

# Activitat EBH

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

**Chapagain, Bibek  
Haris, Muhammad**

**Escenari 08**

**Data: 13/10/2023**

# 1.-Descripció bàsica

**TAULA 1: ESCENARI ORIGINAL: EXTRET DE L'ENUNCIAT. OMPLIU EL QUE HI HA EN GRIS.**

Nombre de Us	892U
Alçada Rack (en Us)	42U
Consum	1812,8kW
Sobreprovisionament d'electricitat	15%
Nombre de servidors	200
Diners Totals	€21.000.000,00
Diners gastats	€15.500.000,00

**taula 2: Elements que escolliu vosaltres**

<b>Elements de mirror i backup</b>	
GB a emmagatzemar al backup	240000
Dies entre 2 backups	1
Còpies senceres a mantenir	10
Opció Backup (1=M-A; 2=MS3; 3=Cintes)	3
Opció Mirror (0=NO; 1=SI)	0
Sistema de backup on-site? (0=N=; 1=SI)	1
<b>Elements de housing</b>	
Opció escollida (1:MOCOSA, 2: CPDs Céspedes, 3: Mordor)	3
Gestió local de <i>backup</i> ? (0=No, 1=Si)	1
Monitorització? (0=NO; 1=SI)	0

Bandwidth provider	
Tipus de línia (1:10Mbps; 2:100Mbps; 3:1Gbps; 4:10Gbps; 5:100Gbps)	4
Número de línies agregades	2
Segon proveïdor? (0=NO, 1=SI)	1
SAN? (0=no, 1=8Gbps, 2=16Gbps, 3=32Gbps, 4=64Gbps, 5=128Gbps)	0
Cabina de discos	
Opció Disc principal (Entre 1 i 10)	9
Nombre de discos a comprar	144
Opció cabina de discos (Entre 1 i 6)	4
Nombre de Cabines	4
Cabina de discos 2 (cas de fer servir dos tipus)	
Opció Disc (Entre 1 i 10)	3
Nombre de discos a comprar	24
Opció cabina de discos (Entre 1 i 6)	2
Nombre de Cabines	1

TAULA 3: OPEX	anual	cinc anys
Consum energètic (hardware només)	€328.381,91	€1.641.909,53
Empresa de Housing escollida	Mordor	
Cost Housing (inclou electricitat addicional)	€427.257,29	€2.136.286,43
Off-site: empresa escollida	Take the tapes and run	
Cost mirror	€0,00	€0,00
Cost backup	€117.600,00	€588.000,00
Cost Bandwidth provider	€21.168,00	€105.840,00

TAULA 4: CAPEX	Cost
Diners gastats en servers, xarxa, etc	€15.500.000,00
SAN	€0,00
Sistema emmagatzematge	€255.480,00

TAULA 5: AJUST AL PRESSUPOST	
Opex a 5 anys, total	€4.472.035,96
Capex a 5 anys, total	€15.755.480,00
Despeses totals a 5 anys	€20.227.515,96
Diferència respecte al pressupost	€772.484,04

## 2.-Anàlisi de necessitats

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

Cada servidor guarda un màxim 1TB de dades al disc centralitzat i tenim 200 servidors en total. A més, es guarden dades històriques que ocupen 40TB. Llavors en total a emmagatzemar en cru seria  **$200 * 1TB + 40TB = 240 TB$**

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

Tenim 200 servidors de càlcul. Cada servidor té un tràfic en mitja de 1 Mbps de dades. El 75% entre servidors (LAN) i el 25% restant al disc. De tant en quant n'hi ha pics de treball de 400 Mbps de tràfic.

Les IOPS considerant el tràfic en mitja serien:

**$200 * ( (0.25 * 1 \text{ Mbps} * 1000) / 8 ) = 6250 \text{ kB/s}$**  i com que cada operació al disc és de 4KB, tenim  **$6250 / 4 = 1563 \text{ IOPS}$**

Les IOPS considerant el tràfic dels pics serien:

**$200 * ( (0.25 * 400 \text{ Mbps} * 1000) / 8 ) = 2500000 \text{ kB/s}$**  i com que cada operació al disc és de 4KB, tenim  **$2500000 / 4 = 625000 \text{ IOPS}$**

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

Considerant un tràfic de 1 Mbps, el tràfic amb el client seria de  **$200 * 0.75 * 1 \text{ Mbps} = 150 \text{ Mbps}$**

Considerant un tràfic de 400 Mbps, el tràfic amb el client seria de  **$200 * 0.75 * 400 \text{ Mbps} = 60000 \text{ Mbps}$**

## 2.4- Tràfic amb el disc:

Considerant un tràfic de 1 Mbps, el tràfic amb el disc seria de  $200 * 0.25 * 1 \text{ Mbps} = 50 \text{ Mbps}$

Considerant un tràfic de 400 Mbps, el tràfic amb el disc seria de  $200 * 0.25 * 400 \text{ Mbps} = 20000 \text{ Mbps}$

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

Tenim una LAN Ethernet de 10 Gbps i el tràfic total en cas pitjor es de 80000 Mbps (80 Gbps). No ens arriba.

En el cas mitjà el tràfic total es de 200 Mbps, per tant si ens arribaria i tindrem de sobres.

# 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

Les decisions preses han sigut en base al tràfic en mitja.

Com tenim dos grups de dades (actuals i històriques), hem decidit que lo millor és tenir-les separades en cabines diferents.

En les dues cabines hem decidit utilitzar RAID 51 ja que ens proporciona la combinació de RAID 5 i RAID 1 (Mirroring) així tot i ser més car tenim gran nivell de tolerància de fallades i en el nostre escenari hem de ser molt estricts en cas de tenir corrupció de dades degut a la gran xifra en SLA.

Per les dades actuals (200 TB) hem triat l'opció 9 de Disc SSD Enterprise (WD Gold S768T1D0 D) de 7.68 TB. El mínim nombre de discos per a assolir Raid 51 en aquest cas seria de 56, però nosaltres hem decidit comprar 144 discos per poder créixer en el futur.

Veient el nombre de discos, necessitem cabines lo més grans possibles i que siguin compatibles amb spare disk, ja que tenim un SLA que ens obliga a pagar molta quantitat de diners per cada hora de downtime. Com que el nombre màxim de badies es de 36 descartem d'entrada les cabines 1, 2, 3 i 6. Entre les cabines 4 i 5 hem triat la 4 per què no ens fa falta suport SSD.

Comprem 4 cabines tipus 4 i en cada cabina posarem 2 clusters de 10 discos i un de 12 discos + 4 spare disc (ocupant així les 36 badies de cada cabina). Així, discos de dades en total hi hauran 52 (13 en cada cabina).  $52 * 7.68 \text{ TB} = 399 \text{ TB}$ . Ocupem només 200 TB en dades actuals, llavors el marge de creixement seria de 199%. En quant a IOPS tindrem  $144 \text{ discos} * 467000 \text{ IOPS Read} = 67.248.000 \text{ IOPS lectura}$  i  $144 * 65000 \text{ IOPS Write} = 9.360.000 \text{ IOPS Escriptura}$ . Només necessitem 1563, tenim un marge de creixement molt gran.

Per les dades històriques (40 TB) hem triat l'opció 3 de Disc HDD Enterprise (Seagate ST10000N M009G) de 10 TB. El mínim nombre de discos per a assolir Raid 51 en aquest cas seria de 10, pero hem comprat 24 i els posem en una cabina de tipus 2 on hi ha 2 spare disk amb 1 cluster de 12 discos i un de 10 discos ((5+4) discs de dades i (1+1) discs de paritat)\*2 de RAID . Ocupen només 40 TB i tenim 90 TB, llavors podem arribar a créixer un 225% en dades històriques.

Resum de les coses comprades:

- 144 discos opció 9 ( $144 * 1545 = 222.480 \text{ €}$  )
- 24 discos opció 3 ( $24 * 350 = 8.400 \text{ €}$ )
- 4 cabines model 4 ( $4 * 5000 = 20.000\text{€}$  )
- 1 cabina model 2 (4.600 €)

Total gastat = 255.480 €

### **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**

El tràfic total en mitja és de només 0.2 Gbps i tenim una LAN de 10 Gbps, llavors en principi no caldria una SAN ja que podem arribar a créixer fins un 5000% amb la LAN que tenim. Pero el nostre escenari contempla un cas extrem, podem arribar a tenir pics de treball de 400 Mbps, lo que pot suposar un tràfic total de 80 Gbps i en aquest cas la LAN no ens arriba i s'hauria de posar una SAN. Degut al poc pressupost que tenim, no ens podem permetre posar una SAN i s'hauria de parlar amb la empresa per trobar una solució per als moments on hi ha aquests pics de treball.

