Session 11: ADVANCED HBASE

Case Study II: Customer _ Transaction

Case Study Description

Let us take up the CUSTOMER and TRANSACTIONS table we have created in the Let's Do Together section. Let us solve the following use cases using these tables :-

1. Find out the number of transactions done by each customer (These should be take up in module 8 itself)

Solution:

```
hive> SELECT a.custid,a.fname, COUNT(*) FROM Customer a JOIN TRANSACTIONS b ON a.custid=b.custno GROUP BY a.custid,a.fname;
```

Output:

```
Stage-Stage-2: Map: 1  Reduce: 1   Cumulative CP
Total MapReduce CPU Time Spent: 5 seconds 500 ms
OK
101     Amitabh 2
102     Sharukh 1
104     Anubahv 1
105     Pawan   1
106     Aamir   1
107     Salman  1
108     Ranbir  1
Time taken: 60.122 seconds, Fetched: 7 row(s)
```

2. Create a new table called TRANSACTIONS_COUNT. This table should have 3 fields - custid, fname and count. (Again to be done in module 8)

Solution:

```
Time taken: 60.122 seconds, Fetched: 7 row(s)
hive> create table transactions_count(custid INT,fname String,count INT)row format delimited fields terminated by ',';
OK
Time taken: 0.506 seconds
```

3. Now write a hive query in such a way that the query populates the data obtained in Step 1 above and populate the table in step 2 above. (This has to be done in module 9).

Solution:

```
hive> INSERT INTO TRANSACTIONS_COUNT (SELECT a.custid,a.fname,COUNT(*) FROM CUSTOMER a JOIN TRANSACTIONS b ON a.custid=b.cust
no GROUP BY a.custid,a.fname);
WARNING: Hive on MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execu
```

```
hive> SELECT * FROM TRANSACTIONS_COUNT;
OK
101     Amitabh 2
102     Sharukh 1
104     Anubahv 1
105     Pawan   1
106     Aamir   1
107     Salman  1
108     Ranbir  1
Time taken: 0.264 seconds, Fetched: 7 row(s)
hive>
```

4. Now lets make the TRANSACTIONS_COUNT table Hbase complaint. In the sence, use Ser Des And Storate handler features of hive to change the TRANSACTIONS_COUNT table to be able to create a TRANSACTIONS table in Hbase. (This has to be done in module 10)

Solution:

5. Now insert the data in TRANSACTIONS_COUNT table using the query in step 3 again, this should populate the Hbase TRANSACTIONS table automatically (This has to be done in module 10)

Solution:

6. Now from the Hbase level, write the Hbase java API code to access and scan the TRANSACTIONS table data from java level.

Solution: