Assignment 2 – Report

key-value pairs that were sent from the mappers to the reducers:

<u>step 1:</u>
  number of pairs – 212796585
  size (bytes) - 4781939956

<u>step2:</u>
  number of pairs – 4125612
  size (bytes) – 238193533

<u>step3:</u>
  number of pairs – 1181055
  size (bytes) – 69416660

<u>step4:</u>
  number of pairs – 334360
  size (bytes) - 12730612

Scalability:

1. Running with 7 instances: 9:44 minutes.
2. Running with 3 instances: 24:21 minutes.

Interesting word pairs: