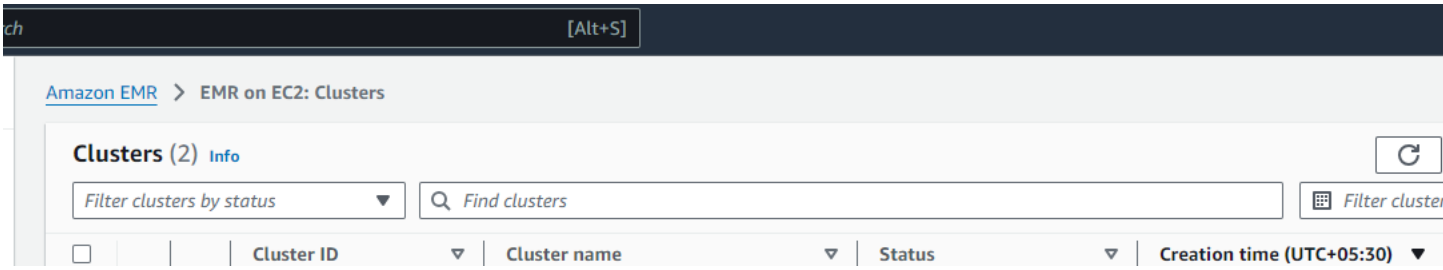
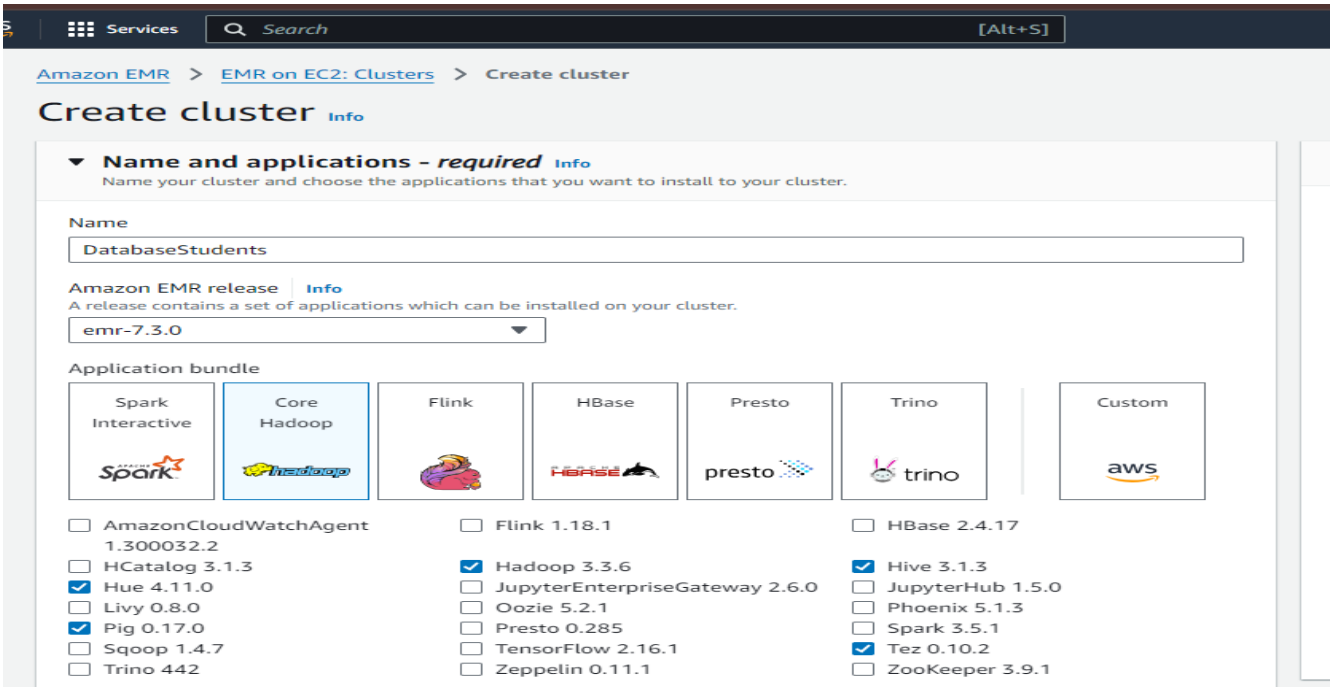
CLOUD COMPUTING PRACTICAL ASSIGNMENT NO:10

Installation and configuration of cloud Hadoop and demonstrate simple query Prepare Screen shots file and also write down the steps. Make single word or PDF file .

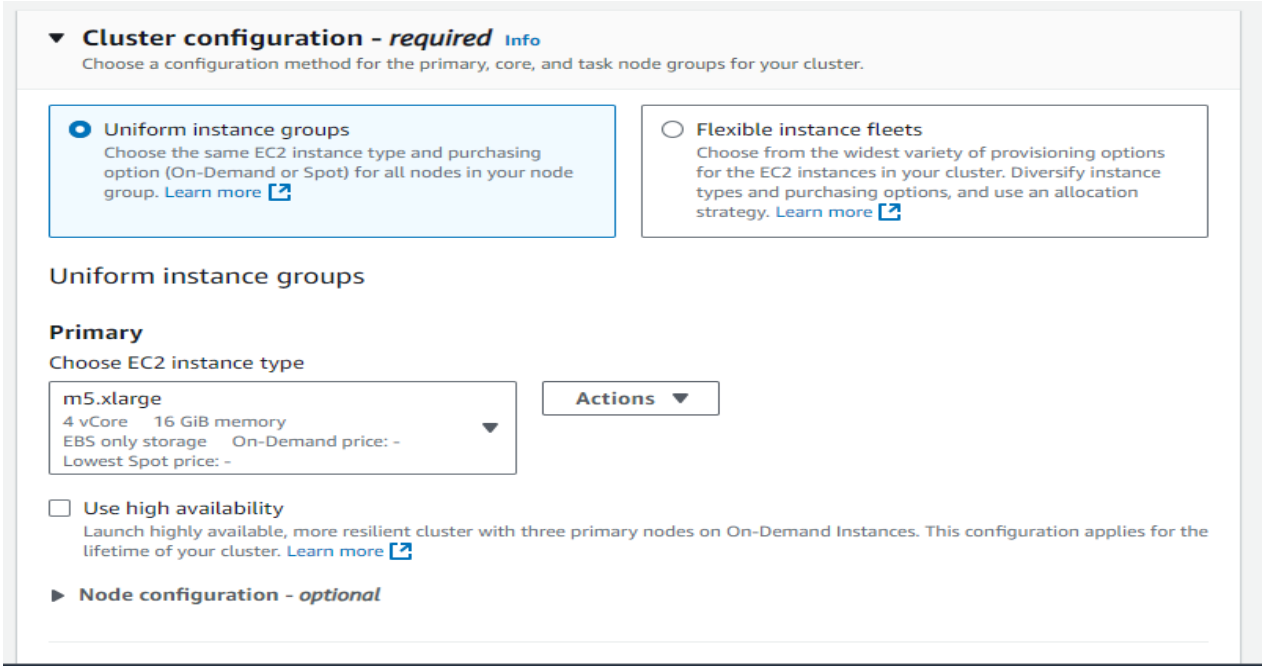
Step I : Go to EMR service and click on to create Cluster.



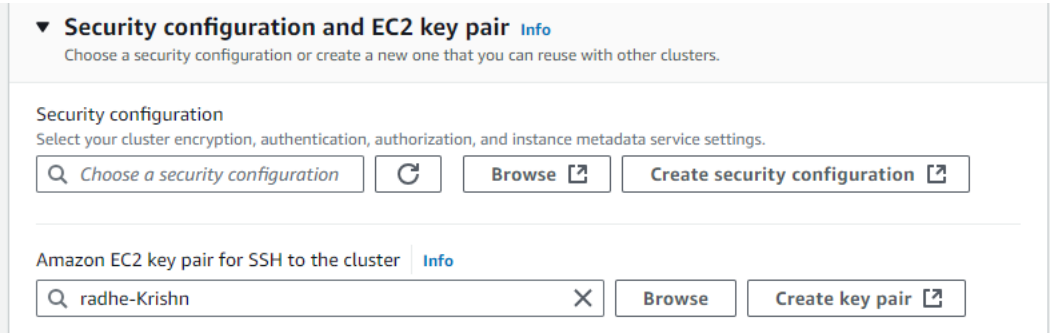
Step II: give the name to cluster and in application Bundle select the Core Hadoop.



