

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Estudiant 1 Algarra Villalba, Victor
Estudiant 2 Peña Peregrina, Oscar**

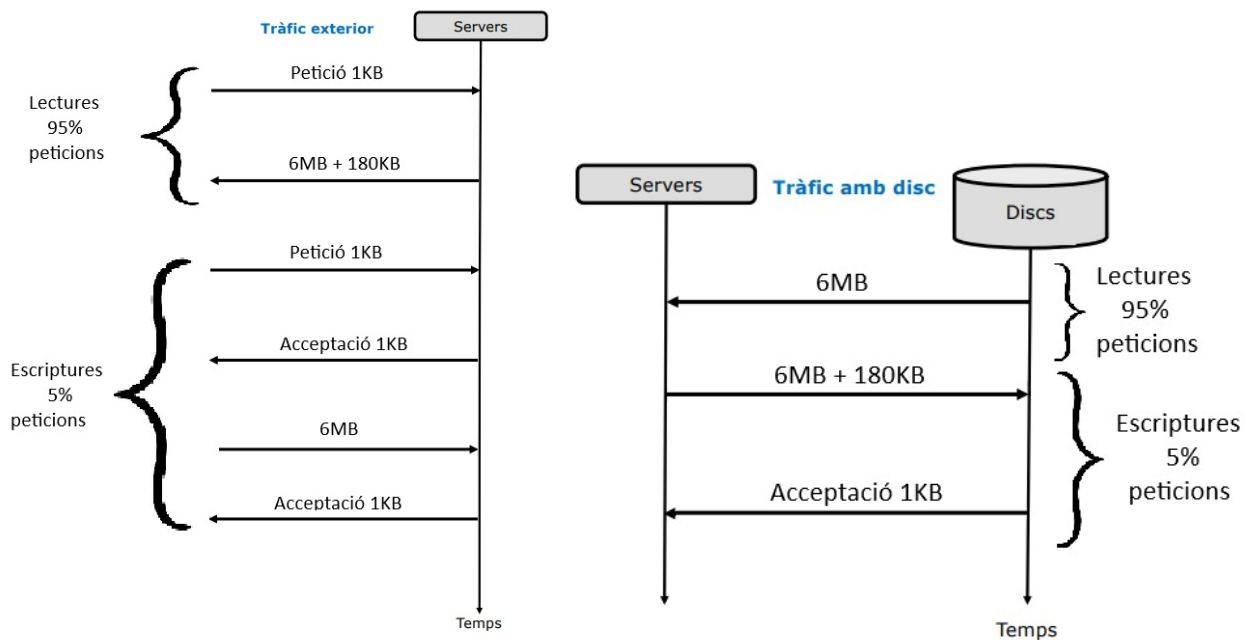
Escenari NET1

Data: 13/10/2023

1.-Descripció bàsica

TAULA 1: ESCENARI ORIGINAL: EXTRET DE L'ENUNCIAT. OMPLIU EL QUE HI HA EN GRIS.		TAULA 3: OPEX		anual
Nombre de Us	972U	Consum energètic (hardware només)	€365.861,27	
Alçada Rack (en Us)	42U	Empresa de Housing escollida	Mordor	
Consum	1979,8kW	Cost Housing (inclou electricitat addicional)	€474.879,19	
Sobreprovisionament d'electricitat	12%	Off-site: empresa escollida	Take the tapes and run	
Nombre de servidors	470	Cost mirror	€0,00	
Diners Totals	€25.000.000,00	Cost backup	€312.600,00	
Diners gastats	€15.500.000,00	Cost Bandwidth provider	€10.584,00	
taula 2: Elements que escolliu vosaltres		TAULA 4: CAPEX		Cost
Elements de mirror i backup		Diners gastats en servers, xarxa, etc	€15.500.000,00	
GB a emmagatzemar al backup	3600000	SAN	€2.868.792,00	
Dies entre 2 backups	1	Sistema emmagatzematge	€996.300,00	
Còpies senceres a mantenir	5			
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			
Elements de housing		TAULA 5: AJUST AL PRESSUPOST		
Opció escollida (1:MOCOSA, 2: CPDs Céspedes, 3: Mordor)	3	Opex a 5 anys, total	€5.819.622,29	
Gestió local de backup? (0=No, 1=SI)	1	Capex a 5 anys, total	€19.365.092,00	
Monitorització? (0=NO; 1=SI)	1	Despeses totals a 5 anys	€25.184.714,29	
Bandwidth provider		Diferència respecte al pressupost	-€184.714,29	
Tipus de línia (1:10Mbps; 2:100Mbps; 3:1Gbps; 4:10Gbps; 5:100Gbps)	4			
Número de línies agregades	1			
Segon proveïdor? (0=NO, 1=SI)	1			
SAN? (0=no, 1=8Gbps, 2=16Gbps, 3=32Gbps, 4=64Gbps, 5=128Gbps)	2			
Cabina de discos				
Opció Disc principal (Entre 1 i 10)	9			
Nombre de discos a comprar	540			
Opció cabina de discos (Entre 1 i 6)	4			
Nombre de Cabines	15			
Cabina de discos 2 (cas de fer servir dos tipus)				
Opció Disc (Entre 1 i 10)	3			
Nombre de discos a comprar	170			
Opció cabina de discos (Entre 1 i 6)	5			
Nombre de Cabines	5			

2.-Anàlisi de necessitats



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

$$600.000 \text{ usuaris} \times \frac{6 \text{ MB}}{\text{fotografia}} \times 1000 \text{ fotografia} = 3.600.000.000 \text{ MB} = 3.600.000 \text{ GB} = 3.600 \text{ TB}$$

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

$$10.000 \frac{\text{peticions}}{\text{minut}} \times \frac{1 \text{ minut}}{60 \text{ segons}} \times \frac{6.009,05 \text{ KB}}{\text{petició}} = 1.001.508,333 \text{ KBps}$$

$$\text{Els discos accedeixen en blocs de 4KB per tant: } \frac{1.001.508,333}{4 \text{ KB/operacio IO}} = 250.377,0825 \text{ IOPS}$$

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

$$\text{Cada petició} = 0,95(1 \text{ KB} + 6180 \text{ KB}) + 0,05(1 \text{ KB} + 1 \text{ KB} + 6000 \text{ KB} + 1 \text{ KB}) = 6172,1 \text{ KB} = \frac{6,1721 \text{ MB}}{\text{petició}}$$

$$\frac{6,1721 \text{ MB}}{\text{petició}} \times 10.000 \frac{\text{peticions}}{\text{minut}} \times \frac{1 \text{ minut}}{60 \text{ segons}} \times \frac{8 \text{ Mb}}{1 \text{ MB}} = 8.229,46 \text{ Mbps}$$

2.4- Tràfic amb el disc:

$$0,95(600 \text{ KB}) + 0,05(6.000 \text{ KB} + 180 \text{ KB} + 1 \text{ KB}) = 6,00905 \text{ MB}$$

$$\frac{6,00905 \text{ MB}}{\text{petició}} \times 10.000 \frac{\text{peticions}}{\text{minut}} \times \frac{1 \text{ minut}}{60 \text{ segons}} \times \frac{8 \text{ Mb}}{1 \text{ MB}} = 8.012,06 \text{ Mbps}$$

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

$$\text{Tràfic de client} = 8.229,46 \text{ Mbps}$$

$$\text{Tràfic de disc} = 8.012,06 \text{ Mbps}$$

$$\text{Trafic total} = 8.229,46 \text{ Mbps} + 8.012,06 \text{ Mbps} = 16.241,52 \text{ Mbps}$$

3.-Decisions preses

El nostre escenari tenia dos funcionalitats ben diferenciades, una primera dedicada a la visualització de fotografies (interactiu) i un altre dedicat al emmagatzemament de fotografies (dropbox). La visualització interactiva requeria de més espai ja que la majoria d'usuaris es troba allà en el nostre cas el 70% mentres que en el dropbox tenim el 30%.

Dades:

Total = 3600 TB

Dropbox = 1080 TB

Interactiu = 2520 TB

IOPS Total = 250380

IOPS Dropbox = 75114

IOPS R Dropbox = 71.358,3

IOPS W Dropbox = 3.755,7

IOPS Interactiu = 175266

IOPS R Interactiu = 166.502,7

IOPS W Interactiu = 8763,3

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

Interactiu: Pel interactiu hem decidit utilitzar discos SSD de tipus 9, ja que volem una resposta ràpida i tindre les dades disponibles el més ràpid possible, hem escollit aquest tipus perquè encara que són més cars que les altres opcions SSD enterprise, amb aquesta necessitem menys quantitat de discos, el que comporta menys gastos en consum i el preu final ens donava similar. Hem optat per una distribució en RAID 5 ja que ens dona una certa seguretat si un disc falla poguer recuperar les dades d'una manera ràpida i sense que afecti al funcionament del sistema. Lo cert es que en un principi ens havíem decantat per un RAID 51 per diferenciar-nos de Dropbox i perquè realment a priori era una molt bona opció, pero a la pràctica suposa un gran nombre de discos i molt pocs amb

dades reals o no redundants. Per això finalment vam escollir RAID 5, que podem optar a més discos de dades no redundants que era els que ens interessa, ja que com veurem més endavant, a interactiu l'espai és el recurs que abans s'esgota. Hem escollit cabines del tipus 4, ja que era la cabina que tenia més badies, no tenia SSD suport (ja que els nostres discos ja son SSD de per si), suportava la nostra distribució de RAID 5. A més hem decidit incorporar 4 discos de spare per afegir més fiabilitat en cas predir la fallada d'un disc i aquesta cabina ens ho permetia. Finalment hem dividit els nostres discos en clusters de 4 discos cadascun (3 de dades i 1 de RAID) per tant ens quedaven 8 clusters per cabina més els 4 discos de spare.

Un cop decidit això al excel de discos requerits ens posava que minim necessitavem 330 discos per RAID 5, vem decidir que posariem 360 discos de dades per sobredimensionar una mica. Incloent els discos de RAID i els de spare en total tenim un total de 540 discos situats en 15 cabines.

Uns 1500 cops superior la nostra capacitat de IOPS. $\frac{540 \times 467000}{166502,7} = 1.514,57$

Tenim un 1,09 vegades la capacitat requerida $\frac{360 \times 7680}{2520000} = 1,09$

Dropbox: Pel dropbox hem decidit utilitzar discos HDD de tipus 3, ja que en aquest cas volem tindre fiabilitat de les dades i no és tan rellevant el temps de resposta en aquest cas. Hem escollit el tipus 3 ja que era el que més ens convenia per les seves característiques, econòmic, amb alta capacitat i un consum raonable ja que el volum de discos respecte a les altres opcions era molt inferior. Hem optat també per una distribució en RAID 5 perquè encara que no tenim les mateixes IOPS que amb els SSD no ens importa trigar una mica més en recuperar les dades d'un disc que hagi fallat, també en aquest cas posarem 4 discos de spare perquè dona una fiabilitat més alta davant de fallades. Hem escollit una cabina de tipus 5, ja que era una cabina que tenia les 36 badies, suport SSD per agilitzar les operacions que calguin fer, suporta RAID 5 i dona la possibilitat de posar 4 discos de spare. En aquest cas hem dividit els nostres discos en clusters de 6 discos (5 de dades i 1 de RAID) per tant ens queden 5 clusters per cabina més els 4 discos de spare i els 2 de suport SSD.

Un cop decidit això al excel de discos requerits ens posava que minim necessitavem 122 discos per RAID 5, vem decidir que posariem 125 discos de dades per cuadrar els discos en les cabines. Incloent els discos de RAID i els de spare en total tenim un total de 170 discos situats en 5 cabines.

Un 1,69 cops superior la nostre capacitat en IOPS $\frac{170 \times 710}{71358,3} = 1,69$

Un 1,15 cops superior la nostre capacitat $\frac{125 \times 10000}{1080000} = 1,15$

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

Sí perquè la pressió del tràfic supera la xarxa LAN de 10 Gbps en 6 Gbps. Llavors com solament es pot dividir en funció de tràfic exterior i tràfic en el disc hem d'agafar una SAN de 16Gbps per tal de suplir un tràfic mínimament superior a 8 que suposa el trafic de disc, ja que la de 8 resultaria ser insuficient. La veritat és que suposa un gran cost afegir una SAN de 16Gbps, és possible que sigui més interessant disminuir el tràfic per tal de tenir suficient amb una SAN de 8Gbps, però ens hem d'ajustar als requisits demanats.

3.3.- Posem un *mirror*?

No, suposa un cost que en qualsevol cas no podem assumir, i no ens dóna una gran ventatge respecte el no tenir-lo.

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

Hem escollit la empresa Mordor perquè és la que ens garanteix que es compleix el nostre SLA en qualsevol cas, tant per la part elèctrica com per la part de xarxa, ja que ens dóna la possibilitat de contractar una segona línia per un 40% del preu i ens interessa, perquè la gestió del backup és relativament barata a diferencia de si la tinguéssim que fer nosaltres i té la seguretat més alta de les tres alternatives.

3.5- Posem monitorització?

Sí perquè al nostre housing ja inclou la monitorització

3.6- Opció de backup?

Cintes, totes les altres alternatives suposen un gran cost que no podem assumir.

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

Degut a que no tenim cap mirror i el nostre backup es per cintes no se'ns afegeix cap trafic amb el exterior. Per tant tindrem un bandwidth igual a la nostre LAN de 10 Gbps, ja que tenim una SAN de 16 Gbps per tant limitaria la conexio de 10 Gbps.

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?**

De un dia, una setmana, un mes, un trimestre i un any.

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

Interactiu: Es pot predir la fallada per SMART (70%) es farà la copia quan el cluster estigui inactiu.

Disc SSD enterprise TLC (0.45% anual).

De les fallades, 30% requereixen reconstrucció, la resta SMART.

$$15 \times 32 \times 0,0045 \times 0,3 = 0,648 \text{ disc/any}$$

$$0,648 \times 5 = 3,24 \text{ discs cada 5 anys}$$

Cada fallada que necessita reconstrucció, com estem en RAID 5 trigo 4h per TB els discos son de 7,68 TB per tant triga 30,72h durant aquest temps els discos d'aquell cluster van a la mitad de velocitat. Tinc 120 clústers de 4 discos, 119 a tot funcionament i 1 a mig funcionament. Cada clúster té 4 discos a 65000 IOPS = 260.000 IOPS per clúster x 119,5 clústers en funcionament = 31.070.000 IOPS, l'escenari demana 175.266 IOPS, sense problemes.

Dropbox: : Es pot predir la fallada per SMART (70%) es farà la copia quan el cluster estigui inactiu.

Disc HDD Enterprise < 10000 rpm (2.84% anual).

De les fallades, 30% requereixen reconstrucció, la resta SMART.

$$5 \times 30 \times 0,0284 \times 0,3 = 1,278 \text{ discos/any}$$

$$1,278 \times 5 = 6,39 \text{ discos cada 5 anys}$$

Cada fallada que necessita reconstrucció, com estem en RAID 5 trigo 4h per TB els discos son de 10 TB per tant triga 40h durant aquest temps els discos d'aquell cluster van a la mitad de velocitat. Tinc 25 clústers de 6 discos, 24 a tot funcionament i 1 a mig funcionament. Cada clúster té 6 discos a 710 IOPS = 42600 IOPS per clúster x 24,5 clústers en funcionament = 1.043.700 IOPS, l'escenari demana 75114 IOPS, sense problemes.

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

Com ja hem vist als calculs de adalt no

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

Estem hostatjats a MODOR? Segona línia elèctrica i SAI, garanteix 99.999% de uptime (menys de 30' en 5 anys).

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

Dues línies: probabilitat downtime entre 0,00034% i 0,00071%

Si SLA = 200.000€ per hora de downtime

- 5 anys (43830 hores) tenim entre un 0,00034% i 0,00071% de possibilitats de downtime, o sigui entre 0,149022 i 0,311193 (menys de 1 hora cada 5 anys). Per tant la penalització que pagarem als clients serà de 0 €.

- **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**

DROPBOX

$$\frac{5TB/hora}{3.600} = \frac{5000000000}{3.600} = 1.388.888 KBps$$

1080 TB discos tipus 3, 710 IOPS

Tenim RAID 5 per tant cada operació de escriptura son 2R+2W, podem fer

$$\frac{710}{2} \text{ escriptures reals per disc} \times 150 \text{ discos} = 53.250 \frac{\text{escriptures}}{s} \times \frac{4KB}{\text{escriptura}} = 213.000 KBps.$$

INTERACTIU

Tenim RAID 5 per tant cada operació de escriptura son 2R+2W, podem fer

$$\frac{65000}{2} \text{ escriptures reals per disc} \times 480 \text{ discos} = 15.600.000 \frac{\text{escriptures}}{s} \times \frac{4KB}{\text{escriptura}} =$$

62.400.000 KBps per tant limiten la capacitat de les cintes

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

DROPBOX

Tinc que recuperar 10.800.000.000 KB i recupero 213.000 KBps, necessito 50.704 segons (14 h)

→ sempre i quan tingui una bona xarxa.

2520 TB discos tipus 9, 467000 R i 65000 W IOPS (limiten les escriptures)

INTERACTIU

Tinc que recuperar 25.200.000.000 KB i recupero 1.388.888 KBps, necessito 18.144 segons (5 h)

→ sempre i quan tingui una bona xarxa

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

DROPBOX

Tinc que recuperar 1.080.000.000.000 KB i recupero 213.000 KBps, necessito 5.070.422 segons

(1.408 h) -> sempre i quan tingui una bona xarxa.

2520 TB discos tipus 9, 467000 R i 65000 W IOPS (limiten les escriptures)

INTERACTIU

Tinc que recuperar 2.520.000.000.000 KB i recupero 1.388.888 KBps, necessito 1.814.401 segons (504 h) -> sempre i quan tingui una bona xarxa

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:

0€ ja que com hem vist és menys de 1h de downtime cada 5 anys per tant no tindria impacte

Fallada de disc

0€ ja que com hem vist les IOPS en ambdós casos superen les exigències que suposa la fallada d'un disc i la seva recuperació de les dades.

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, ...)

Total = 3600 TB → 4.320 TB

Dropbox = 1080 TB → 1296 TB

Interactiu = 2520 TB → 3024 TB

IOPS = 250380 → 300456

IOPS Dropbox = 75114 → 90136,8

IOPS R Dropbox = 71.358,3 → 85629,96

IOPS W Dropbox = 3.755,7 → 4506,84

IOPS Interactiu = 175266 → 210319,2

IOPS R Interactiu = 166.502,7 → 199803,24

IOPS W Interactiu = 8763,3 → 10515,96

El primer recurs que s'esgota és la capacitat dels discos, tant en interactiu com en dropbox, al ajustar-nos tant al pressupost hem hagut de retallar en quantitat de discos per oferir els mínims requeriments donant una funcionalitat completa. Per tant un creixement d'un 20% tal i com està distribuït ara se'ns desbordaria la capacitat per les dues funcionalitats.

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, etc.

Les inversions més importants ronden sobre el nostre recurs limitant el qual és l'espai. Ara mateix solament tenim un marge de creixement de un 9% limitat específicament pel sistema interactiu, augmentar el nombre de discos suposaria una inversió considerable però es una opció a valorar. En quant al DropBox ara mateix compta amb un marge de creixement del 15% i en cas que volguéssim augmentar el marge per tot el sistema no supondria un gran cost augmentar aquest en comparació amb l'interactiu, ja que amb pocs discs podríem augmentar fàcilment l'espai. A més, si féssim això també augmentaríem els IOPS en Dropbox i disminuiria significativament el temps de recuperació de dades.

Un altre aspecte en el que es podria invertir és en el sistema de backup, actualment i degut al pressupost que teniem hem hagut de deixar un backup molt bàsic i poc freqüent, seria interessant valorar una inversió en millorar en aquest aspecte ja que la pèrdua del 100% de les dades suposa un downtime molt gran.