

2 Lezione

martedì 19 febbraio 2019

14:09

Importante trovare una soluzione parallela che spreca anche poca energia (top500, green500)

PAR/DISTR COMPUTER

TOP 500

|

SERVICES

|

PERSONAL DEVICES

laptop

mobile

IoT

Translate the book:

divisione pagine + unione pagine tradotte

↑
extra work

1 traduzione

2 persone

translate myself

↓
1/2

n persone
→ ideally I can go to $\frac{t}{n}$

Non conviene più quando il tempo che serve per distribuire le pagine e rimetterle insieme è maggiore del tempo necessario per fare le traduzioni.

Esempio: contare le monete in tasca

① ognuno conta le sue monete

② we sum up (il conteggio lo facciamo considerando che non è importante l'ordine con cui sommo).

Rispetto all'esempio della traduzione → qua abbiamo qui l'input, non devono fare il task da eseguire.

GOOGLE MAP REDUCE

$f: \alpha \rightarrow \beta$ map function
 $\langle k, v \rangle$

$\oplus \langle k, v \rangle^* \rightarrow \langle k, v \rangle^* \quad v \oplus v \rightarrow v$

$\langle "k", 1 \rangle, \langle "k", 2 \rangle \Rightarrow \langle "k", 3 \rangle$

Mapreduce usato anche per conteggio delle parole in un testo.

Map phase → shuffle phase → reduce phase

for all the tuples group the tuples with the same value:

Site A

⊙
⊙
⊙

$\langle a, 1 \rangle, \langle a, 1 \rangle, \langle b, 1 \rangle, \langle c, 1 \rangle$

I can do a local reduction

$\langle a, 2 \rangle, \langle b, 1 \rangle, \langle c, 1 \rangle$

Site B

⊙
⊙
⊙

$\langle a, 1 \rangle, \langle c, 1 \rangle$

$\langle a, 1 \rangle, \langle c, 1 \rangle$

Site C

⊙
⊙
⊙

$\langle b, 1 \rangle, \langle b, 1 \rangle, \langle b, 1 \rangle, \langle a, 1 \rangle$

$\langle b, 3 \rangle, \langle a, 1 \rangle$

$\langle a, 1 \rangle$ ①

$\langle b, 1 \rangle$ ③

$\langle c, 1 \rangle$ ②

→ Merge

No centralization point:

Site A $\langle a, 2 \rangle, \langle b, 1 \rangle, \langle c, 1 \rangle$ → Site X

Site B $\langle a, 1 \rangle, \langle c, 1 \rangle$ → Site Y

Site C $\langle a, 1 \rangle, \langle b, 3 \rangle$

I can use hash.

I'm moving the work from one single server to local work (computation of hash)