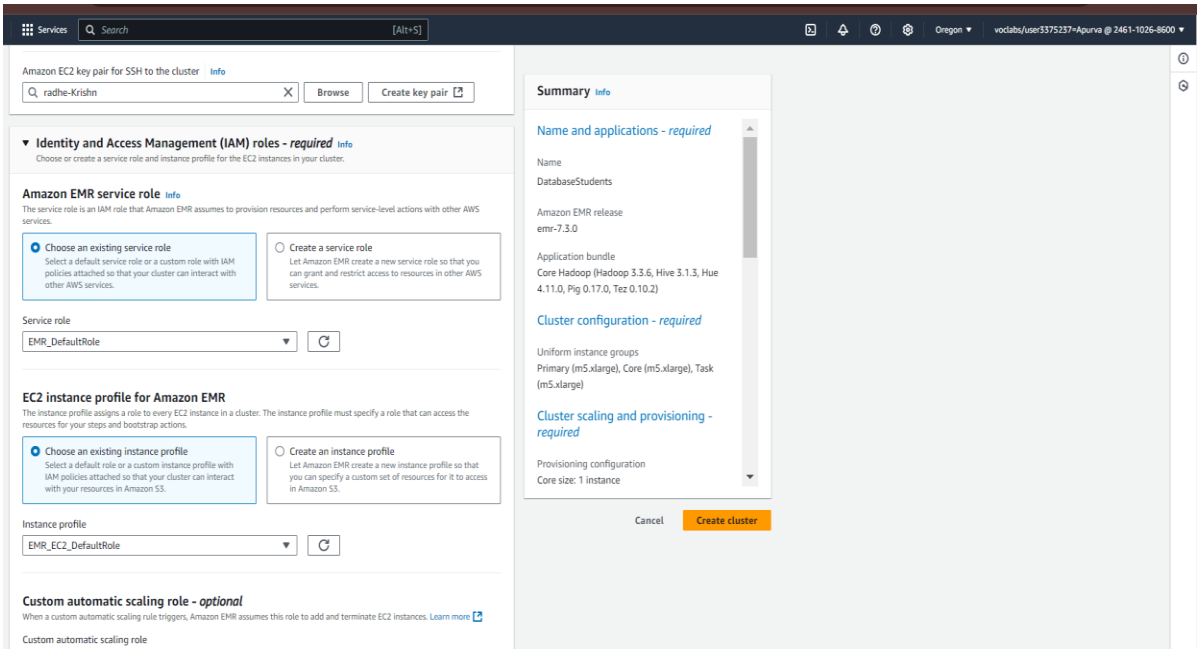
Step III: In Cluster Configuration keep uniform instance group as it is.



Step IV: Keep all as it is only in Security Configuration and EC2 key pair select the key-pair.

