

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Estudiant 1 Amat, Alejandro
Estudiant 2 Oliva, Gerard**

Escenari NET2

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	cinc anys
Nombre de Us	134U	Consum energètic (hardware només)		€35,840.61	€179,203.07
Alçada Rack (en Us)	42U	Empresa de Housing escollida	Mordor		
Consum	180.5kW	Cost Housing (inclou electricitat addicional)		€131,376.09	€656,880.46
Sobreprovisionament d'electricitat	7%	Off-site: empresa escollida	Take the tapes and run		
Nombre de servidors	60	Cost mirror		€0.00	€0.00
Diners Totals	€10,000,000.00	Cost backup		€43,814.29	€219,071.43
Diners gastats	€7,500,000.00	Cost Bandwidth provider		€1,058.40	€5,292.00
taula 2: Elements que escolliu vosaltres		TAULA 4: CAPEX		Cost	
Elements de mirror i backup		Diners gastats en servers, xarxa, etc		€7,500,000.00	
GB a emmagatzemar al backup	500000	SAN		€0.00	
Dies entre 2 backups	7	Sistema emmagatzematge		€926,490.00	
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		€1,060,446.96	
Gestió local de backup? (0=No, 1=SI)	1	Capex a 5 anys, total		€8,426,490.00	
Monitorització? (0=NO; 1=SI)	1	Despeses totals a 5 anys		€9,486,936.96	
Bandwidth provider		Diferència respecte al pressupost		€513,063.04	
Tipus de línia (1:10Mbps; 2:100Mbps; 3:1Gbps; 4:10Gbps; 5:100Gbps)	3				
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)	0				
Cabina de discos					
Opció Disc principal (Entre 1 i 10)	8				
Nombre de discos a comprar	300				
Opció cabina de discos (Entre 1 i 6)	4				
Nombre de Cabines	9				
Cabina de discos 2 (cas de fer servir dos tipus)					
Opció Disc (Entre 1 i 10)	9				
Nombre de discos a comprar	42				
Opció cabina de discos (Entre 1 i 6)	4				
Nombre de Cabines	2				
Cabina de discos 3 (cas de fer servir tres tipus)					
Opció Disc (Entre 1 i 10)	8				
Nombre de discos a comprar	0				
Opció cabina de discos (Entre 1 i 6)	5				
Nombre de Cabines	0				

2.-Anàlisi de necessitats

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

En el nostre escenari, la meitat dels usuaris escullen el sistema de tipus dropbox i l'altra meitat el sistema interactiu. En el cas de dropbox obtenim un emmagatzematge total de

$$A_{dropbox} = 60000 \text{ usuarios} * 1/2 * 1000 \text{ fotos/usuario} * 6MB/foto = 180 TB$$

El cas interactiu és idèntic així que l'emmagatzematge total és

$$A_{total} = 2A_{dropbox} = 360 TB$$

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

Presió de lectura

$$\begin{aligned} P_L &= 0.95 * 3000 \text{ peticiones/min} * 1 \text{ min}/3600 \text{ sec} * C_{L_{int}} * 1/4 \text{ KB operaciones} \\ &= 1187.5 \text{ OPS} \end{aligned}$$

Presió d'escriptura

$$\begin{aligned} P_E &= 0.05 * 3000 \text{ peticiones/min} * 1 \text{ min}/3600 \text{ sec} * C_{E_{int}} * 1/4 \text{ KB operaciones} \\ &= 62.5 \text{ IPS} \end{aligned}$$

Presió total

$$P_T = P_E + P_L = 1250 \text{ IOPS}$$

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

Lectura: Tràfic exterior

Cada petició des del punt de vista del servidor és

$$C_{L_{ext}} = 1KB (http) + 6MB (foto) + 180KB (info adicional) = 6181 KB$$

Tenint en compte que tenim un cert número de peticions per minut de mitja:

$$\begin{aligned} T_{ext} &= 0.95 * 3000 \text{ peticiones/min} * 1 \text{ min}/3600 \text{ sec} * C_{L_{ext}} * 8 \text{ bits/B} * 1MB/1000KB \\ &= 39.14 \text{ Mbps} \end{aligned}$$

Escritura: Tráfico exterior

Cada petició des del punt de vista del servidor és

$$C_{Ext} = 3 * 1KB (http) + 6MB (foto) = 6003 KB$$

Tenint en compte que tenim un cert número de peticions per minut de mitja:

$$\begin{aligned} T_{ext} &= 0.05 * 3000 \text{ peticiones/min} * 1 \text{ min}/3600 \text{ sec} * C_{Ext} * 8 \text{ bits/B} * 1MB/1000KB \\ &= 2.001 \text{ Mbps} \end{aligned}$$

2.4- Tràfic amb el disc:

Esriptura: Tràfic interior

Cada petició des del punt de vista del servidor és

$$C_{int} = 6MB (foto)$$

Por tanto:

$$T_{ext} = 0.05 * 3000 \text{ peticiones/min} * 1 \text{ min}/3600 \text{ sec} * C_{Eint} * 8 \text{ bits/B} = 2 \text{ Mbps}$$

Lectura: Tràfic interior

Cada petició des del punt de vista del servidor és

$$C_{Lint} = 6MB (foto)$$

Per tant:

$$T_{ext} = 0.95 * 3000 \text{ peticiones/min} * 1 \text{ min}/3600 \text{ sec} * CL_{int} * 8 \text{ bits/B} = 38 \text{ Mbps}$$

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

Tràfic total

$$C_{TotExt} = C_{Lext} + C_{Eext} = 41.147 \text{ Mbps}$$

$$C_{TotInt} = C_{Lint} + C_{Eint} = 40 \text{ Mbps}$$

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

Segons els requisits d'emmagatzemament que hem calculat a l'apartat anterior, necessitem un total de 360TB, però hem decidit posar 500TB (38,89% més del requerit) per poder augmentar el número d'usuaris. D'aquests 500TB, la meitat (250TB), són d'usuaris interactius, i els altres 250TB d'usuaris tipus "dropbox". En el cas dels 250TB dels usuaris interactius hem decidit utilitzar un RAID 51, per aconseguir la màxima redundància possible i intentar que el *uptime* sigui el màxim possible. Per aquesta mateixa raó (màxima redundància possible), hem decidit utilitzar l'opció de disc 8, ja que ens ofereix un MTBF, més baix que no pas en discos durs magnètics. També aconseguim amb aquests discs uns dels IOPS més elevats, el qual en el nostre escenari no és necessari però ens pot ser útil en cas d'haver de fer un recovery del backup. Per últim també hem decidit utilitzar 4 spare discs per cada cabinet, per la màxima redundància possible. Seguint el que hem dit, hem seleccionat el cabinet número 4, el qual ens permet 36 discs per cada cabinet, incloent 4 spare discs. Per poder encabir 250TB en RAID 51, necessitem 264 discs de l'opció 8, i per poder encabir aquests discs més els spare discs per cada cabinet, necessitem 9 cabinets, i per tant 4 discs addicionals per cada cabinet.

En el cas dels 250TB dels usuaris, hem decidit utilitzar un RAID 5, ja que ens podem permetre que les dades estiguin no disponibles més temps. Per aquest motiu hem decidit utilitzar l'opció de discs 9, la qual té una densitat de capacitat més alta, però els IOPS de lectura són molt més baixos. Per aconseguir els 250TB en RAID 5, només necessitem 34 discs, i igual que abans hem decidit posar 4 spare discs a cada cabinet, i hem decidit utilitzar els mateixos cabinets que abans, perquè tot i que sobren molts discos, de l'altre manera no tindrem gens de marge per afegir-ne, i amb només 2 cabinets de l'opció 4 en tenim de sobres pels discs i els spare discs.

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

En el nostre cas, no necessitem un SAN, ja que amb la LAN que tenim en tenim de sobres com pel nostre tràfic actual (utilitzem només 40Mbps dels 10Gbps que disposem).

3.3.- Posem un *mirror*?

En el nostre cas no teníem els recursos econòmics com per contractar un sistema de backup extern que també inclou un *mirror*, entrarem més en detall a l'apartat de back-up. Ja que no posem un *mirror*, posem un backup de cintes on site, el qual ens ofereix un ample de banda suficientment alt, com perquè en cas d'haver de recuperar les dades del bac-up on site, el coll de botella estarà sempre en els nostres discs (tot i estar utilitzant discs SSD amb un nombre de IOPS d'escriptura alt).

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

L'empresa de housing que hem escollit és l'opció 3, ja que és la que ens ofereix més opcions, tant de monitoratge, com de redundància, que com ja hem dit amb anterioritat és una de les nostres prioritats. També inclou la gestió de local de back-up on site, la qual tenim a més de l'empresa de backup.

3.5- Posem monitorització?

Com que utilitzem spare discs, volem posar monitorització per canviar els discs quan s'escaigui, i en l'empresa de housing que hem escollit, aquest servei ja ve inclòs en el preu.

3.6- Opció de backup?

Com ja hem comentat en l'apartat de mirror, amb el pressupost que nosaltres tenim no tenim suficient com per contractar una empresa de back-up extern, per això hem decidit utilitzar el mètode de "takes de tapes and run", en el qual fem un backup en una cinta local, i després una empresa s'encarrega de venir a buscar la cinta, i emmagatzemar-la a les seves instal·lacions. Una de les coses que ens agradaria afegir si tinguéssim els diners necessaris per fer-ho seria poder contractar un servei de backup extern i mirror, d'aquesta manera en cas que passés alguna cosa i s'hagués de recuperar les dades d'un backup, ens disminuiria dràsticament el temps per recuperar-ho.

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

Com que el sistema de backup que estem fent servir no utilitza la xarxa externa, no tenim cap necessitat d'afegir bandwidth per poder poder fer els backups.

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 s'ha comentat fem servir el mètode "take tapes and run"; En el nostre cas hem escollit fer un full backup cada 7 dies i s'emmagatzemen els últims 5. Per tant les versions a recuperar podrien ser en el pitjor cas de fa 6 dies.

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

Fem servir molts mètodes de seguretat per intentar minimizar el temps on s'ha d'aturar el negoci. És per això que no només tenim backups sino RAID 5 i RAID 51 on en principi si només es trenca un disc podem recuperar la informació en poc temps. En el cas de 51 no cal reconstrucció ja que podem agafar les dades del MIRROR. En el cas del 5 es pot fer la reconstrucció en calent (Unes 4 hores per TB tenint 7,68TB/disc en opció 9 són unes 30.72 hores) i funcionarà a la meitat de IOPS. A més disposem de Spare Disks per tant podem canviar els discos operatius i no haver

d'aturar el negoci; Amb SMART podem preveure la fallada de discs (70%). A més estem fent servir discs SSD amb TLC enterprise (8 i 9) per tant la probabilitat de fallida anualment és 0.45.

Tenim en el primer cas 9 cabines amb 32 discs funcionant en cadascuna (288) i en l'altre cas 2 cabines de 17 discs operatius. En el primer cas la probabilitat de fallida en una cabina és $32 \cdot 0,3 \cdot 0,45 = 4,24\%$ i en l'altre $17 \cdot 0,3 \cdot 0,45 = 2,295\%$.

En cap cas tindrem problemes amb els IOPS, ja que aquests discs son de >200k en R i >60k en W i el nostre sistema requereix 1250 per tant en cap moment serà un error per molt que funcion a meitat de velocitat per reconstrucció. Ja hem vist la baixa probabilitat i que el temps en RAID 5 és de 30.72 hores.

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

Tal com s'ha explicat en la pregunta anterior, utilitzant tants sistemes de seguretat l'impacte de la fallida d'un disc es veu reduït degut a las backups, sistemes de RAID i SPARE DISKS a les cabines.

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

En cas que la línia elèctrica caigués no passaria res i podríem seguir operatius, ja que el collocation center, té generadors per suportar la potència pic durant un màxim de 72 hores, i a més es podem re-omplir amb calent, amb la qual cosa no tenim perill de quedar-nos sense electricitat.

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

Tenim dos proveïdors amb una línia de xarxa agregada en cadascun per tant si en algún punt cau la línia d'un proveïdor podem fer servir l'altre. Si tinguéssim més diners, i el "collocation center" ho permetés, podríem afegir-hi més línies agregades.

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

1)

9 cabines amb 32 discs operatius (no spare) de 1,72TB = 495.36 TB. Dels quals de dades son 15, per tant $15 \cdot 1,72\text{TB} = 25,8\text{TB/cabina}$ a recuperar. IOPS: 540K /205K. Cada write és 4R+4W per tant podem fer $205\text{k}/4 = 51.25 \text{ KW/disc} \cdot 32 \text{ discos} = 1640\text{kW} \cdot 4\text{KB/W} = 6.560.000 \text{ KB/s}$.

Per tant $25,8 \cdot 10^9 / 6,56 \cdot 10^6 \simeq 1,1\text{h}$.

2) 2 cabines amb 16 discs útils (1 de RAID paridad i 4 Spare/cabina). raid 5-> 16 discs dades/cabina. $16 * 7,68\text{TB} = 122,68\text{ TB}$. W 65K. Write és $65\text{k}/2 = 32.5\text{ k} * 16\text{ discs} * 4\text{KB}/\text{W} = 2.080.000\text{ KB/s}$

$122,68\text{TB} / 2.080.000 = 16,38\text{ h}$

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 quan perdo en diners per tenir-lo aturat i quan em costaria evitar aquesta situació.

Caiguda de la xarxa de dades:

Com ja hem dit amb anterioritat, tenim redundància en la la xarxa de dades, la qual cosa fa molt poc probable un caiguda de les dues línies de xarxa. Si en posem en el cas que les dues línies de xarxa fallesin, gràcies a la certificació del colocation tier que hem escollit, com a màxim estarem 1,6 hores aturats a l'any, i amb el SLA, que tenim amb els nostres clients, això no ens suposaria cap penalització econòmica.

Fallada de disc

En el cas que estiguem parlant de la fallada d'un sol disc, no hi hauria cap mena d'implicació pel negoci, ja que tenim prou mesures de seguretat com perquè una fallada d'un sol disc no aturi l'empresa. Per altra banda si del que estem parlant és d'una fallada total del sistema de discs, el que ens limitaria és la velocitat d'escriptura dels discs, ja que podem recuperar 5TB/h, de la màquina de backup local. En els nostres discs màxim podem escriure a una velocitat de 2,952TB/h, amb la qual cosa tardarem $250 / 2,952\text{TB/h} \approx 84,69$ hores, restant les 24 hores que tenim al nostra SLA, hauriem de pagar $61\text{ hores} * 20000\text{ €} / \text{hora} = 1,22\text{ M€}$ per no complir amb el SLA.

Per poder evitar aquesta situació, hauriem de disposar d'un sistema de backup extern, amb opció de mirròr, per en el moment que passés algo en el sistema de discs poder utilitzar el mirròr. Aquesta opció de backup extern + mirròr, ens costaria uns 2,4M€ més dels que ens costa l'opció actual, que és més del que s'ha de pagar en indemnitzacions per no complir el SLA, però quan tenim en compte el que suposa a nivell de marca no poder complir amb el SLA, es fa bastant evident la importància del backup extern + mirròr.

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, ...)

El que més ens limita el creixement és l'emmagatzematge, ja que a nivell de xarxa, la nostra aplicació no és gaire intensa. Dit això, sense modificar el nostre CPD, podem assumir fins a un nombre de clients: $500.000.000 \text{ MB} / (6\text{MB/foto} * 1000 \text{ fotos/usuari}) \approx 83.333$ usuaris. el qual és un increment del 38% en el nombre d'usuaris respecte els actuals. El següent serien el creixement del nombre de peticions per segon, assumint un augment de les peticions per segon lineal respecte el nombre d'usuaris tindrem un total de tràfic extern de $\approx 56,78\text{Mbps}$, que encara està molt lluny del 1Gbps contractat, i com que en el cas del tràfic intern és una mica menor que l'extern, però la nostra LAN és de 10Gbps, encara estem més lluny del límit.

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 i seguretat.

Com ja hem dit abans en la nostra manera de veure, el més important seria poder tenir un sistema de backup extern amb un mirròr, per poder tenir uns temps de recuperació més baixos, ja que en una empresa que es dedica a l'emmagatzematge de fotos, i a la seva distribució en directe (50% del negoci), creiem que està molt temps per recuperar el sistema de disc, és una cosa que l'empresa no es pot permetre. I com ja hem dit costaria uns 2,4M€ més del cost actual del CPD poder posar un sistema de backup extern amb mirròr.