

# 38| MapReduce

MapReduce
 Google
 GFS
 Bigtable
 PageRank

MapReduce

MapRedue

divide and conquer
 n

- 
- 
- 
- 
- 
- 
- 

$$1)/2$$

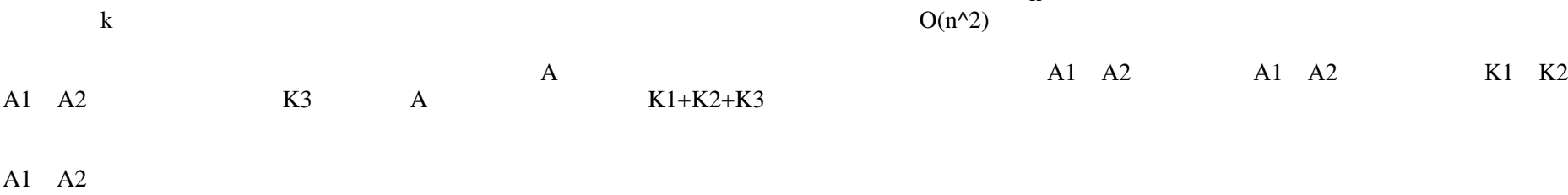
$$n$$

$$n(n-1)/2$$

$$0$$

$$0$$

$$n(n-$$





```
private int num = 0; //

public int count(int[] a, int n) {
    num = 0;
    mergeSortCounting(a, 0, n-1);
    return num;
}

private void mergeSortCounting(int[] a, int p, int r) {
    if (p >= r) return;
    int q = (p+r)/2;
    mergeSortCounting(a, p, q);
    mergeSortCounting(a, q+1, r);
    merge(a, p, q, r);
}

private void merge(int[] a, int p, int q, int r) {
    int i = p, j = q+1, k = 0;
    int[] tmp = new int[r-p+1];
    while (i<=q && j<=r) {
        if (a[i] <= a[j]) {
            tmp[k++] = a[i++];
        } else {
            num += (q-i+1); //      p-q      a[j]
            tmp[k++] = a[j++];
        }
    }
    while (i <= q) { //
        tmp[k++] = a[i++];
    }
    while (j <= r) { //
        tmp[k++] = a[j++];
    }
    for (i = 0; i <= r-p; ++i) { //      tmp      a
        a[p+i] = tmp[i];
    }
}
```

- $n$
- $n \times n$   $A \quad B \qquad C=A \times B$

' , p

A UdFYXi WY

10GB

10GB

2 3GB

10GB

10GB

1 100

10GB

101 200

GFS

10GB

## MapReduce

10GB

1T    10T    100T

## MapReduce

GFS  
Borg

Borg

GFS

Borg

## MapReduce

Digitized by Google

## MapReduce

## MapReduce

“

”

n

## MapReduce

Z\Y.###. #hYa d#[ YY\_h]a Y#

 $\#$ 

A UdFYXi WY

"\ha `Q&\$%#/#%`%.'\*.&)Q

“ ”

- Williamzhang 2018-12-20 02:22:11

num+= [3 ]

2018-12-21 02:07:54

- Yves 2018-12-19 04:10:10

1 num += (q - i + 1)  
i < r - p + 1 i < r - p

a[i] <= a[j]

2 while (i <= q)

num += (q - i + 1) 3

```
private void merge(int[] a, int p, int q, int r) {
    int i = p, j = q + 1, k = 0;
    int[] tmp = new int[r - p + 1];
    while (i <= q && j <= r) {
        if (a[i] <= a[j]) {
            tmp[k++] = a[i++];
        } else {
            num += (q - i + 1);
            tmp[k++] = a[j++];
        }
    }
    while (i <= q) {
        tmp[k++] = a[i++];
    }
    while (j <= r) {
        tmp[k++] = a[j++];
    }
    for (i = 0; i < r - p + 1; ++i) {
        a[p + i] = tmp[i];
    }
} [4 ]
```

2018-12-19 11:40:48

- 2018-12-19 14:27:02

[3 ]



2018-12-20 02:01:20

- 2018-12-20 02:03:38

6

v.  $2*2$   $8$   $7$   
n  $7/8 \quad ^{(\log n)} [2 \quad ]$

- 2019-01-11 03:39:48

10GB

1.  $2 \quad 3\text{GB}$  10GB  
2. buffer  $[1 \quad ]$

- Sharry 2018-12-20 03:38:35

, ! , (j-q-1) A tmp ,  
else ,  $[1 \quad ]$

- ALAN 2018-12-19 12:22:20

$[1 \quad ]$

- Smallfly 2018-12-19 00:47:19

$[1 \quad ]$

- 2018-12-19 00:19:56

if (a[i] <= a[j]) {

```
num += (j - q - 1);  
tmp[k++] = a[i++];  
} else {  
tmp[k++] = a[j++];  
}
```

[1 ]

2018-12-19 01:38:37

- qinggeouye 2019-01-14 17:19:07

12

<https://time.geekbang.org/column/article/41913>