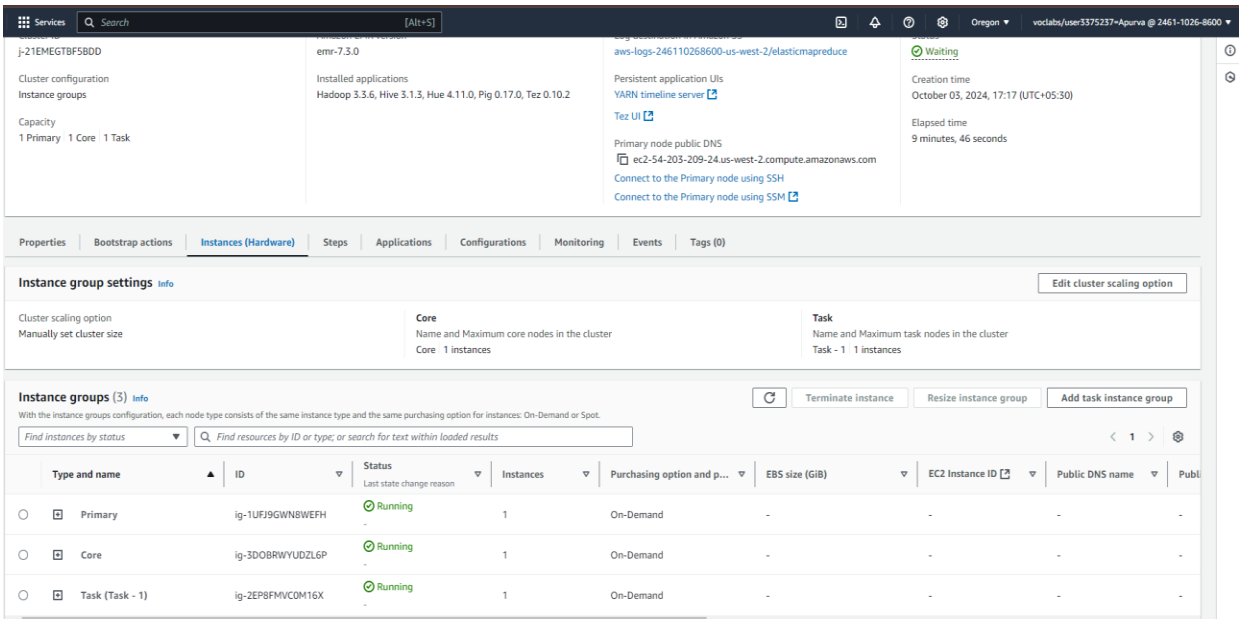


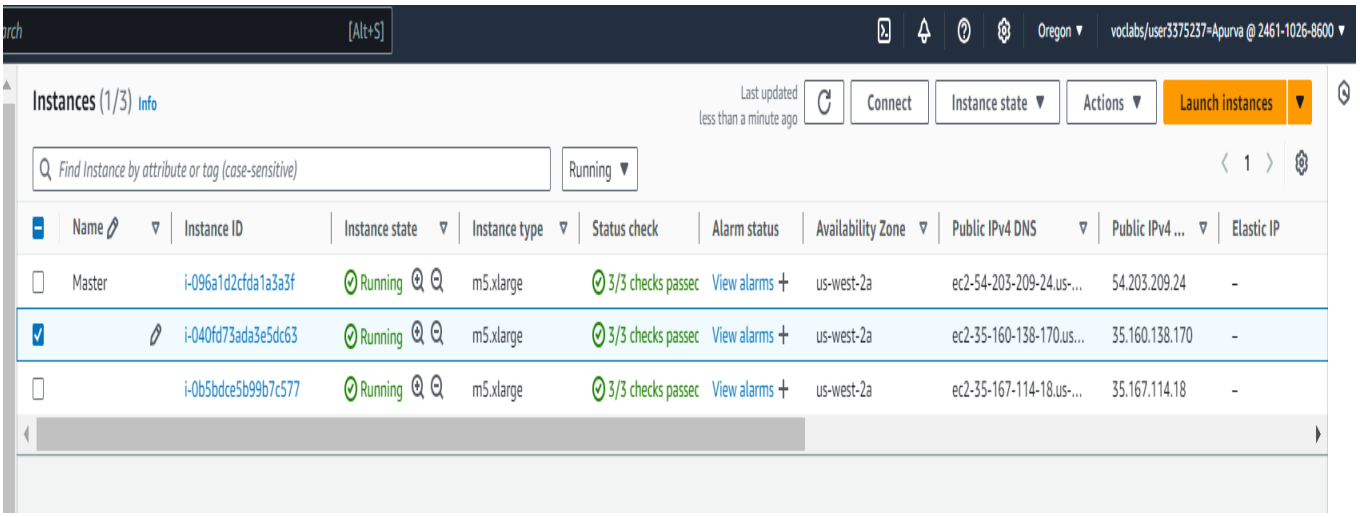
Step V: In Amazon EMr service role and ec2 instance profile for Amazon EMR sleect EMR DefaultRole and EMR_EC2_DefaultRole respectively.and click on to create cluster.



