

# Activitat EBH

**Emmagatzematge, *backup* i *housing***

**Bonilla Pérez, Laia  
Casado Esteras, Diego**

**Escenari CRB2**

**Data: 29/09/2023**

# 1.-Descripció bàsica

Copy & Paste del full de càlcul "Resum"

AQUEST APARTAT NOMÉS PEL LLIURAMENT FINAL

**IMPORTANT: Pels apartats següents no poseu només el número, justifiqueu el perquè. És més important el perquè que el número en si.**

## 2.-Anàlisi de necessitats

### 2.1- Número de GB a emmagatzemar (en cru).

Tenim 10 grups de recerca, i cada grup té un espai on pot emmagatzemar 30 TB.  
Per tant,

$$10 \text{ grups} * \frac{30 \text{ TB}}{1 \text{ grup}} * \frac{1024 \text{ GB}}{1 \text{ TB}} = 307200 \text{ GB} = \mathbf{3,072 * 10^5 \text{ GB}}$$

### 2.2- Velocitat requerida del sistema de disc (IOPS).

Cada rack intercanvia 8 MBps de dades amb el disc, i tenim 18 racks.  
Per tant,

$$8 \text{ MBps} * 18 \text{ racks} = 144 \text{ MBps} = 147456 \text{ KBps}$$

Els discs accedeixen en blocs de 4KB, per tant  $\frac{147456 \text{ KB / s}}{4 \text{ KB / operació IO}} = \mathbf{36864 \text{ IOPS}}$

### 2.3- Tràfic amb el client (entre servers i de server a switch de connexió a xarxa):

Cada rack intercanvia 4 MBps de dades amb l'exterior, i tenim 18 racks.  
Per tant,

$$18 \text{ racks} * \frac{4 \text{ MBps}}{1 \text{ rack}} * \frac{8 \text{ Mbps}}{1 \text{ MBpS}} = \mathbf{576 \text{ Mbps}}$$

### 2.4- Tràfic amb el disc:

Cada rack intercanvia 8 MBps de dades amb el sistema d'emmagatzematge, i tenim 18 racks.  
Per tant,

$$18 \text{ racks} * \frac{8 \text{ MBps}}{1 \text{ rack}} * \frac{8 \text{ Mbps}}{1 \text{ MBpS}} = \mathbf{1152 \text{ Mbps}}$$

### 2.5- Pressió sobre la xarxa (ample de banda mínim necessito per servir el tràfic de client i disc). M'arriba?:

El tràfic total és la suma del tràfic amb el client i el tràfic amb el disc.

Per tant,

$$\text{Tràfic total} = 576 \text{ Mbps} + 1152 \text{ Mbps} = 1728 \text{ Mbps} = \mathbf{1,688 \text{ Gbps}}$$

Com que tenim una xarxa de 2 Gbps, no anem sobrats però ens arriba.

### 3.-Decisions preses

#### 3.1- Descripció dels elements d'emmagatzematge escollits, en funció de les necessitats. Quants tipus de cabines? (i perquè), RAID escollit a cadascuna d'elles. Nombre de cabines de cada tipus

En els detalls del nostre entorn ens diu que el 70% de les dades emmagatzemades no es consulten habitualment i que els servidors sabem quines són. Per tant, tenim dos tipus de dades: aquest 70% que no es consulten amb gaire freqüència i l'altre 30% de les dades que es consulten amb regularitat. Per això hem decidit organitzar l'emmagatzematge per grups. Cada grup de recerca disposarà de dues cabines: una on estaran emmagatzemades el primer tipus de dades i una altra on estaran les del segon tipus.

Com ens diu l'enunciat, cada grup pot emmagatzemar 30TB de dades. Per tant, el 70% d'aquestes dades aniran a una cabina (21504 GB), i l'altre 30% a l'altra cabina (9216 GB).

#### **DISSENY CABINES TIPUS A** (contenen el 70% de dades consultades amb poca freqüència)

Per a aquestes cabines hem decidit utilitzar discs magnètics (HDD), ja que no es necessària una resposta ràpida (en l'enunciat ens diuen que els clients accepten que aquestes dades puguin trigar més temps en ser accedides). També ens hem decantat per aquest tipus de disc perquè creiem que la quantitat d'espai que ofereixen en relació amb el preu és més adient.

