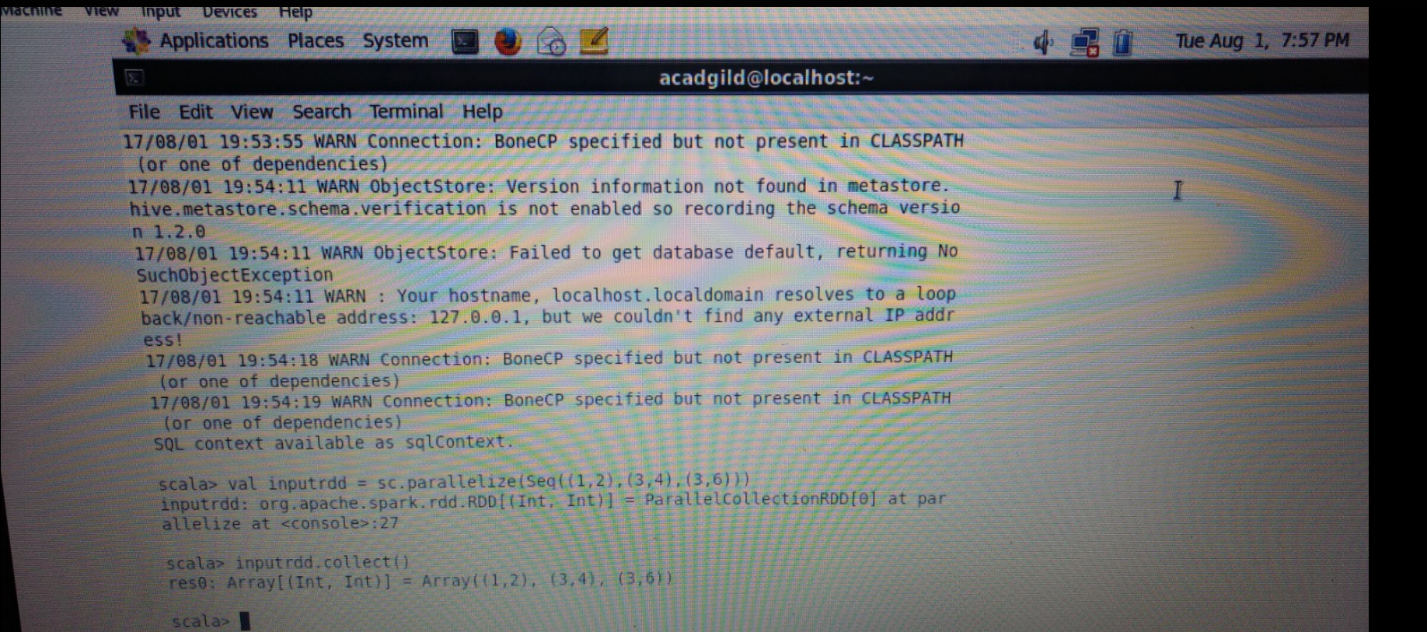
Start the spark shell

spark-shell

1. Create a simple PairRDD of (1, 2), (3, 4), (3, 6).

val inputrdd = sc.parallelize (Seq ((1, 2), (3, 4), (3, 6)))

inputrdd.collect ()



2. Transform an RDD of ("a","b","c","d","e") to PairRDD (a,0), (b,1), (c,2), (d,3), (e,4)

val inputA = sc.parallelize(Seq(“a”, “b”, “c”, “d”, “e”,))

val inputB = inputA.zipwithIndex

inputB.collect()

