

## CAPSTONE PROJECT INSTANT HEALTH ALERT SYSTEM SUBMITTED BY : Sushant Sarswat (C-38)

### SQOOP COMMAND

1. The following command can be used for Sqoop Job:

```
sqoop import \
--connect jdbc:mysql://upgradtest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
--username student \
--password STUDENT123 \
--table patients_vital_info \
--direct \
-m 1 \
--hive-import \
--create-hive-table \
--hive-table Patients_Contact_Info \
--fields-terminated-by ';
```

```
[hadoop@ip-172-31-39-26 ~]$ sqoop import --connect jdbc:mysql://upgradtest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase --username student --password STUDENT123 --table patients_informat
ion --direct -m 1 --hive-import --create-hive-table --hive-table PatientContactInfo --fields-terminated-by ';'
Warning: /usr/lib/sqoop/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hive/lib/log4j-slf4j-impl-2.17.1.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]
22/11/27 20:38:02 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
22/11/27 20:38:02 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
22/11/27 20:38:02 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
22/11/27 20:38:02 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 cl
ass is generally unnecessary.
22/11/27 20:38:03 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'patients_information' AS t LIMIT 1
22/11/27 20:38:03 INFO manager.DirectMySQManager: Beginning mysqldump fast path import
22/11/27 20:38:03 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce
Note: /tmp/sqoop-hadoop/compile/5209fbc347517cb7d8102a8bd2433a40/patients_information.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
22/11/27 20:38:07 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-hadoop/compile/5209fbc347517cb7d8102a8bd2433a40/patients_information.jar
22/11/27 20:38:07 INFO mapreduce.ImportJobBase: Beginning import of patients_information
22/11/27 20:38:07 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
22/11/27 20:38:08 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
22/11/27 20:38:08 INFO client.RMProxy: Connecting to ResourceManager at ip-172-31-39-26.ec2.internal/172.31.39.26:8032
22/11/27 20:38:08 INFO client.AMProxy: Connecting to Application History server at ip-172-31-39-26.ec2.internal/172.31.39.26:10200
22/11/27 20:38:10 INFO db.DBInputFormat: Using read committed transaction isolation
22/11/27 20:38:10 INFO mapreduce.JobSubmitter: number of splits:1
22/11/27 20:38:11 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1669577452461_0004
22/11/27 20:38:11 INFO conf.Configuration: resource-types.xml not found
22/11/27 20:38:11 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
22/11/27 20:38:11 INFO resource.ResourceUtils: Adding resource type - name = memory-mb, units = Mi, type = COUNTABLE
22/11/27 20:38:11 INFO resource.ResourceUtils: Adding resource type - name = vcores, units = , type = COUNTABLE
22/11/27 20:38:11 INFO impl.YarnClientImpl: Submitted application application_1669577452461_0004
22/11/27 20:38:11 INFO mapreduce.Job: The url to track the job: http://ip-172-31-39-26.ec2.internal:20888/proxy/application_1669577452461_0004/
22/11/27 20:38:11 INFO mapreduce.Job: Running job: job_1669577452461_0004
22/11/27 20:38:18 INFO mapreduce.Job: Job job_1669577452461_0004 running in uber mode : false
22/11/27 20:38:18 INFO mapreduce.Job: map 0% reduce 0%
22/11/27 20:38:24 INFO mapreduce.Job: map 100% reduce 0%
22/11/27 20:38:25 INFO mapreduce.Job: Job job_1669577452461_0004 completed successfully
```

2. Snapshot of collected data

```
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      Karan C Aws Address           8923739282      3      56      NULL
4      Dara M India Address          2182739282      4      67      NULL
5      Pam ABC test Address          4982739282      5      72      NULL
Time taken: 2.316 seconds, Fetched: 5 row(s)
```