application will send patient vital signs to the Kafka topic named 'patients_vital_topic'.
2. Implement a Spark Streaming job to consume data from the 'patients_vital_topic' Kafka topic. This job should append a timestamp column to the data and save it in HDFS using the Parquet format.
3. Set up an external Hive table called 'Patients_Vital_Info' to continuously ingest and query the streaming data stored in the specified HDFS location.

Script to download and extract kafka:

```
wget https://dlcdn.apache.org/kafka/3.9.0/kafka\_2.12-3.9.0.tgz
tar -xzf kafka_2.12-3.9.0.tgz
cd kafka_2.12-3.9.0
ls
```



```

aws CloudShell Search [Alt+S] United States (N. Virginia) vocabs/user3424149+nguyenthang187xnm@gmail.com @ 6634-7945... ▾
[2025-02-23 10:58:09,639] INFO [KafkaServer id=0] End processing authorizer futures (kafka.server.KafkaServer)
[2025-02-23 10:58:09,639] INFO [KafkaServer id=0] Start processing enable request processing future (kafka.server.KafkaServer)
[2025-02-23 10:58:09,640] INFO [KafkaServer id=0] End processing enable request processing future (kafka.server.KafkaServer)
[2025-02-23 10:58:09,643] INFO Kafka version: 3.9.0 (org.apache.kafka.common.utils.AppInfoParser)
[2025-02-23 10:58:09,643] INFO Kafka commitId: 0 (org.apache.kafka.common.utils.AppInfoParser)
[2025-02-23 10:58:09,643] INFO Kafka startTimeMs: 1740308289640 (org.apache.kafka.common.utils.AppInfoParser)
[2025-02-23 10:58:09,644] INFO [KafkaServer id=0] started (kafka.server.KafkaServer)
[2025-02-23 10:58:09,718] INFO [zk-broker-0-to-controller-alter-partition-channel-manager]: Recorded new ZK controller, from now on will use node ip-172-31-42-66.ec2.internal:9092 (id: 0 rack: null) (kafka.server.NodeRecordControllerRequestThread)
[2025-02-23 10:58:09,718] INFO [zk-broker-0-to-controller-forwarding-channel-manager]: Recorded new ZK controller, from now on will use node ip-172-31-42-66.ec2.internal:9092 (id: 0 rack: null) (kafka.server.NodeRecordControllerRequestThread)

hadoop@ip-172-31-42-66 kafka_2.12-3.9.0$ bin/kafka-topics.sh --create --topic patients_vital_topic --bootstrap-server localhost:9092 --partitions 1 --replication-factor 1
WARNING: Due to limitations in the current implementation, replication factor must be 1 or higher. If you need a replication factor greater than 1, please use brokers.
[2025-02-23 10:59:05,158] INFO Creating topic patients_vital_topic with configuration () and initial partition assignment: Map(0 => ArrayBuffer(0)) (kafka.zk.AdminClient)
[2025-02-23 10:59:05,249] INFO [ReplicaFetcherManager on broker 0] Removed fetcher for partitions Set(patients_vital_topic-0) (kafka.server.ReplicaFetcherManager)
[2025-02-23 10:59:05,299] INFO [LogLoader partition.patients_vital_topic-0, dir=/tmp/kafka-logs/patients_vital_topic-0] Loading producer state till offset 0 with message format version 1 (kafka.log.UnifiedLog$)
[2025-02-23 10:59:05,312] INFO Created log for partition patients_vital_topic-0 in /tmp/kafka-logs/patients_vital_topic-0 with properties () (kafka.log.LogManager)
[2025-02-23 10:59:05,316] INFO [Partition patients_vital_topic-0 broker=0] No checkpoint highwatermark is found for partition patients_vital_topic-0 (kafka.cluster.Partition)
[2025-02-23 10:59:05,316] INFO [Partition patients_vital_topic-0 broker=0] Log offset for partition patients_vital_topic-0 with initial high watermark 0 (kafka.cluster.Partition)
Created topic patients_vital_topic.
hadoop@ip-172-31-42-66 kafka_2.12-3.9.0$ pip install --user kafka-python mysql-connector-python mysql-connector-repackaged
patients_vital_topic created.
hadoop@ip-172-31-42-66 kafka_2.12-3.9.0$ pip install --user kafka-python mysql-connector-python mysql-connector-repackaged
Collecting kafka-python
  Downloading https://files.pythonhosted.org/packages/2e/31/f41c1362aa0e12b7bddc1bd60eca57893f2dc84a1f5a5b9936ffef81b51/kafka_python-2.0.4-py2.py3-none-any.whl (251kB)
  100% |██████████| 251kB 5.0MB/s
Collecting mysql-connector-python
  Downloading https://files.pythonhosted.org/packages/e0/a0/7fae7fa0ecd701126d4b0-1b64f5b868b9-56155ef23c80b7858e22cda9c/mysql_connector_python-0.0.33-cp37-cp37-manylinux1_x86_64.whl (27.4MB)
  100% |██████████| 27.4MB 4.7KB/s
Collecting mysql-connector-repackaged
  Downloading https://files.pythonhosted.org/packages/00/c8/0f20b32691c3fe7dc013ca3971dac923b687c125c7e521f338e4a74ebef5/mysql-connector-repackaged-0.3.1.tar.gz (132kB)
  100% |██████████| 132kB 8.9MB/s
Collecting protobuf<=3.20.3,>=3.11.0 (from mysql-connector-python)
  Downloading https://files.pythonhosted.org/packages/4c/12/2e21d550c172e1a7f803d83b0b1693f7952c3c271eb2f155703012ae67a/protobuf-3.20.3-cp37-cp37-manylinux_2_5_x86_64_manylinux1_x86_64.whl (1.0MB)
  100% |██████████| 1.0MB 1.3MB/s
Installing collected packages: kafka-python, protobuf, mysql-connector-python, mysql-connector-repackaged
  Running setup.py install for kafka-python ... done
Successfully installed kafka-python-2.0.4 mysql-connector-python-0.0.33 mysql-connector-repackaged-0.3.1 protobuf-3.20.3
hadoop@ip-172-31-42-66 kafka_2.12-3.9.0$ i-0857d5ce17830a01f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

Create the python file, then copy the content, run the producer script and view log data:

`vim kafka_produce_patient_vitals.py`

`vim kafka_spark_patient_vitals.py`

`python kafka_produce_patient_vitals.py > kafka_logs.log 2>&1 &`
`cat kafka_logs.log`

