

3)

Spark sample program
val textFile = spark.textFile("hdfs:")
val count = textFile.flatMap(line => line.split(" "))
count.filter(f => f.contains('='))
count.map({ (person, friend) } => buildSortedKey(person, friend))
count.reduceByKey(output => { intersection(list1, list2, ..., listN) })
count.saveAsTextFile("hdfs://...")