

Introducció al BigData

Problemes-1

Ll. Gesa

Arquitectura i tecnologies software

UAB, 2017

Problema 1

Com comptar lletres/paraules

Com comptaríeu quantes lletres/paraules hi ha en un arxiu?

Problema 1

Com comptar lletres/paraules

Com comptaríeu quantes lletres/paraules hi ha en un arxiu?

Com comptar lletres/paraules

I en un arxiu de 2Gb? I en 10 arxius de 2Gb cadascú?

Problema 1

Com comptar lletres/paraules

Com comptaríeu quantes lletres/paraules hi ha en un arxiu?

Com comptar lletres/paraules

I en un arxiu de 2Gb? I en 10 arxius de 2Gb cadascú?

MapReduce

La metodologia MapReduce permet aplicar el concepte de 'divide & conquer' de forma neta i clara

Problema 1

Com comptar lletres/paraules

Com comptaríeu quantes lletres/paraules hi ha en un arxiu?

Com comptar lletres/paraules

I en un arxiu de 2Gb? I en 10 arxius de 2Gb cadascú?

MapReduce

La metodologia MapReduce permet aplicar el concepte de 'divide & conquer' de forma neta i clara

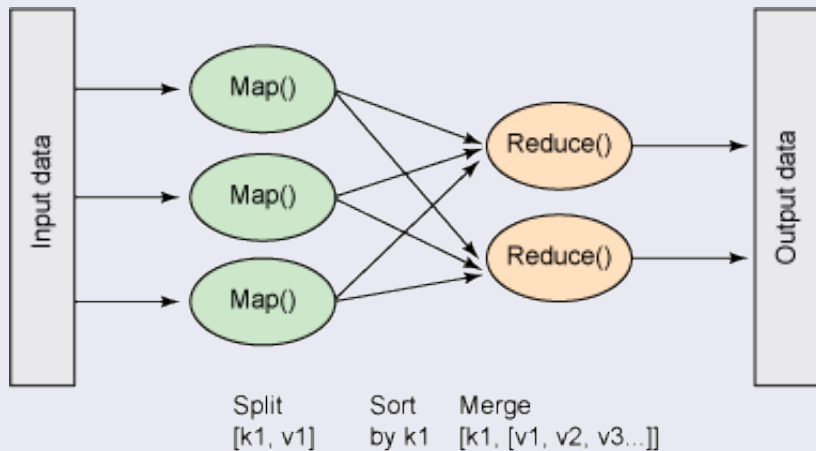
Sintaxi

$\text{Map}(k1, v1) \rightarrow \text{list}(k2, v2)$

$\text{Reduce}(k2, \text{list}(v2)) \rightarrow \text{list}(v3)$

Problema 1

MapReduce diagrama basic



Problema 1

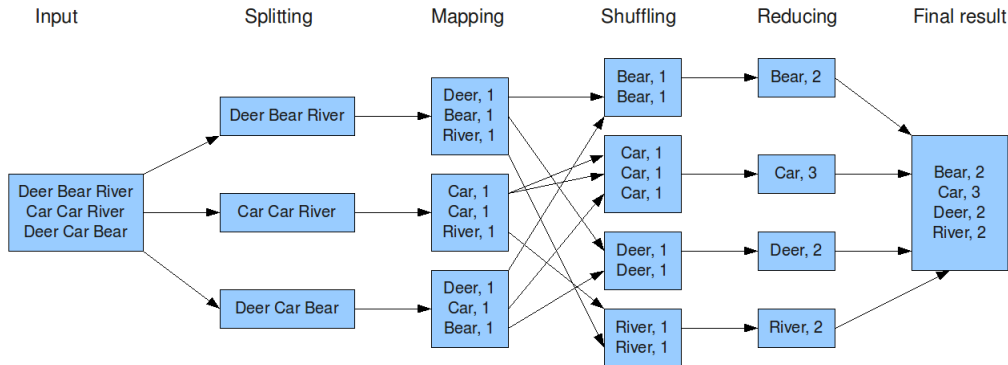
Comptem paraules amb pseudocodi

```
function map(String name, String document):  
    // name: document name  
    // document: document contents  
    for each word w in document:  
        emit (w, 1)  
  
function reduce(String word, Iterator partialCounts):  
    // word: a word  
    // partialCounts: a list of aggregated partial counts  
    sum = 0  
    for each pc in partialCounts:  
        sum += pc  
    emit (word, sum)
```

Problema 1

MapReduce diagrama estès

The overall MapReduce word count process



Tamany de les paraules

Quantes paraules (i quines) hi ha de 2 caràcters, de 3, de... en un arxiu?

Problema 3

Amics comuns

Facebook ens contracta per implantar una rutina que actualitzi unes taules que ens diguin directament quants amics comuns tenim amb algun contacte nostre en el moment d'entrar al seu perfil.

Problema 3

Amics comuns

Facebook ens contracta per implantar una rutina que actualitzi unes taules que ens diguin directament quants amics comuns tenim amb algun contacte nostre en el moment d'entrar al seu perfil.

Amics comuns

A — > B C D

B — > A C D E

C — > A B D E

D — > A B C E

E — > B C D

Problema 3

Amics comuns

Facebook ens contracta per implantar una rutina que actualitzi unes taules que ens diguin directament quants amics comuns tenim amb algun contacte nostre en el moment d'entrar al seu perfil.

Amics comuns

A — > B C D
B — > A C D E
C — > A B D E
D — > A B C E
E — > B C D

Solució

<http://stevekrenzel.com/finding-friends-with-mapreduce>

Multiplicació de matrius

Multiplicar Una matriu $N \times M$ per una altra $M \times N$.

Problema 4: GPU

Problema 4: GPU

Graphic Process Unit



CPU vs GPU