```

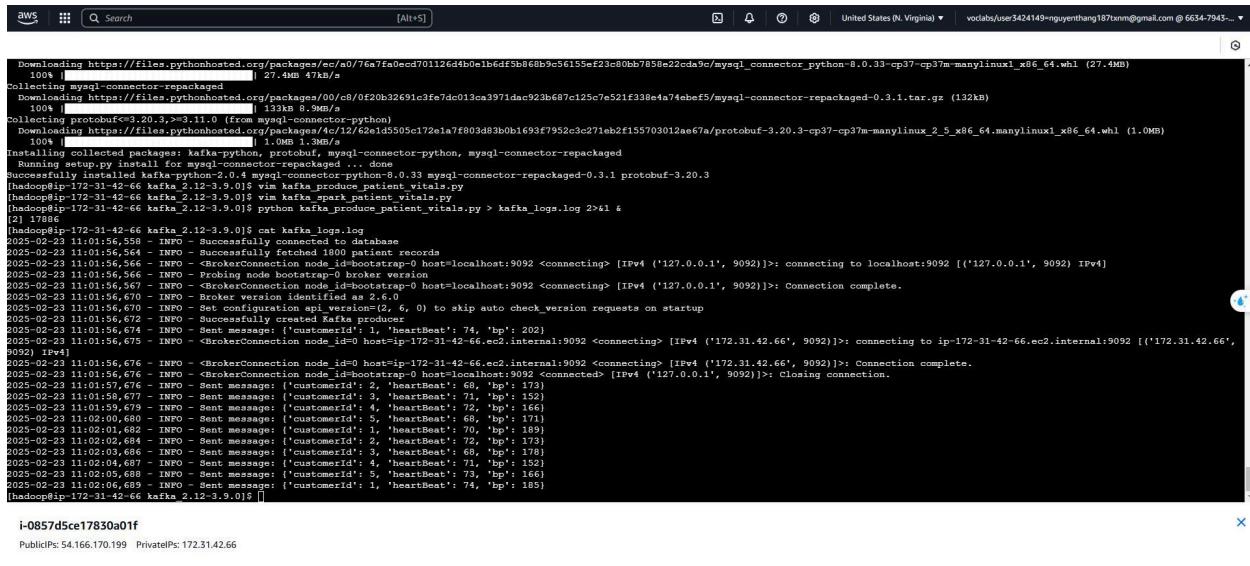
aws CloudShell Search [Alt+S] United States (N. Virginia) vocabs/user3424149+nguyenthang187xnm@gmail.com @ 6634-7945... ▾
[2025-02-23 11:01:56,558] INFO Successfully connected to database
[2025-02-23 11:01:56,558] INFO - Sent connection metadata to bootstrap
[2025-02-23 11:01:56,566] INFO - <BrokerConnection node_id=bootstrap-0 host=localhost:9092 <connecting> [IPv4 ('127.0.0.1', 9092)]>: connecting to localhost:9092 [("127.0.0.1", 9092)]
[2025-02-23 11:01:56,670] INFO - Probing node bootstrap-0 broker version
[2025-02-23 11:01:56,670] INFO - <BrokerConnection node_id=bootstrap-0 host=localhost:9092 <connecting> [IPv4 ('127.0.0.1', 9092)]>: Connection complete.
[2025-02-23 11:01:56,670] INFO - Broker version identified as 2.6.0
[2025-02-23 11:01:56,670] INFO - Sent configuration apply (version 5, 0) to skip auto check_version requests on startup
[2025-02-23 11:01:56,672] INFO - Successfully connected to Kafka producer
[2025-02-23 11:01:56,674] INFO - Sent message: ('customerid': 1, 'heartBeat': 74, 'bp': 202)
[2025-02-23 11:01:56,675] INFO - <BrokerConnection node_id=0 host=ip-172-31-42-66.ec2.internal:9092 <connecting> [IPv4 ('172.31.42.66', 9092)]>: connecting to ip-172-31-42-66.ec2.internal:9092 [("172.31.42.66", 9092)]
[2025-02-23 11:01:56,676] INFO - <BrokerConnection node_id=0 host=ip-172-31-42-66.ec2.internal:9092 <connecting> [IPv4 ('172.31.42.66', 9092)]>: Connection complete.
[2025-02-23 11:01:56,676] INFO - <BrokerConnection node_id=0 host=ip-172-31-42-66.ec2.internal:9092 <connected> [IPv4 ('172.31.42.66', 9092)]>: Closing connection.
[2025-02-23 11:01:56,677] INFO - Sent message: ('customerid': 3, 'heartBeat': 71, 'bp': 152)
[2025-02-23 11:01:56,677] INFO - <BrokerConnection node_id=0 host=ip-172-31-42-66.ec2.internal:9092 <connected> [IPv4 ('172.31.42.66', 9092)]>: Closing connection.
[2025-02-23 11:02:00,680] INFO - Sent message: ('customerid': 5, 'heartBeat': 68, 'bp': 171)
[2025-02-23 11:02:01,682] INFO - Sent message: ('customerid': 1, 'heartBeat': 70, 'bp': 189)
[2025-02-23 11:02:02,684] INFO - Sent message: ('customerid': 2, 'heartBeat': 72, 'bp': 173)
[2025-02-23 11:02:03,686] INFO - Sent message: ('customerid': 3, 'heartBeat': 68, 'bp': 178)
[2025-02-23 11:02:04,688] INFO - Sent message: ('customerid': 4, 'heartBeat': 73, 'bp': 172)
[2025-02-23 11:02:05,688] INFO - Sent message: ('customerid': 5, 'heartBeat': 73, 'bp': 166)
[2025-02-23 11:02:06,689] INFO - Sent message: ('customerid': 1, 'heartBeat': 74, 'bp': 185)
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ i-0857d5ce17830a01f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

`i-0857d5ce17830a01f`

Spark Streaming Job 1

`spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.11:2.4.5 kafka_spark_patient_vitals.py`



```

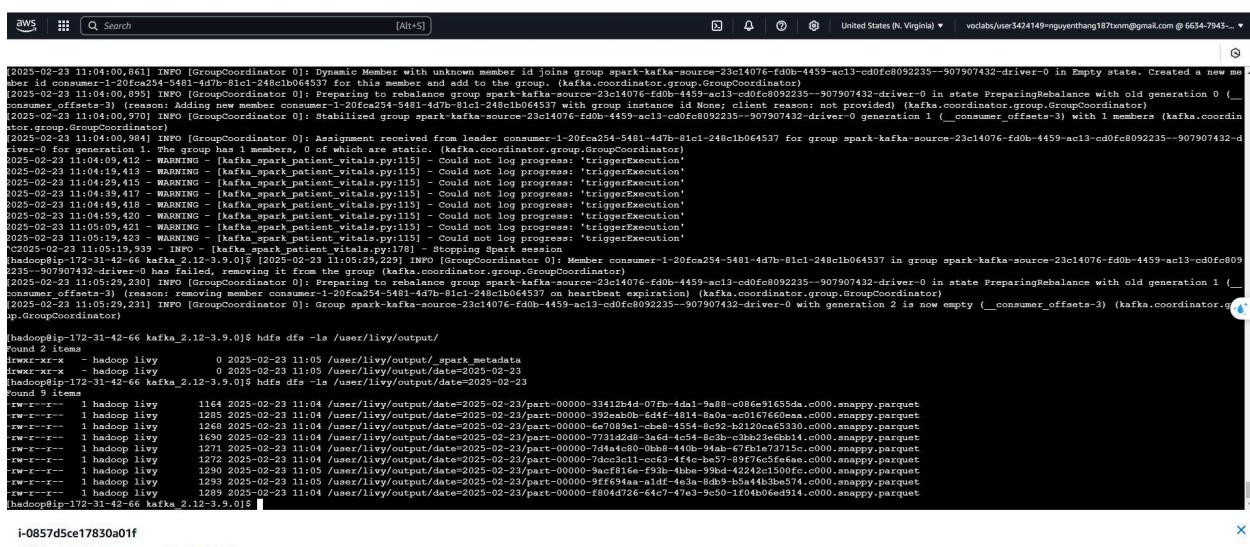
aws CloudShell Search [Alt+S] United States (N. Virginia) vocabs/user342419-nnguyenthang187txnm@gmail.com @ 6634-7943... ▾

[2] 17886
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ cat kafka_logs.log
2025-02-23 11:01:56.564 - INFO - Starting to fetch offsets from database
2025-02-23 11:01:56.564 - INFO - Successfully fetched 100 offset records
2025-02-23 11:01:56.564 - INFO - <BrokerConnection node_id=bootstrap-0 host=localhost:9092 <connecting> [IPv4 ('127.0.0.1', 9092)]>: connecting to localhost:9092 [('127.0.0.1', 9092) IPv4]
2025-02-23 11:01:56.567 - INFO - Probing node bootstrap-0 broker version
2025-02-23 11:01:56.567 - INFO - <BrokerConnection node_id=bootstrap-0 host=localhost:9092 <connecting> [IPv4 ('127.0.0.1', 9092)]>: Connection complete.
2025-02-23 11:01:56.600 - INFO - Broker version identified as 2.6.0
2025-02-23 11:01:56.670 - INFO - Using configuration setting auto.create.topics.enable (true, 6) to skip auto check_version requests on startup
2025-02-23 11:01:56.672 - INFO - Successfully created Kafka producer
2025-02-23 11:01:56.674 - INFO - Sent message: {'customerid': 1, 'heartBeat': 74, 'bp': 202}
2025-02-23 11:01:56.675 - INFO - <BrokerConnection node_id=0 host=ip-172-31-42-66.ec2.internal:9092 <connecting> [IPv4 ('172.31.42.66', 9092)]>: connecting to ip-172-31-42-66.ec2.internal:9092 [('172.31.42.66', 9092) IPv4]
2025-02-23 11:01:56.676 - INFO - <BrokerConnection node_id=0 host=ip-172-31-42-66.ec2.internal:9092 <connected> [IPv4 ('172.31.42.66', 9092)]>: Connection complete.
2025-02-23 11:01:56.676 - INFO - <BrokerConnection node_id=bootstrap-0 host=localhost:9092 <connected> [IPv4 ('127.0.0.1', 9092)]>: Closing connection.
2025-02-23 11:01:57.676 - INFO - Sent message: {'customerid': 2, 'heartBeat': 68, 'bp': 173}
2025-02-23 11:01:57.679 - INFO - Sent message: {'customerid': 3, 'heartBeat': 72, 'bp': 166}
2025-02-23 11:01:57.679 - INFO - Sent message: {'customerid': 4, 'heartBeat': 71, 'bp': 152}
2025-02-23 11:02:00.680 - INFO - Sent message: {'customerid': 5, 'heartBeat': 68, 'bp': 171}
2025-02-23 11:02:01.682 - INFO - Sent message: {'customerid': 1, 'heartBeat': 70, 'bp': 169}
2025-02-23 11:02:02.684 - INFO - Sent message: {'customerid': 2, 'heartBeat': 72, 'bp': 173}
2025-02-23 11:02:02.687 - INFO - Sent message: {'customerid': 3, 'heartBeat': 70, 'bp': 170}
2025-02-23 11:02:02.687 - INFO - Sent message: {'customerid': 4, 'heartBeat': 71, 'bp': 152}
2025-02-23 11:02:05.688 - INFO - Sent message: {'customerid': 5, 'heartBeat': 73, 'bp': 166}
2025-02-23 11:02:06.689 - INFO - Sent message: {'customerid': 1, 'heartBeat': 74, 'bp': 165}
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ i-0857d5ce17830a0f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

View data in HDFS:

`hdfs dfs -ls /user/livy/output/
hdfs dfs -ls /user/livy/output/date=2025-02-23`



```

aws CloudShell Search [Alt+S] United States (N. Virginia) vocabs/user342419-nnguyenthang187txnm@gmail.com @ 6634-7943... ▾

[2] 19701
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ hdfs dfs -ls /user/livy/output/
Found 2 items
-rw-r--r-- 1 hadoop livy 0 2025-02-23 11:05 /user/livy/output/_spark_metadata
-rw-r--r-- 1 hadoop livy 0 2025-02-23 11:05 /user/livy/output/date=2025-02-23
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ hdfs dfs -ls /user/livy/output/date=2025-02-23
Found 9 items
-rw-r--r-- 1 hadoop livy 1164 2025-02-23 11:05 /user/livy/output/date=2025-02-23/part-00000-3412b4d-07fb-4de1-9e80-c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1285 2025-02-23 11:04 /user/livy/output/date=2025-02-23/part-00000-392aab0b-644f-4814-8a0a-ec0167660eas.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1268 2025-02-23 11:04 /user/livy/output/date=2025-02-23/part-00000-6c7089be-1ches-454-8c92-b2120c65330.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1690 2025-02-23 11:04 /user/livy/output/date=2025-02-23/part-00000-77312d8-3ad6-4c54-8e9b-c3bb23e6b14.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1271 2025-02-23 11:04 /user/livy/output/date=2025-02-23/part-00000-7d4a4c80-02b8-4419-94ab-67f81e73715c.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1270 2025-02-23 11:04 /user/livy/output/date=2025-02-23/part-00000-7d4a4c80-02b8-4419-94ab-67f81e73715c.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1290 2025-02-23 11:05 /user/livy/output/date=2025-02-23/part-00000-9acf816e-f93b-4bba-99bd-42242c1500fc.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1293 2025-02-23 11:05 /user/livy/output/date=2025-02-23/part-00000-9eff94aa-aldf-4e3a-8cb5-b5a4b3be574.c000.snappy.parquet
-rw-r--r-- 1 hadoop livy 1289 2025-02-23 11:04 /user/livy/output/date=2025-02-23/part-00000-f804d726-64c7-47e3-9c50-1f04b06ed914.c000.snappy.parquet
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ i-0857d5ce17830a0f
PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

```

Script to create database and create table Patients_Vital_Info:

hive

