

EXAMEN DE CBDE

12 de Gener del 2016

Instruccions: Respon cada pregunta al full corresponent. L'examen dura 1h 30'

Nom i Cognoms:

Pregunta 1. [4p]

MongoDB ha decidit fer servir el motor *WiredTiger* com a motor per defecte a partir de la versió 3.2. *WiredTiger* implementa un catàleg global basat en un arbre B+ distribuït (LSM), tal i com fa HBase. **Justifica una avantatge i un inconvenient** de l'LSM enfront del consistent hashing que implementava a les versions anteriors.

.....
.....
.....
.....
.....

A més, *WiredTiger* permet elegir entre emmagatzemar les dades *row* o *column-oriented*. **Justifica una avantatge i un inconvenient** de cada tipus d'emmagatzematge i exemplifica'ls amb un cas real que justifiqués una o altra elecció.

.....
.....
.....
.....
.....

A part, *WiredTiger* permet especificar el tipus de compressió de dades (considera les següents opcions: *run length-encoding*, *Lempel-Ziv*, *Huffman*, *cap*) i el format dels fitxers (*variable* o *fixed-record size*). Suposa ara que has de crear el magatzem de dades d'una empresa i decideixes emprar MongoDB. Justifica, de les següents opcions que et permet MongoDB, **quines elegiries i perquè** (fes les assumpcions que creguis necessàries a l'hora de fer la teva elecció):

- Motor: *WiredTiger* (basat en LSM), *MMAP V1* (basat en Consistent-Hashing)
- Format de fragmentació: *row-oriented* o *column-oriented* (només amb *WiredTiger*)
- Compressió de dades: *run length-encoding* (només amb *WiredTiger*), *Lempel-Ziv*, *Huffman*, *cap*
- Format dels fitxers: *variable* o *fixed-record size*

.....
.....
.....

.....

.....

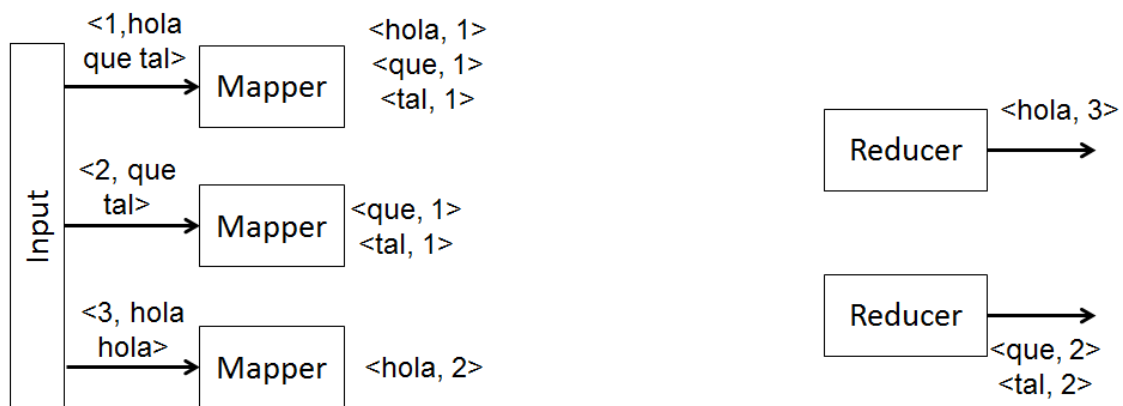
.....

.....

.....

Pregunta 2. [4p]

Completa la fase intermitja que falta entre els mappers i els reducers dibuixats per tal que aquest MapReduce job sigui coherent respecte a la teoria explicada en classe. Cal deixar ben clar com es porten a terme cadascun dels passos de MapReduce. Empra l'espai sota de la figura per fer qualsevol justificació en aquest aspecte que no puguis fer a la figura.



.....

.....

.....

.....

.....

Considera el següent codi pel Map:

```
public void map(LongWritable key, Text value) {  
    String line = value.toString();  
    StringTokenizer tokenizer = new StringTokenizer(line);  
    while (tokenizer.hasMoreTokens()) {  
        write(new Text(tokenizer.nextToken()), new IntWritable(1));  
    }  
}
```

Amb aquest codi, com és possible que el tercer Map emeti un <key, value> com el de la figura?

.....

.....

.....

Nom i Cognoms:

Pregunta 3. [2p]

Anomena dos dels tres esquemes que hi ha al catàleg global d'una base de dades distribuïda

.....
.....
.....

Quin factor, que no sigui la data de l'última actualització, influeix en la mesura de qualitat de '*Timeliness*'?

.....
.....
.....