**Name : Hari Yadav**

**Class Id: 15**

**Source Code :**

**Video/Demo :**

**Introduction**

This lab assignment deals with understanding the concepts of hadoop - map reduce and also implementing a map reduce algorithm to find the mutual friends concept. The second part deals with comparison of Hbase and Cassandra based on a use case and user own data set.

**Objective**

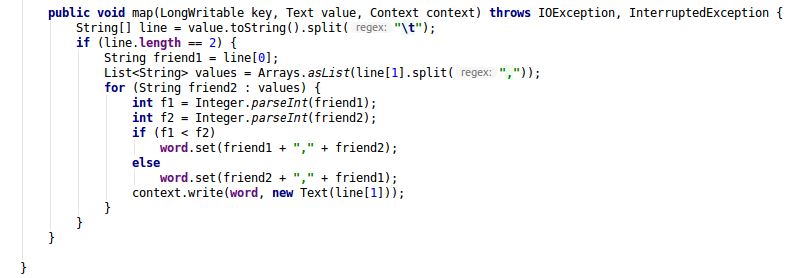
**1. Implement MapReduce algorithm for finding Facebook common friends problem and run the MapReduce job on Apache Hadoop. Show your implementation through map-reduce diagram**

**Approach**

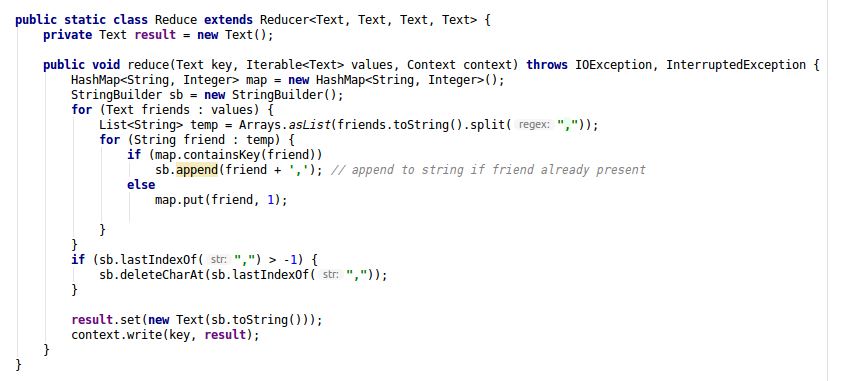
First, take the input as discussed in the use case as "A -> B C D, B -> A C D E, C -> A B D E, D -> A B C E, E -> B C D" Then in map phase find the mutual friends of two people. Group them based on the mutual pair key and finally reduce them to get the mutual friends list.

**Workflow**

Create a mapper class as shown in the code snippet below. Each line of the input file is split based on "tab". Then its length is computed = 2, where the first part is source or base user and the rest of the split is considered as list of friends of the user. Then the keys are prepared as (A,B) or (B,A) based on the integer values of A & B in the input.



Create a reducer class where the data is grouped based on the key values (A,B) or (B,C) and their list of friends as produced. Then finally reduced to find the mutual friends of (A,B).



A main method which acts as a driver to set mapper and reducer class which takes the input and produces the output.



**Data set and Parameter**

The input file for this can be found at <https://github.com/Ruthvicp/CS5590_BigDataProgramming/blob/master/Lab/Lab1/Source/MutualFriend/Input3/demo.txt>

**Evaluation**

Hadoop map reduce is very efficient in finding the common/mutual friends of two users when their list of friends are grouped together and filtered using reduce operation

**conclusion**

Representing the mutual friends problem using the map reduce diagram 