```
create database if not exists patient_health_care;
```

```
use patient_health_care;
```

```
CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Vital_Info (
```

```
CustomerID INT,
```

```
BP INT,
```

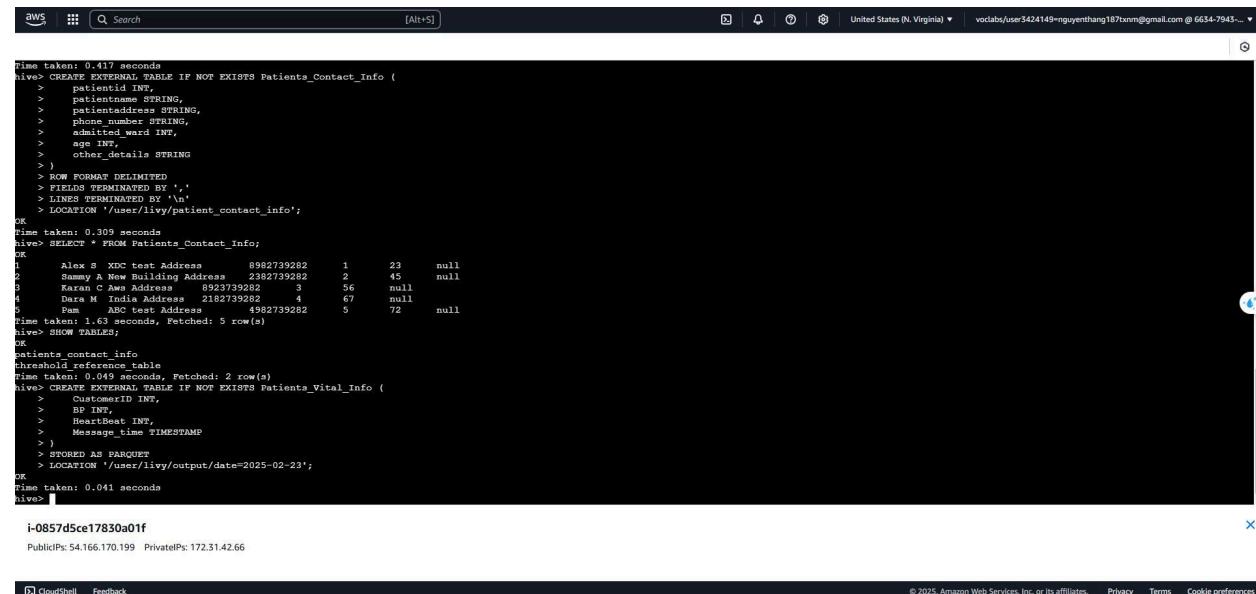
```
HeartBeat INT,
```

```
Message_time TIMESTAMP
```

```
)
```

```
STORED AS PARQUET
```

```
LOCATION '/user/livy/output/date=2025-02-23';
```



```

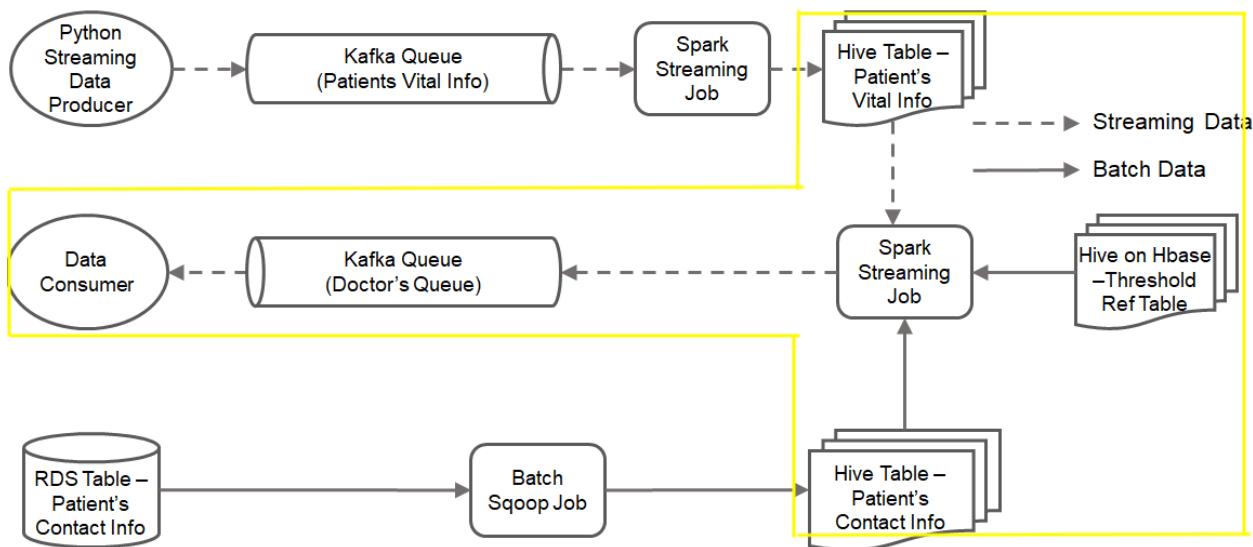
Time taken: 0.417 seconds
hive> CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Contact_Info (
    >     patientid INT,
    >     patientname STRING,
    >     patientaddress STRING,
    >     phone_number STRING,
    >     admission_id INT,
    >     age INT,
    >     other_details STRING
    > )
    > ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ','
    > LINES TERMINATED BY '\n'
    > LOCATION '/user/livy/patient_contact_info';
OK
Time taken: 0.309 seconds
hive> SELECT * FROM Patients_Contact_Info;
OK
1   Alex S  XDC test Address  8982739282  1    23    null
2   Sammy A New Building Address 2382739282  2    45    null
3   Ramya D New Address 6923739282  3    56    null
4   Dara M India Address  2182739282  4    67    null
5   Pam M ABC test Address  4982739282  5    72    null
Time taken: 1.63 seconds, Fetched: 5 row(s)
hive> SHOW TABLES;
OK
patients contact info