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

Seria molt interessant tenir un mirror pero no ens arriba el pressupost. Farem backups on-site i off-site.

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

Hem escollit la empresa de housing Mordor Colocation Center, encara que es una mica més cara que les altres dues ens garanteix un temps de downtime per any molt més baix (1,6 hores al any) i això en el nostre cas es essencial ja que tenim un SLA que ens fa pagar 1.000.000 € per cada hora o fracció de downtime.

### **3.5- Posem monitorització?**

Si, ja tenim monitorització inclosa en la empresa de housing que hem triat.

### 3.6- Opció de backup?

Com no tenim mirror fem backup on-site i també off-site. Per el backup off-site hem escollit l'opció 3 (Take the tapes and run) ja que ens surt molt més barat que les altres opcions al tenir una gran quantitat de dades. Hem decidit fer un backup diari i que es guardi durant 10 dies (ja que en el nostre escenari les dades només són útils durant 10 dies).

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

Tenim una LAN Ethernet de 10 Gbps i considerant el cas mitjà, el tràfic es de 200 Mbps. Veient això tenim la probabilitat de créixer en el futur. Doncs, hem escollit la opció 4 que té la velocitat de sortida de 10 Gbps. Encara que lo òptim seria poder comptar amb la velocitat de sortida de 100 Gbps (per el tràfic en pics) però com ja hem mencionat abans el nostre escenari no permet tenir en compte això ja que el pressupost és massa petit.

## 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?**

Com fem backups diaris i els guardem durant 10 dies, podem recuperar versions de fins a 10 dies abans.

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

Si es trenca un disc, com que tenim RAID 51 no perderiem dades i es recuperarien de forma immediata gràcies al mirror que proporciona RAID 51. També RAID 51 tolera la fallada d'un disc i les dades continuaran sent accessibles si això arriba a passar. El negoci no s'hauria de aturar en cap cas.

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

No hauríem de tenir cap problema ja que treballem amb RAID 51.

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

Si cau una línia elèctrica no tenim cap problema ja que la nostra empresa de housing ens proporciona dues línies d'entrada d'electricitat i a més a més també té un generador diesel amb capacitat per aguantar la potència pic durant 72 hores, que a més pot fer re-ompliment en calent. La seva certificació garanteix un 99,982% de uptime (és a dir, un màxim de 1,6 hores de downtime a l'any).

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

La nostra empresa de housing permet tenir duplicades les línies de connexió a la xarxa. Llavors hem decidit contractar una línia de backup d'un segon proveïdor de 1 Gbps per si de cas.

- **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**

**Dades actuals (200 TB):**

Raid 51

1% de 200 TB = 2 TB = 2000 GB

Recuperem com a molt a 10 Gbps →  $10 \text{ GBps} / 8 = 1.25 \text{ GBps}$

Necessitem  $2000 \text{ GB} / 1.25 \text{ GBps} = 1600 \text{ segons}$  (uns 27 minuts)

**Dades històriques (40 TB):**

Raid 51

1% de 40 TB = 0.4 TB = 400 GB

Recuperem com a molt a 10 Gbps →  $10 \text{ GBps} / 8 = 1.25 \text{ GBps}$

Necessitem  $400 \text{ GB} / 1.25 \text{ GBps} = 320 \text{ segons}$  (uns 6 minuts)

Encara que tenint RAID 51 tenim una còpia, llavors no faria faltar recuperar les dades.

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

**Dades actuals (200 TB):**

Raid 51

200 TB = 200000 GB

Recuperem com a molt a 10 Gbps →  $10 \text{ GBps} / 8 = 1.25 \text{ GBps}$

Necessitem  $200000 \text{ GB} / 1.25 \text{ GBps} = 160.000 \text{ segons}$  (unes 45 hores)



### Dades històriques (40 TB):

Raid 51

40 TB = 40000 GB

Recuperem com a molt a 10 Gbps  $\rightarrow 10 \text{ Gbps} / 8 = 1.25 \text{ GBps}$

Necessitem  $40000 \text{ GB} / 1.25 \text{ GBps} = 32000 \text{ segons}$  (unes 9 hores)

Veient aquests números tenim una altre raó per poder augmentar la LAN. S'hauria de parlar amb el client.