Entre les 5 opcions de discs HDD que se'ns ofereixen hem descartat l'opció 1, ja que creiem que és convenient apostar pels discs Enterprise respecte als Consumer en perquè ens ofereixen una major fiabilitat, tenen un període més llarg de garantia i proporcionen un major MTBF (Mean Time Between Failure) i recuperació d'errors. Ens interessa molt la fiabilitat ja que estarà emmagatzemant un gran conjunt de dades dels clients i seria bo que tingués una bona durabilitat a llarg termini. D'entre les opcions Enterprise hem descartat les opcions 2 i 3 perquè la primera tant com la segona requerim d'una quantitat elevada de discs per poder satisfer els IOPS corresponents i alhora obtenim una capacitat bastant desorbitada respecte l'emmagatzematge a priori necessari. Les opcions 4 i 5 creiem que ofereixen una millor relació en quant a IOPS i preu, envers a la reducció de capacitat proporcionada. Finalment, hem optat per l'opció 5 perquè creiem que tot i que a l'escenari no siguin necessaris una gran quantitat d'operacions d'entrada i sortida dóna una bona resolució a les peticions que haurà de proporcionar el sistema. Tot i que no alberguin una quantitat molt gran ofereixen un bon consum i un preu considerable per tal de poder escalar d'una forma econòmica i en base a les necessitats de l'escenari.

Pel que fa al sistema RAID, donat que no s'emmagatzemen dades sensibles o de vital importància, creiem que un sistema 61 i 6 no són necessaris. Tampoc creiem que un sistema 51 sigui adient per l'escenari ja que augmenta considerablement el nombre de

discos al tenir mirroring i per tant succeeix el mateix amb el preu. A més, tenim discs Enterprise i el Backup en cas que passés alguna cosa, així que informació no es perdria. Trigaríem una mica més a recuperar-la, però com gairebé no s'accedeixen, no és cap problema.

Tampoc ens convé un 10, que està més a prop en seguretat al 5 que al 51, però el preu està més a prop del 51 que del 5. No creiem que valgui la pena, mai.

Degut a la quantitat de dades necessitem 9 discs (21TB entre 2,4TB que permet cada disc de tipus 5). Hem decidit escollir la cabina 2, ja que no necessitem JBOD, no necessitem 36 badies (però sí més de 12), i no necessitem ssd support. A més, com escollim la 2 en comptes de la 1, tenim més flexibilitat per tal de créixer. Aprofitant aquest espai extra, dividirem els discs en 5 grups de 4. Cada grup serà un RAID5, format per un disc de paritat i 4 usables. A més, comprarem un spare disk per cada grup, per compensar que tenim RAID 5 en comptes de 51.

### **DISSENY CABINES TIPUS B (contenen el 30% de dades consultades amb regularitat)**

Per a aquestes cabines hem decidit utilitzar discs d'estat sòlid (SSD), perquè volem que la resposta sigui ràpida ja que aquestes dades es consulten en regularitat. De nou hem decidit apostar per els discos Enterprise respecte al Consumer per la fiabilitat, tenen un període més llarg de garantia i proporcionen un major MTBF (Mean Time Between Failure) i recuperació d'errors. Per tant les opcions 6 i 7 queden descartades. De les opcions candidates hem fet la escollit l'opció 8, ja que no necessitem discs grans, donat que només hem de guardar 9GB per grup, per velocitat ens sobra, i són molt més econòmics que el 9 i 10.

En quant al sistema RAID a escollir, com que el sistema necessita una resposta més ràpida i a més no ens podem arriscar a tenir parat el sistema per tal de recuperar del backup, hem decidit implementar un RAID 51. Així doncs, necessitem 12 discs per grup.

Pel que fa al model de cabina, al tractar amb discos SSD pensem que no fa falta incorporar la opció de SSD perquè creiem que els discos pels que està format tenen la suficient rapidesa per poder dur a terme una bona gestió. No creiem que sigui necessari un spare disk, ja que tenim un RAID 51 i, en cas de caure un disc, està la part redundada per recuperar tota la informació. Al no tenir un spare disk, podem tenir aquests 12 discs guardats en una cabina 2 (i no 1 perquè no permet RAID 51), i deixar les 12 badies restants disponibles per tal d'ampliar el sistema en un futur amb dos grups nous de 5+1.

En resum, per al 30% de les dades que es consulten en freqüència comprarem 1 cabina de model 2 per cada grup. Per tant, un total de 10 cabines.

Així doncs, tindrem 20 cabines de tipus 2, amb 12 discs de tipus 8 i 24 discs de tipus 5 per cabina.

### **3.2- Es justifica la necessitat d'un SAN? Si la resposta és si, raonar si el cost és assumible o no, i cas de no ser-ho calcular l'impacte sobre el rendiment del CPD**

De moment no necessitem un SAN, ja que la nostra infraestructura pot amb el que tenim, gràcies, en part, al sistema de backup contractat, i no hi ha previsions de créixer gaire. Tot i així, queda molt poc marge disponible. Donat això, podríem tenir problemes en un futur, en cas de necessitar més discos, cosa probable.

El cost, no és assumible. Hauria de ser prioritària la inversió, ja que és l'únic punt del sistema que està apurat.

### 3.3.- Posem un *mirror*?

Donat el percentatge de fallada que tenen els nostres discos Enterprise, les mesures de seguretat ja implementades, la relativament baixa importància de la feina que realitzem (i, per tant, relativament poca necessitat d'aprofitar el màxim temps possible, hem considerat que un *mirror* és totalment prescindible.

Pel sistema lent,  $2,84\% * 2,84\% * 0,3 * 13$  dona 0,3146% de probabilitats de que falli un disc dels que tenim en marxa ara mateix per a cada grup durant un any.

Pel sistema ràpid, amb un sistema 51 amb 12 discos que tenen un 0,45% de probabilitats de fallada, arribem a un resultat de  $4,92 * 10^{-7}\%$  de probabilitats que falli un disc durant l'any a un grup de treball. No necessitem un *mirror*.

### 3.4- Empresa de *housing* escollida i perquè (relació entre el que ofereix, el que necessito i el que costa)

Clarament escollirem l'opció 2. La 1 té un problema fonamental, i és l'obligatorietat de contractar serveis en pacos tan grans. Nosaltres tenim ara mateix 18 racks amb un total de 756Us, i necessitem per 770Us, per tant només necessitem un rack més, i ens sobra. Així doncs, la 1 no és adequada. La tres tampoc ho és, ja que està pensada per serveis molt més crítics i amb més probabilitats de ser atacats, i no és el nostre cas. Per tant, la dos és la més adient. Permet backup on-site i el downtime està dintre de les nostres necessitats.

### 3.5- Posem monitorització?

Al utilitzar un sistema de spare disk en ambdós sistemes és necessària la monitorització per poder preveure la fallada del disc i per tal d'evitar la pèrdua de dades i activar els discos perquè puguin fer una còpia del que es preveu que falli.

### 3.6- Opció de backup?

Es farà en cintes, per dos motius. Primer perquè els altres són més cars, i no ens sobren els calers. Segon, perquè la nostra limitació és la xarxa, principalment. Per tant, ens dona igual que ens enviïn la informació ràpid o lenta, si no podem gravar-la a més velocitat.

Donada la baixa probabilitat de patir un atac, perquè no som gent que destaquem, i la baixa probabilitat de que un disc caigui, hem decidit fer un backup cada 10 dies. Guardaríem el que acabem de fer, l'anterior, el trimestral i l'annual, seguint les recomanacions.

### 3.7- Tràfic amb l'exterior afegit pel sistema de *backup/mirror* escollit. Quin *bandwidth* caldria?

Com que la capacitat de la xarxa LAN provoca un coll d'ampolla, no té sentit escollir un *bandwidth* de més de 2Gbps perquè un cop les dades arriben a la nostra xarxa LAN només es poden transmetre dades a, com a màxim, 2Gbps. Així que escollirem un *bandwidth* de 2 línies agregades de 1 Gbps cadascuna.

## 4.-Recomanacions als inversors

### 4.1.- Anàlisi de Riscos (*Risk Analysis*)

Quines desgràcies poden passar i com les hem cobert?

Al menys s'han de cobrir els següents casos:

- **Hi ha pèrdua d'un fitxer (per error o corrupció). De quan puc recuperar versions?**

Si perdem un fitxer, la darrera copia del backup serà de fa dues setmanes. En cas que la corrupció fos anterior, ens hauríem de remuntar un més abans del moment que ens vam adonar. Igualment, al tenir un sistema amb 51 a les dades més usades, la possibilitat de patir un problema per corrupció és mínima.

- **Es trenca un disc (es perden dades? quan trigo en recuperar-me? el negoci s'ha d'aturar?)**

En cas que es trenqui un disc, en cap cas suposaria un problema. Al sistema lent, tenim tolerància a fallades mitjançant un RAID 5 + un spare disk per cada grup de discos, de manera que estem coberts. Pel cas del ràpid, que requereix més seguretat, tenim un RAID 51, que no només té el disc de paritat, sinó que té tot el sistema redundat.

- **Puc tenir problemes de servei si falla algun disc?**

Haurien de fallar mínim dos discs en el sistema lent, i a més fer-ho de forma sobtada, per haver de tirar de backup. En el cas del ràpid, serà encara més difícil haver de tirar de backup.

- **Cau la línia elèctrica. Què passa?**

No gaire. Tenim un sistema amb SAI, usant diesel, per tant estem coberts. I ja hem tingut en compte que el downtime que ens ofereixen és molt bo.

- **Cau una línia de xarxa. Què passa?**

Tenim tres línies, per tant si cau una no passa res. Seria diferent si caigués el proveïdor sencer, ja que només tenim un, al no tenir contractat Mordor.

- **En cas de pèrdua o detecció de corrupció de dades no ens podem permetre seguir treballant fins que recuperem les dades correctes. Calculeu temps i costos de recuperació en cas de**

- **Pèrdua/ corrupció d'un 1% de les dades**

Com és una petita part de les dades, comptem que es pot seguir treballant sense problema.