patients contact info
threshold reference table
Time taken: 0.049 seconds, Fetched: 2 row(s)
hive> CREATE EXTERNAL TABLE IF NOT EXISTS Patients_Vital_Info (
    >     CustomerID INT,
    >     BP INT,
    >     HeartBeat INT,
    >     Message_time TIMESTAMP
    > )
    > STORED AS PARQUET
    > LOCATION '/user/livy/output/date=2025-02-23';
OK
Time taken: 0.041 seconds
hive>

```

i-0857d5ce17830a01f
 PublicIPs: 54.166.170.199 PrivateIPs: 172.31.42.66

CloudShell Feedback © 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

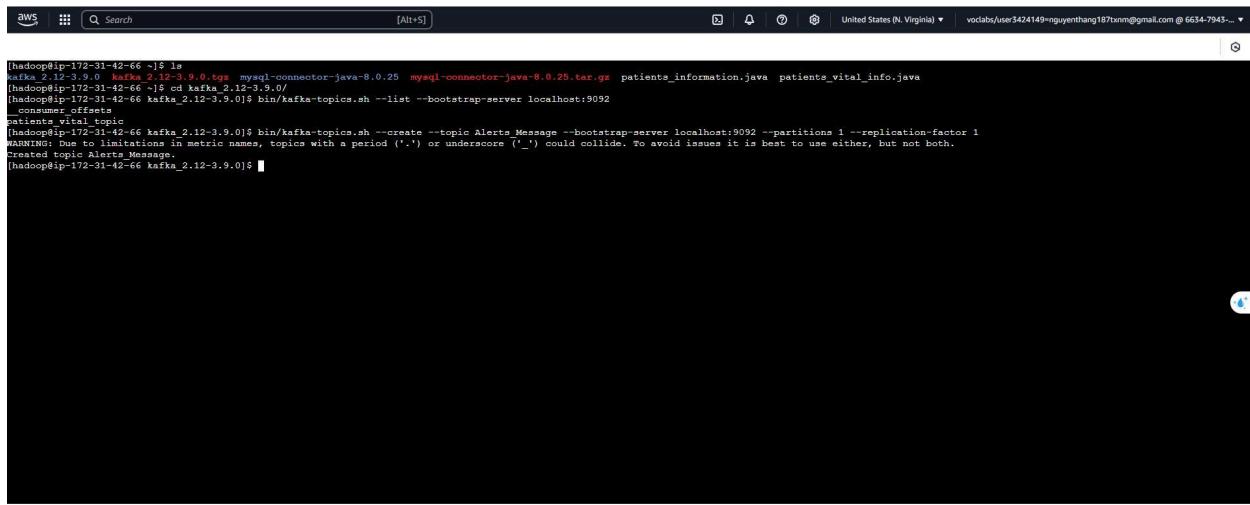
PART 4: Comparing vital information with threshold values, analyzing the data, and sending notifications if it exceeds the limits.



1. Establish a Kafka topic named **Alerts_Message** to capture and store irregular patient vital signs.
2. Develop a Spark Streaming job to retrieve and analyze data from three Hive tables. If any patient vitals are identified as irregular, the job should forward these records to the **Alerts_Message** Kafka topic.
3. Implement a Kafka consumer application to process and read messages from the **Alerts_Message** topic.
4. Configure the Kafka consumer to trigger email notifications via Amazon SNS for each message it processes, ensuring timely alerts for irregular patient vitals.

Create Alerts Message topic:

```
bin/kafka-topics.sh --create --topic Alerts_Message --bootstrap-server localhost:9092
--partitions 1 --replication-factor 1
```



```
[hadoop@ip-172-31-42-66 ~]$ ls
kafka_2.12-3.9.0.tgz  mysql-connector-java-8.0.25  mysql-connector-java-8.0.25.tar.gz  patients_information.java  patients_vital_info.java
[hadoop@ip-172-31-42-66 ~]$ cd kafka_2.12-3.9.0/
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ bin/kafka-topics.sh --list --bootstrap-server localhost:9092
com.vital.topic
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$ bin/kafka-topics.sh --create --topic Alerts_Message --bootstrap-server localhost:9092 --partitions 1 --replication-factor 1
WARNING: Due to limitations in metric names, topics with a period ('.') or underscore('_') could collide. To avoid issues it is best to use either, but not both.
Created topic Alerts_Message.
[hadoop@ip-172-31-42-66 kafka_2.12-3.9.0]$
```

i-0857d5ce17830a01f X

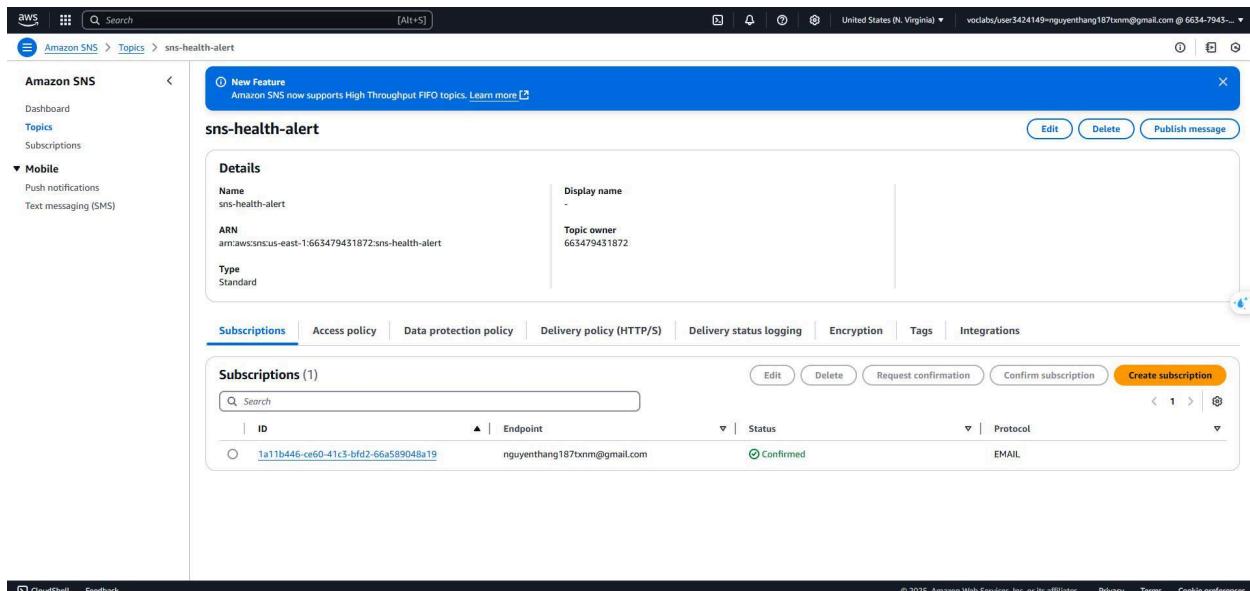
Public IPs: 54.166.170.199 Private IPs: 172.31.42.66

[CloudShell](#) [Feedback](#) © 2025, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

Spark Job to push irregular patient vital to topic Alerts Message:

`spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.11:2.4.5 --jars /usr/lib/hive/lib/hive-hbase-handler.jar,/usr/lib/hbase/lib/hbase-client-1.4.13.jar,/usr/lib/hbase/lib/hbase-common-1.4.13.jar,/usr/lib/hbase/lib/hbase-server-1.4.13.jar kafka_spark_generate_alerts.py`

Create topic SNS:



Amazon SNS > Topics > sns-health-alert

sns-health-alert

Details

Name	sns-health-alert	Display name	-
ARN	arn:aws:sns:us-east-1:663479431872:sns-health-alert	Topic owner	663479431872
Type	Standard		

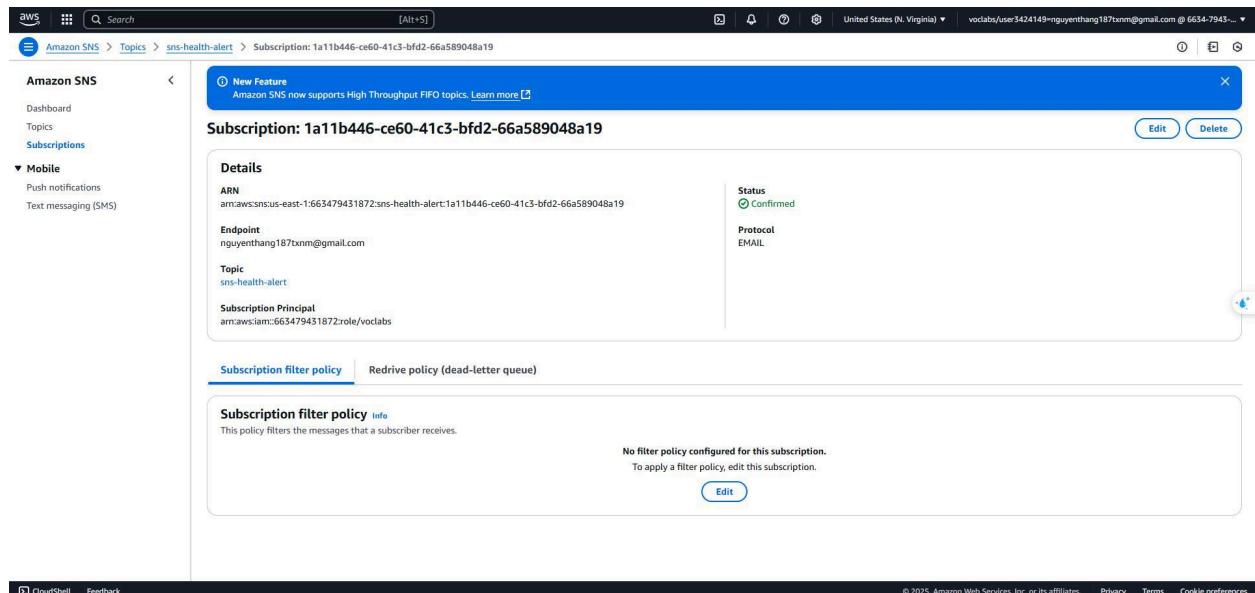
Subscriptions (1)

ID	Endpoint	Status	Protocol
Ja11b446-ce60-41c3-bfd2-66a589048a19	nguyenthang187xnm@gmail.com	Confirmed	EMAIL

[Edit](#) [Delete](#) [Publish message](#) [Create subscription](#)

[CloudShell](#) [Feedback](#) © 2025, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

Create subscriptions:



The screenshot shows the AWS SNS console with the following details:

- Subscription:** 1a11b446-ce60-41c3-bfd2-66a589048a19
- ARN:** arn:aws:sns:us-east-1:663479431872:sns-health-alert:1a11b446-ce60-41c3-bfd2-66a589048a19
- Status:** Confirmed
- Protocol:** EMAIL
- Endpoint:** nguyenthang187txnm@gmail.com
- Topic:** sns-health-alert
- Subscription Principal:** arn:aws:iam::663479431872:role/voclabs

Below the subscription details, there is a section for "Subscription filter policy" which states: "No filter policy configured for this subscription. To apply a filter policy, edit this subscription." There is an "Edit" button next to this message.

SNS Subscription Confirmation Email:

AWS Notification - Subscription Confirmation



AWS Notifications <no-reply@sns.amazonaws.com>
đến nguyenthang187txnm ▾

You have chosen to subscribe to the topic:
arn:aws:sns:us-east-1:663479431872:sns-health-alert

To confirm this subscription, click or visit the link below (If this was in error no action is necessary):
[Confirm subscription](#)

Please do not reply directly to this email. If you wish to remove yourself from receiving all future SNS subscription confirmation requests please send an email to [sns-opt-out](#)

Kafka Consumer:

```
pip install boto3
python kafka_consume_alerts.py
```

Alert Email:

Health Alert Notification [Hộp thư đến](#) x

 AWS Notifications <no-reply@sns.amazonaws.com>
đến nguyenthang187txnm ▾

```
{"patient_name": "Karan C",
"age": 56,
"patient_address": "Aws Address",
"phone_number": "8923739282",
"admitted_ward": 3,
"bp": 171,
"heart_beat": 56,
"input_message_time": "2025-02-23T20:58:12.698Z",
"alert_message": "Low Heart Rate than Normal",
"alert_generated_time": "2025-02-23T20:58:16.266Z"}
```

--
If you wish to stop receiving notifications from this topic, please click or visit the link below to unsubscribe:
<https://sns.us-east-1.amazonaws.com/unsubscribe.html?SubscriptionArn=arn:aws:sns:us-east-1:663479431872:sns-health-alert:1a11b446-ce60-41c3-bfd2-66a589048a19&Endpoint=nguyenthang187txnm@gm>

Please do not reply directly to this email. If you have any questions or comments regarding this email, please contact us at <https://aws.amazon.com/support>