

NAMA : ISEP LUTPI NUR
NPM : 2113191079
KELAS : INFORMATIKA A2 2019
MATA KULIAH : BIG DATA
PERTEMUAN : MINGGU 5 MAP REDUCE

1. Jelaskan apa yang dimaksud dengan Map Reduce.

MapReduce adalah kerangka kerja perangkat lunak dan model pemrograman yang digunakan untuk memproses data dalam jumlah besar. **Program MapReduce** bekerja dalam dua tahap, yaitu Map dan Reduce. **Program MapReduce** bekerja dalam dua tahap, yaitu Map dan Reduce.

2. Beri contoh dan jelaskan langkah-langkah proses Map Reduce.

- **Input Splits:**

Suatu input ke kerja MapReduce dibagi ke dalam potongan-potongan berukuran tetap bernama **input splits**. Input split merupakan potongan input yang akan diproses oleh suatu map tunggal.

- **Mapping**

Ini merupakan fase paling awal dalam eksekusi program map-reduce. Dalam fase ini data dalam setiap split dilewatkan ke fungsi mapping untuk menghasilkan nilai-nilai output. Dalam contoh kita, suatu kerja fase mapping adalah menghitung jumlah kemunculan setiap kata dari input splits (lebih rinci mengenai input-split dibahas di halaman selanjutnya) dan menyiapkan suatu daftar (list) berbentuk <kata, frekuensi>.

- **Shuffling**

Fase ini menerima output dari fase Mapping. Tugasnya adalah menggabungkan record-record yang relevan dari output fase Mapping. Dalam contoh di atas, kata yang sama dikumpulkan bersama dengan frekuensinya masing-masing.

- **Reducing**

Dalam fase ini, nilai output dari fase Shuffling diaggregasikan (digabungkan). Fase ini mengkombinasikan nilai-nilai dari fase Shuffling dan mengembalikan suatu nilai output tunggal. Singkatnya, fase ini merangkum dataset lengkap.