

# EXAMEN DE CBDE

20 de Novembre del 2013

*Instruccions: Respon cada pregunta al full corresponent.*

*L'examen dura 1 hora.*

Nom i Cognoms: .....

**PREGUNTA 1.** Justifica les següents afirmacions.

1.1. MongoDB implementa *journaling*. Segons la definició del manual de MongoDB, això implica que implementa un sistema de REDO logging WAL (Write-Ahead Logging).

Segons aquesta definició, explica breument què necessita guardar MongoDB al dietari (log):

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

Què significa que el sistema de recuperació de MongoDB implementi el protocol WAL?

.....  
.....

1.2. Justifica breument si el framework MapReduce implementa una estratègia data shipping, query shipping o una versió híbrida de les dues.

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

1.3. Justifica breument quina és la principal diferència entre el Aggregation Framework de MongoDB i MapReduce respecte al paral·lisme que proporciona cadascun d'ells.

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

**PREGUNTA 2.** Considera els dos sistemes que es mostren a continuació:

Sistema 1: Base de dades centralitzada

- És una base de dades centralitzada amb un únic node
- La base de dades ocupa 12TB. Per simplicitat, assumeix que conté una única taula T amb totes les dades i n'hi ha 100.000.000 de tuples
- La latència de llegir de disc és de 5ms
- L'ample de banda màxim que pot aconseguir el disc és de 100MB/s

Sistema 2: Base de dades distribuïda

- Consta de 3 nodes connectats per una LAN. Qualsevol d'ells pot llençar queries
- La base de dades ocupa 12TB. Pots assumir que la taula T s'ha distribuït de forma uniforme entre els nodes mitjançant una fragmentació horitzontal. No hi ha replicació
- La latència de llegir de disc és de 5ms i la de la xarxa 1 ms
- L'ample de banda màxim que pot aconseguir el disc és de 100MB/s
- Els nodes estan connectats a través d'una LAN amb ample de banda màxim de 10MB/s

Suposa que l'única *query* del sistema és:

```
SELECT SUM(a) FROM T, on a és un atribut de T
```

No hi ha índexs ni cap altra estructura definida en el sistema.

En el millor cas, quant trigarà (en segons) en fer un accés seqüencial de T en el sistema 1?

Latència .....

Lectura seqüencial .....

Total .....

I en el sistema 2 en el millor cas?

Latència .....

Lectura seqüencial .....

Total .....

Suposa ara que la única query del sistema fos:

```
SELECT * FROM T WHERE pk = 1, on pk és la clau primària de T i té un B+ associat (pots suposar que el B+ està a memòria)
```

Què canviaria respecte a la pregunta anterior?

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