***About the First Problem :***

I think there is a problem in the Pipeline.

I recommend you the following pipeline:

/////////////////////////

**Input:** sentences

**Mapper 1**: sentence -> (word, count)

For word in sentence:

Emit (word, 1)

**Reducer1:** list of (word, count) -> (word, total\_count)

(count1, count2) -> count1 + count2(total\_count)

Emit (word, total\_count)

**Mapper2:** list of (word, total\_count) -> (letter, (word, total\_count))

For (word, total\_count) in list:

Emit (word[0], (word, total\_count))

**Reducer2:** list of (letter, (word, total\_count)) -> (letter, list of (word, total\_count))

**Reducer3:** list of (letter, list of (word, total\_count)) -> list of (letter, words)

For letter, list in list of (letter, list of (word, total\_count)):

List = SortByValue(list of (word, total\_count))

Emit letter

For word in List:

Emit word