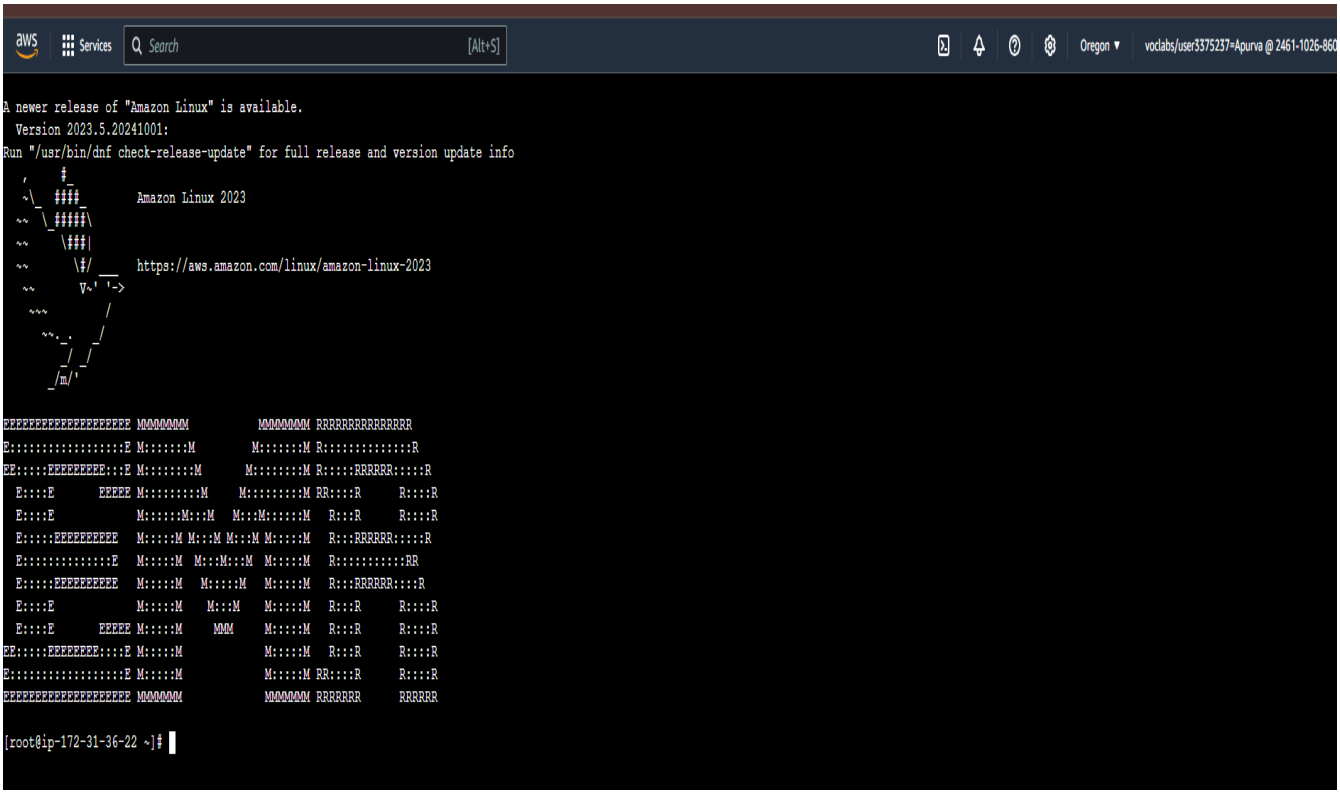
Step VI: Wait for few minute for creation of cluster.

Step VII: Once Instances are in running state then go to EC2 Service and connect the Ec2 Instance.





Step VIII: then connect the instance.it is connected successfully.



