HIROSHIMA UNIVERSITY 広島大学

課題 1 MapReduce (Homework 1)

Big Data KA218001

ビッグデータ KA218001

Submission Information

Date	Student ID	Name
22/12/2024 (DD/MM/YYYY)	C240424	Yousef Ibrahim Gomaa Mahmoud Mabrouk

第 1 問の答え:

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}, \qquad v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

which is stored as input coordinates in this form:

$$(i, j, m_{ij})$$

$$\rightarrow (1,1,1), (1,2,2), (1,3,3) \ldots (4,4,16)$$

1. *Map Step:*

- Divide the idle mapper workers where each is assigned a chunk of matrix M.
- For each element m_{ij}, multiply by v_j.
- The output of the Map function is:

$$(key, value) = (i, m_{ij} \cdot v_j)$$

- By Iterating over j, the resultant key-value pairs are as follows: (Answer for 1-1)
  - (j=1, where 1≤i≤4)

$$(1, 1 \cdot 1) = (1, 1)$$
$$(2, 5 \cdot 1) = (2, 5)$$
$$(3, 9 \cdot 1) = (3, 9)$$
$$(4, 13 \cdot 1) = (4, 13)$$

  - (j=2)

$$(1, 2 \cdot 2) = (1,4)$$
$$(2, 6 \cdot 2) = (2, 12)$$
$$(3, 10 \cdot 2) = (3, 20)$$
$$(4, 14 \cdot 2) = (4, 28)$$

  - (j=3)

$$(1, 3 \cdot 3) = (1, 9)$$
$$(2, 7 \cdot 3) = (2, 21)$$
$$(3, 11 \cdot 3) = (3, 33)$$
$$(4, 15 \cdot 3) = (4, 45)$$

  - (j=4)

$$(1, 4 \cdot 4) = (1, 16)$$
$$(2, 8 \cdot 4) = (2, 32)$$
$$(3, 12 \cdot 4) = (3, 48)$$
$$(4, 16 \cdot 4) = (4, 64)$$

1.5. Grouping Step:

- *Group the values by key to form the input for the Reduce step.*
- *The output form of this step:*

$$(i, [m_{i1}, m_{i2}, \ldots, m_{in}])$$

2. Reduce Step

- Each reducer worker sums the values associated with key *i* to compute the final result for each X_i row. (which is commutative and associative).

$$X_i = \sum_{j=1}^{n} m_{ij} v_j$$

- The input for the Reduce function is: (Answer for 1-2)

$$(1, [1,4,9,16])$$
$$(2, [5,12,21,32])$$
$$(3, [9,20,33,48])$$
$$(4, [13,28,45,64])$$

- The output for Reduce function is:

$$X = \begin{bmatrix} 30 \\ 70 \\ 110 \\ 150 \end{bmatrix}$$

第 2 問の答え:

R		S	
A	B	B	C
0	1	0	1
1	2	1	2
2	3	2	3

1. *Map Step:*
   - Generate the key-value pairs (b, (R, a)) and (b, (S, c)) for each (a, b) and (b, c). (Cartesian Product)
   - Output of the Map function: (Answer for 2-1)

$$(1, (R, 0)), (0, (S, 1))$$
$$(2, (R, 1)), (1, (S, 2))$$
$$(3, (R, 2)), (2, (S, 3))$$

1.5. Grouping Step:

- Group the key-value pairs generated by the Mappers by key B, and pass the output to the Reducers: (Answer for 2-2)

$$(0, [(S, 1)])$$
$$(1, [(R, 0), (S, 2)])$$
$$(2, [(R, 1), (S, 3)])$$
$$(3, [(R, 2)])$$

2. *Reduce Step:*
    - For each pair (R, a), (S, c) produce the tuple (a, b, c) with a key (irrelevant)
        - At B = 0 → No values from R, and at B = 3 → No values from S.

$$B = 1:$$

$$(R, 0), (S, 2) \rightarrow (0,1,2)$$

$$B = 2:$$

$$(R, 1), (S, 3) \rightarrow (1,2,3)$$

    - The output is a list of all key-value pairs in the form of (a, b, c): (Answer for 2-3)
        - [(0,1,2), (1,2,3)]