Així doncs, comptem que la xarxa està utilitzada, igual que la resta, però la xarxa és, i ha sigut sempre, el nostre problema. Parlem, doncs, de 0,08Gbps disponibles, que són 0,01GBps disponibles. Un 1% de la informació emmagatzemada són 3TB. 3TB → 3072GB.

$3072\text{GB} \cdot (100\text{s}/1\text{GB}) = 307200\text{s} \rightarrow 85,33\text{h} \rightarrow 3,55\text{ dies}$ . És a dir, tres dies i mig.

Aquesta quantitat de temps es podria veure rebaixada en cas que a la nit no hi treballés ningú.

- **Pèrdua/ corrupció de la totalitat de les dades**

Si perdem totes les dades, ningú pot treballar, així que tenim totes les eines a la nostra disposició.

De nou, la nostra limitació és la xarxa, ja que els nostres discos lents poden escriure a una velocitat de 10Gbps (5210IOPS per disc, 4KB per IOP, 4 operacions d'escriptura per dada real escrita degut al sistema 51, per 240 discos), i la ràpida parlem de 393,6Gbps (205K IOPS per disc, 4KB per IOP, 2 operacions d'escriptura per dada real degut al raid 5, 120 discos). Així doncs, tenim 2Gbps per 300TB d'informació, el que ens deixa amb

#### **4.2.- Anàlisi de l'impacte al negoci (*Business Impact Analysis*)**

En funció de l'anàlisi de riscos anterior i del que costa estar amb la màquina aturada o no donar el servei complet, calcular quant perdo en diners per tenir-lo aturat i quan em costaria evitar aquesta situació.

##### Caiguda de la xarxa de dades:

Tenir aturada la xarxa de dades significa no poder treballar. Parlem d'aturades de 0-1h cada 18 mesos, 1-3h cada 3 anys i 3-9h cada 6 anys. Això ens deixa 1,69 hores l'any de caiguda, el que significa 33800 euros l'any, que són 169.000 euros per 5 anys. Per canviar-ho, hauríem de contractar mordor, amb una segona companyia de xarxa, i dues línies per cada companyia, la qual cosa suposaria 332.000 euros adients per 5 anys. No val la pena, si no és per comoditat i imatge.

##### Fallada de disc

En cas de fallada del disc no perdríem res, ja que tenim 51 i 5 respectivament en els lents i ràpids. El pitjor que passaria seria tenir la meitat de velocitat al sistema ràpid durant 4 hores, però els IOPS disponibles són molt superiors a la capacitat de la xarxa, motiu pel qual no es notaria la parada.

#### **4.3.- Creixement**

**Si creix el nombre de clients/ màquines/ dades (depèn de l'escenari), hem d'estar preparats.**

Quin creixement (en nombre de clients, etc...) podem assumir sense canviar el sistema (sobreprovisionament)? Quin és el recurs que s'esgota abans? Feu un informe de les implicacions que suposaria un increment d'un 20% en el volum de negoci (tot, clients, dades, ...)

Hauríem de contractar un SAN sí o sí, o ampliar la xarxa LAN. El SAN més econòmic ens serviria, però són 900000€ més. També hauríem de comprar més discos ràpids, ja que de lents tenim pel doble del que usem ara mateix. Estaríem parlant de 3 discos ràpids + 1 pel raid 5, 4 discos. Parlem de 14.880 euros en total. Com ens han sobrat 18000 euros, tenim suficient per créixer, exceptuant el tema del SAN.

#### **4.4.- Inversions més urgents**

Donat el CPD resultant és possible que no haguem escollit la millor opció per manca de diners. El CPD no és nostre, nosaltres només ho dissenyem, així que al final s'hauria de fer un informe als que posen els diners de en què valdria la pena invertir per millorar rendiment, seguretat o...

Com comentàvem unes línies més amunt, la inversió més urgent, amb molta diferència, és el SAN. Només cal que es guardin unes poques dades més o que creixi el nombre de grups d'investigació, per a que el nostre sistema sigui completament insuficient i comenci a anar molt

més lent del que hauria de ser, veient aquesta lentitud reflectida en el rendiment de l'equip i en els ànims.

La segona inversió necessària seria reduir el temps entre backups a una setmana, i guardar un backup més, per tal de garantir la recuperació de dades en cas d'haver comés un error o no haver detectat adequadament la corrupció d'un fitxer.

La tercera, però igualment recomanable, seria moure al housing 3 la nostra infraestructura. És recomanable principalment per la segona línia elèctrica, ja que el downtime de la que tenim contractada, tot i complir les nostres necessitats, està molt apurada, i perdem força diners amb ella, a més de possible reputació.