### **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 complert, calcular quant perdo en diners per tenir-lo aturat i quan em costaria evitar aquesta situació.

#### Caiguda de la xarxa de dades:

Tenint només una línia la probabilitat downtime es entre 0,017% i 0,036% i en canvi amb dues línies la probabilitat downtime seria entre 0,00034% i 0,00071%.

Tenim un SLA de 1M € per hora downtime. En 5 anys (43830 hores) tenim entre un 0,017% i 0,036% de possibilitats de downtime, o sigui entre 7.45 (8 hores) i 15.77 (16 hores). Per tant la penalització que pagariem als clients estarà entre 8 i 16 milions €. Però la empresa de housing ens assegura que no tindrem més de 1.6 hores a l'any de downtime (8 hores en 5 anys). Llavors pagariem 8 milions en 5 anys i descartem totalment els 16 millions.

Però ens estalviem encara més, ja que nosaltres al tenir dues línies el màxim de downtime en 5 anys seria de 20 minuts ( $0,00071\% * 43830 \text{ hores}$ ), o sigui 334.000 euros.

#### Fallada de disc

#### Disc de dades actuals:

Estem en RAID 51. Si falla un disc podem copiar-lo del mirròr, no cal reconstruir. Si es pot predir la fallada per SMART (70%) es farà la còpia quan el clúster estigui inactiu. Tenim 128 discos (52 de dades i 76 per RAID). Disc SSD enterprise TLC (probabilitat de fallada 0,45% anual). De les fallades, 30% requereixen reconstrucció, la resta SMART.  $128 \text{ discos} \times 0,45\% \text{ fallada} \times 30\% \text{ fallades amb reconstrucció} = 17,3\% \text{ de probabilitat de fallar/any}$ , 86,4% probabilitat falli un disc en 5 anys.

Com tenim més IOPS de les que necessitem, si necessitem copiar un disc en un altre és ràpid.

### Disc de dades històriques:

Estem en RAID 51. Si falla un disc podem copiar-lo del mirròr, no cal reconstruir. Tenim 22 discs(9 de dades i 13 de RAID). El disc falla 2,84% anual (HDD Enterprise <10000 rpm). Fallen  $22 \times 0,0284 = 0,6248$  discos per any  $\times 5$  anys = 3.124 discos. El 70% fallades es poden predir per SMART, s'han de reconstruir el restant 30%  $\Rightarrow 3.124 \times 0.3 = 0.94$  discos en total (posem 1).

#### **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, ...)

Si creix el nombre de dades un 20% no implicaria res ja que nosaltres ens hem preparat per créixer encara més (199% en el cas de les dades actuals i 225% en les històriques).

Si creix la pressió sobre la xarxa (del tràfic en mitja) un 20% també estem preparats a assumir fins al 5000% de creixement. En IOPS també tenim bastant marge com ja hem vist en els apartats anteriors.

#### **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...

Primer de tot, en tots els cas hem assumit que el tràfic de xarxa era pel cas mitjà per això tenim la possibilitat de créixer fins a 5000% ja que el tràfic mitja total es de 0,2 Gbps i la nostre xarxa Ethernet es de 10 Gbps, pero en el cas pitjor pot haver pics que poden arribar a augmentar el tràfic fins a 80 Gbps (40.000% de creixement), doncs per poder atendre a aquest tràfic hauríem de tenir pressupost per optar per la SAN d'opció 5, de capacitat 128 Gbps o en canvi augmentar la LAN fins a 100 Gbps.

Si haguéssim fet la pràctica tenint en compte el cas dels pics màxims, no hauríem pogut dissenyar res ja que lo primer que hauríem de fer seria posar la SAN d'opció 5 que costa uns 5 milions d'euros i ens hauríem quedat sense diners per comprar discos, pagar el SLA etc.

També estaria bé poder tenir un mirròr.