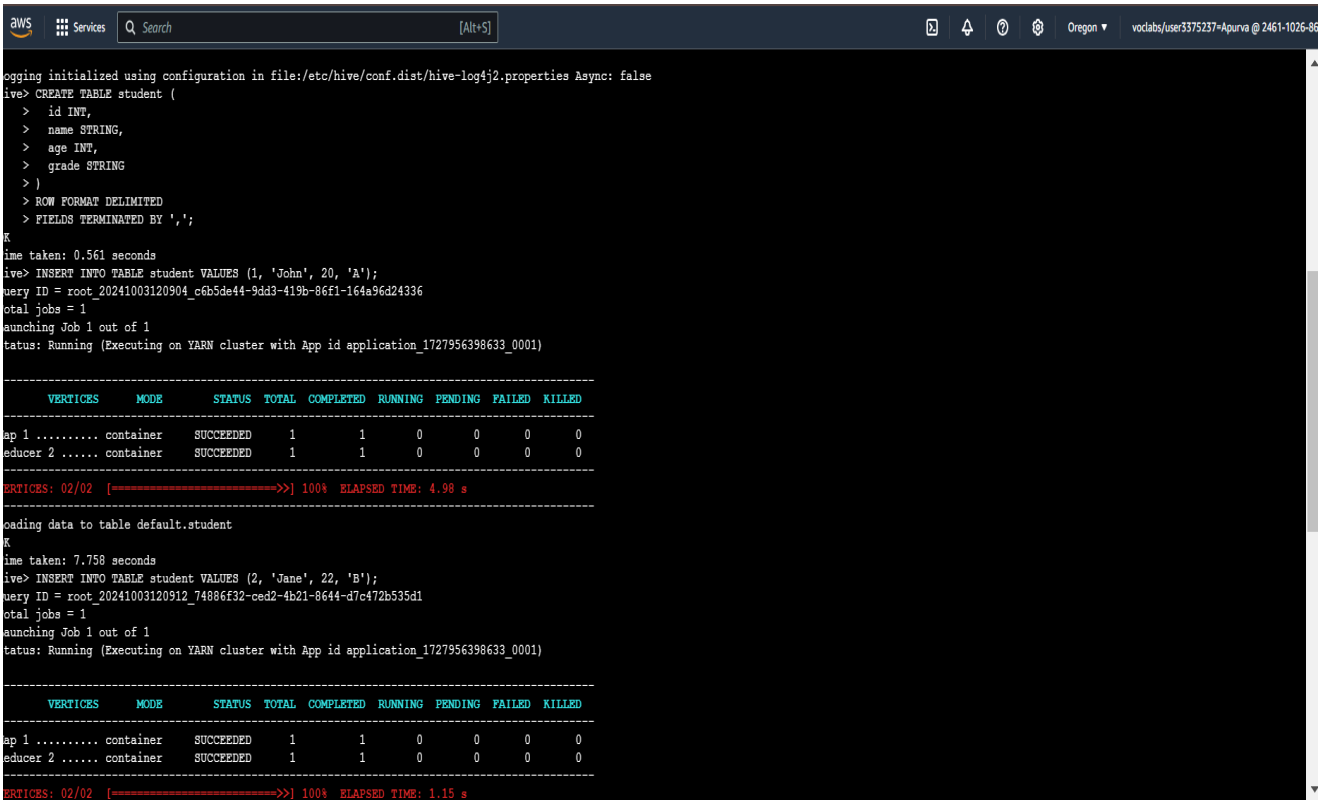
Step IX: then give command hive to terminal.The Hive command-line interface (CLI) is used for querying and

managing data stored in Hadoop's distributed file system (HDFS) using SQL-like syntax.

Step X: in that Create the Student Table

```
CREATE TABLE student (
  id INT,
  name STRING,
  age INT,
  grade STRING
)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';
```

Step XI: then Insert the data in it



Step XII: run Queries on terminal.

to fetch specific student data. For instance, to retrieve students with an 'A' grade:

SELECT * FROM student WHERE grade='A';

```
--------------------------------------------------------------------------------
        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: 0.13 s
--------------------------------------------------------------------------------
Loading data to table default.student
OK
Time taken: 1.039 seconds
hive> INSERT INTO TABLE student VALUES (4, 'Bob', 23, 'C');
Query ID = root_20241003120915_f7fc186a-107c-4c4b-8348-093dff164686
Total jobs = 1
Launching Job 1 out of 1
Status: Running (Executing on YARN cluster with App id application_1727956398633_0001)

--------------------------------------------------------------------------------
        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: 1.01 s
--------------------------------------------------------------------------------
Loading data to table default.student
OK
Time taken: 1.851 seconds
hive> SELECT * FROM student WHERE grade='A';
OK
1       John    20      A
3       Alice   21      A
Time taken: 0.294 seconds, Fetched: 2 row(s)
hive>
```

Step XIII: Retrieve All Students Ordered by Age:

This query will return all students and order them by their age in ascending order.

SELECT * FROM student ORDER BY age ASC;

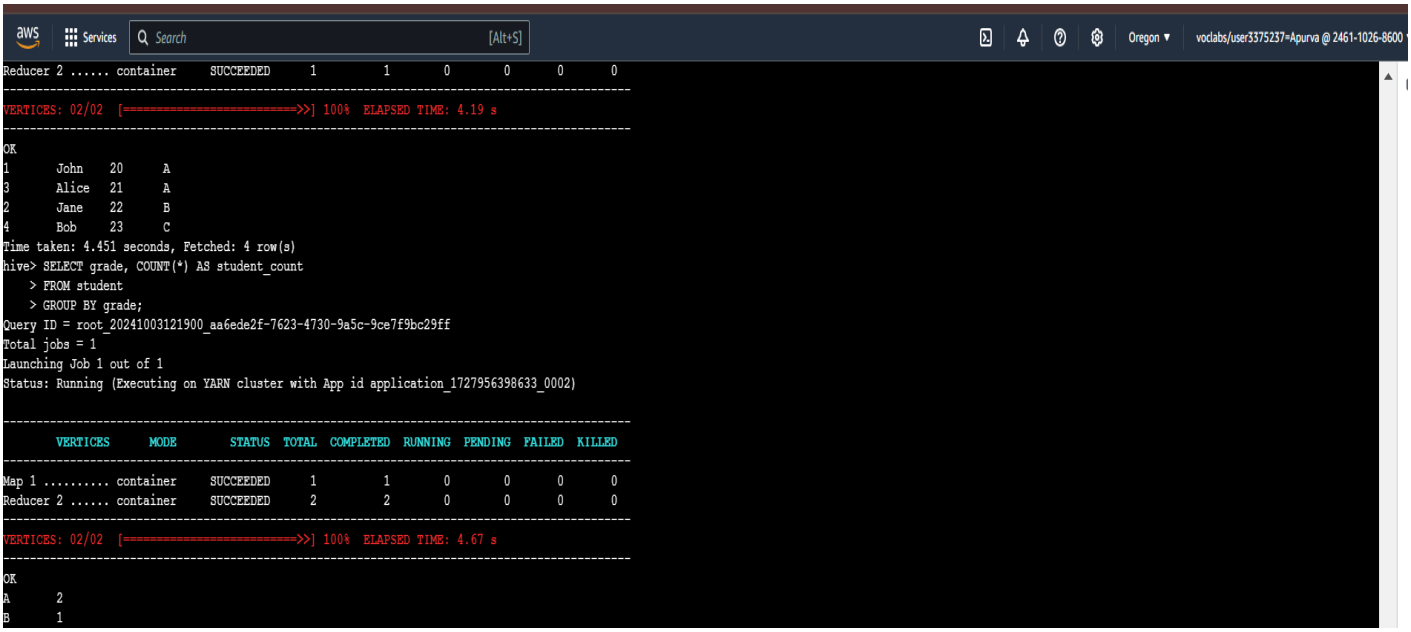
```
Time taken: 0.294 seconds, Fetched: 2 row(s)
hive> SELECT * FROM student ORDER BY age ASC;
Query ID = root_20241003121619_bb5d4f2b-f6f8-474b-9f84-a4c99f2dc9af
Total jobs = 1
Launching Job 1 out of 1
Tez session was closed. Reopening...
Session re-established.
Session re-established.
Status: Running (Executing on YARN cluster with App id application_1727956398633_0002)

--------------------------------------------------------------------------------
        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: 4.36 s
--------------------------------------------------------------------------------
OK
1       John    20      A
3       Alice   21      A
2       Jane    22      B
4       Bob     23      C
Time taken: 9.567 seconds, Fetched: 4 row(s)
```

Step XIV: Count the Number of Students in Each Grade:

This query will give you a count of students grouped by their grade.

SELECT grade, COUNT(*) AS student_count

FROM student GROUP BY grade;



Then delete the instances and cluster properly.