ELEN4020: Data Intensive Computing
Lab 3

School of Electrical & Information Engineering, University of the Witwatersrand, Private Bag 3, 2050, Johannesburg, South Africa

Elias Sepuru 1427726 Boikanyo Radiokana 1386807 Lloyd Patsika 1041888

I. INTRODUCTION

In this report the mapreduce framework is used to count the occurance of each word in a text file. On top of finding the occurance of each word, the top K, where K=10,20, occuring words are distinguished. The framework is also used to find the indices of words in a text file. The time taken to count the occurances of each word in a text file and the indices the words appear on are recorded. Mrs-MapReduce, a lightweight implementation of MapReduce is used.

II. DESIGN & IMPLEMENTATION

input : Two 2D matrices of sizes $n \times n$
output: A resultant C 2D matrix
for $i \leftarrow 1$ **to** n **do**
  for $j \leftarrow 1$ **to** n **do**
    for $k \leftarrow 1$ n **do**
      tempSum $\leftarrow$ tempSum + $A[i,k] \times B[j,k]$;
    end
    tempVector $\leftarrow$ tempSum;
  end
  C $\leftarrow$ tempVector;
end
return C

Algorithm 1: `rank2DTenosrMult(A,B)`: 2D Matrix Multplication