

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Arias Brian Fernando
Pla Sergi**

Escenari 03 NETPRO

Data: 28/09/2023

1.-Descripció bàsica

TAULA 1: ESCENARI ORIGINAL: EXTRET DE L'ENUNCIAT. OMPLIU EL QUE HI HA EN GRIS.	
Nombre de Us	134U
Alçada Rack (en Us)	42U
Consum	180.5kW
Sobreprovisionament d'electricitat	7%
Nombre de servidors	60
Diners Totals	€75,000,000.00
Diners gastats	€10,000,000.00

TAULA 2: Elements que escolliu vosaltres	
Elements de mirròr i backup	
GB a emmagatzemar al backup	60000000
Dies entre 2 backups	1
Còpies senceres a mantenir	8
Opció Backup (1=M-A; 2=MS3; 3=Cintes)	3
Opció Mirròr (0=NO; 1=SI)	0
Sistema de backup on-site? (0=N=; 1=SI)	1
Elements de housing	
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)	1
Bandwidth provider	
Tipus de línia (1:10Mbps; 2:100Mbps; 3:1Gbps; 4:10Gbps; 5:100Gbps)	4
Número de línies agregades	3
Segon proveïdor? (0=NO, 1=SI)	1
SAN? (0=no, 1=8Gbps, 2=16Gbps, 3=32Gbps, 4=64Gbps, 5=128Gbps)	3
Cabina de discos	
Opció Disc principal (Entre 1 i 10)	2
Nombre de discos a comprar	24444
Opció cabina de discos (Entre 1 i 6)	5
Nombre de Cabines	679
Cabina de discos 2 (cas de fer servir dos tipus)	
Opció Disc (Entre 1 i 10)	8
Nombre de discos a comprar	9396
Opció cabina de discos (Entre 1 i 6)	4
Nombre de Cabines	261

Cabina de discos 3 (cas de fer servir tres tipus)	
Opció Disc (Entre 1 i 10)	0
Nombre de discos a comprar	0
Opció cabina de discos (Entre 1 i 6)	0
Nombre de Cabines	0

TAULA 3: OPEX	anual	cinc anys
Consum energètic (hardware només)	€310,239.20	€1,551,195.98
Empresa de Housing escollida	Mordor	
Cost Housing (inclou electricitat addicional)	€1,418,535.88	€7,092,679.40
Off-site: empresa escollida	Take the tapes and run	
Cost mirror	€0.00	€0.00
Cost backup	€6,087,600.00	€30,438,000.00
Cost Bandwidth provider	€31,752.00	€158,760.00

TAULA 4: CAPEX	Cost
Diners gastats en servers, xarxa, etc	€10,000,000.00
SAN	€1,760,932.00
Sistema emmagatzematge	€21,245,692.00

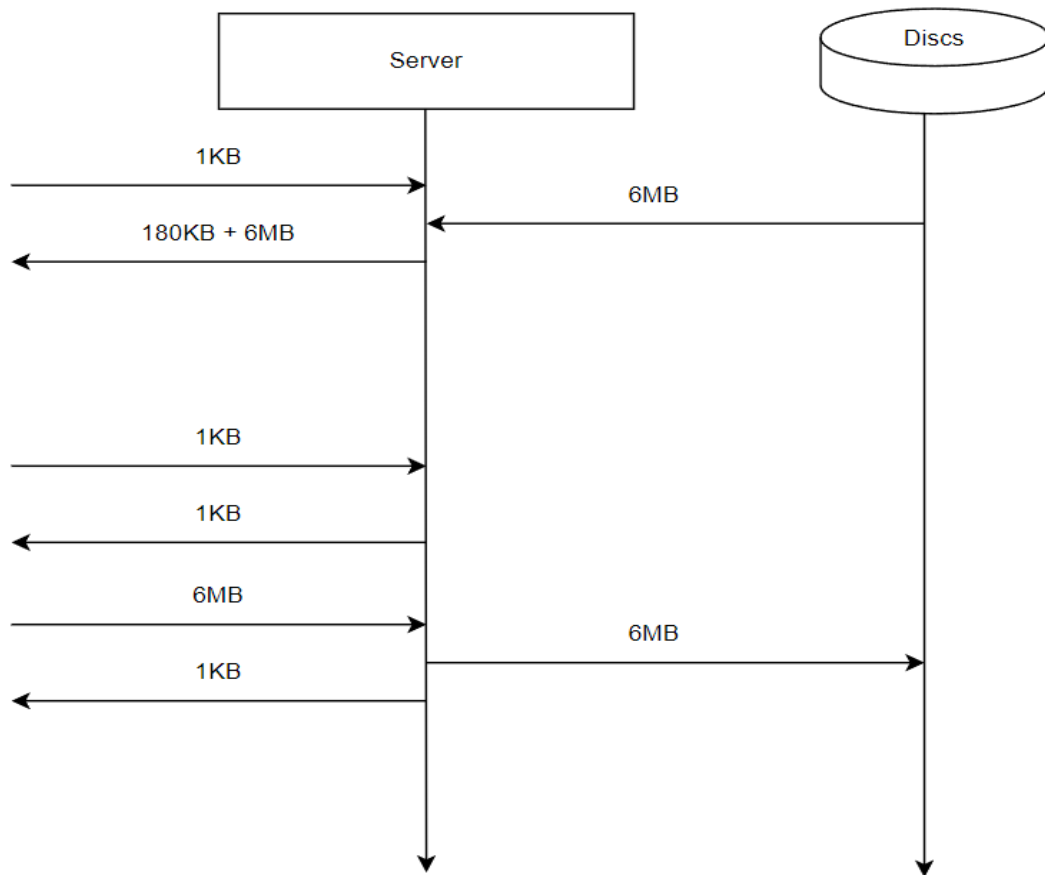
TAULA 5: AJUST AL PRESSUPOST	
Opex a 5 anys, total	€39,240,635.37
Capex a 5 anys, total	€33,006,624.00
Despeses totals a 5 anys	€72,247,259.37
Diferència respecte al pressupost	€2,752,740.63

2.-Anàlisi de necessitats

Pels següents càlculs hem de diferenciar quantes peticions son de lectura i quantes d'escriptura:

Dropbox		Interactiu	
10%		90%	
Lectura	Espectura	Lectura	Espectura
99%	1%	60%	40%

Lectura = $0.1 * 0.99 + 0.9 * 0.6 = 0.639$ (63,9%)
 Escriptura = $0.1 * 0.01 + 0.9 * 0.4 = 0.361$ (36,1%)



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

$1000 \text{ usuaris} * 10.000.000 \text{ fotos/usuari} * 6 \text{ MB/foto} = 60.000 \text{ TB}$

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

$15.000 \text{ petició/min} * 1/60 \text{ min/seg} * 6000 \text{ kB/petició} = 1.500.000 \text{ KBps}$

$1.500.000 \text{ KBps} / 4\text{kB/operació} = 375.000 \text{ IOPS}$

Lectura = (63,9%)
 Escriptura = (36,1%)

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

Diferenciem peticions d'escriptura i de lectura.

Cada petició d'escriptura: 1 Kb (petició) + 1 Kb (acceptació) + 6 MB (foto) + 1 Kb (acceptació) = **6.003 MB**

Cada petició de lectura: 1 Kb (petició HTTP) + 6 MB (foto) + 180 Kb (info adicional) = **6.181 MB**

Petició -> $6.181 \text{ MB} * 0.639 \text{ (\% lectura)} + 6.003\text{MB} * 0.361 \text{ (\% escriptura)} = \mathbf{6.1167 \text{ MB}}$

$15.000 \text{ petició/min} * 1/60 \text{ min/seg} * 6.1167 \text{ MB/petició} * 8\text{Mb/1MB} = \mathbf{12233.4 \text{ Mbps}}$

2.4- Tràfic amb el disc:

Les peticions a disc son sempre de 6 MB (mida de la foto) siguin lectures o escriptures, per tant:
 $15.000 \text{ petició/min} * 1/60 \text{ min/seg} * 6 \text{ MB/petició} * 8\text{Mb/1MB} = \underline{\underline{12000 \text{ Mbps}}}$

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

Tràfic amb el client = 12233.4 Mbps

Tràfic amb el disc = 12000 Mbps

Tràfic total = 12233.4 Mbps + 12000 Mbps = 24233.4 Mbps = **24.2334 Gbps**

El nostre escenari té una xarxa LAN de 10 Gbps, per tant no ens arribarà.

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

Per començar escollirem dos tipus de cabines, una dedicada al dropbox i l'altre per l'interactiu. La cabina del dropbox és el model de cabina 5 i contindrà discos HDD opció 2 ja que són els que ens ofereixen més capacitat per disc. En aquesta cabina ens interessa tenir la màxima capacitat possible ja que el dropbox ens ocupa el 95% de les dades totals per això agafem la que més badies té amb suport SSD. El nombre mínim de discs que necessitem per emmagatzemar les dades del dropbox son 4072 i 151 cabines.

Escollirem RAID51 ja que ens ofereix redundància de dades sense sacrificar molt la capacitat i a més tindrem un mirall de les dades per seguretat.

Finalment decidim 679 cabines model 5 dedicades al dropbox amb 24444 discos HDD opció 2.

CABINA DROPBOX

HDD opció 2 14TB

Cabina model 5 de 36 badies

RAID51 Cluster = 4 discs útils + 6 disc RAID

Cada cabina tindrà 3 clusters.

4 spare disk + 2 SSD discs + 18 discs RAID + 12 discs útils

$679 \text{ cabines} * 36 \text{ badies} = 24444 \text{ discos totals}$

$679 \text{ cabines} * 12 \text{ discos útils} * 14 \text{ TB} = 114072 \text{ TB espai útil}$

$60000 * 0.95 = 57000 \text{ TB espai ocupat}$

Per l'interactiu no ens cal tanta capacitat sinó que volem més velocitat ja que hi van el 90% de les transaccions. Per això escollirem el disc SSD opció 8 que és el que més IOPS ofereix per disc.

Fem servir la cabina 4 ja que ens ofereix el màxim nombre de badies i no ens cal suport SSD.

El nombre mínim de discs que necessitem per emmagatzemar les dades de l'interactiu son 1563 i 66 cabines. Escollirem RAID51 perquè com hem mencionat abans ofereix una bona capacitat amb molta seguretat. Per tant escollirem 261 cabines model 4 per l'interactiu amb 9396 discos SSD opció 8.

CABINA INTERACTIU

SSD opció 8 1.92TB

Cabina model 4 de 36 badies

RAID51 Cluster = 6 discs útils + 10 disc RAID

Cada cabina tindrà 2 clusters.

4 spare disk + 20 discs RAID + 12 discs útils

261 cabines * 36 badies = 9396 discos totals

261 cabines * 12 discos útils * 1.92 TB = 8478.72 TB espai útil

60000 * 0.05 = 3000 TB espai ocupat

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 que necessitem un SAN ja que inicialment tenim una LAN de 10Gbps i el tràfic intern amb els discs és de 12Gbps. Per tant, hem decidit contractar la SAN de 32Gbps per tenir una major capacitat d'expansió i mantenir l'estabilitat del servei.

3.3.- Posem un *mirror*?

No tenim un mirror contractat però amb el RAID 51 ja disposem d'un mirror local el qual podem fer servir davant fallades o errors. Per tant no considerem la necessitat de contractar un mirror.

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

L'empresa de housing escollida és Mordor Colocation Center. El primer motiu és la seguretat. El segon motiu és el downtime que ens ofereix que és del 1,6h a l'any ja que té dues línies d'electricitat i un generador dièsel. L'últim motiu ha estat el fet de poder contractar una segona línia de xarxa.

3.5- Posem monitorització?

Sí que contractarem monitorització ja que farem ús dels spare disk per incrementar la seguretat de les nostres dades y poder oferir un servei sense interrupcions.

3.6- Opció de backup?

Farem backup on-site i contractarem una empresa de transport de cintes per que mantingui els nostres backups desats i segurs. Tindrem 8 còpies, una mensual dels darrers 6 mesos i una diària dels últims 2 dies.

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

Com que no contractarem una empresa externa, no necessitarem incrementar el nostre bandwidth perquè suporti el servei.

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?**

Tenim un RAID 51 a cada cabina, es pot reconstruir el fitxer o copiar-lo del mirròr si aquest no s'ha perdut. Si s'ha perdut de manera que no es pot recuperar pel propi RAID llavors tenim varis backups per cintes que es poden utilitzar, tenim 8 còpies, una mensual dels darrers 6 mesos i una diària dels últims 2 dies.

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

Tenim 4 spare discs per cabina, per tant, es pot substituir el disc al moment i tenim dues formes de recuperar la informació: per el propi RAID utilitzant el bits de comprovació o amb una còpia directa del mirròr. Per tant és poc probable que hi hagi pèrdua de dades.

Com la còpia directa és més ràpida el temps de recuperació serà molt curt i no hauria de ser necessari aturar el negoci ja que la còpia dels discos seria ràpida i en el cas del dropbox qui té una major probabilitat de tenir discos trencats, només suposa el 10% de les nostres transaccions per tant no creiem que l'impacte sigui molt significatiu, en canvi en la cabina de l'interactiu sí que hi haurà la resta de transaccions i per això ens interessa una baixa probabilitat de fallada ja que no es podem permetre aturar el servei.

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

En general no hi hauria d'haver problema ja que tenim un mirròr inclòs en el RAID51 i per tant permet continuar amb el servei només reduint el rendiment d'escriptura a la meitat en el cluster.

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

En el nostre cas al tenir contractada l'empresa de housing Mordor no ens passaria res ja que contem amb una segona línia elèctrica i un generador dièsel capaç d'aguantar durant 72h.

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

Al tenir contractada Mordor també ens ofereix una segona línia de xarxa. També comptem amb un segon proveïdor de xarxa amb una altra línia contractada per si el problema fos de l'operador.

- 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

Cabina Dropbox

679 cabines tipus 5 amb 24444 discos tipus 2 (14TB) en RAID51

570 TB a recuperar (1% cabina)

Els discos tipus 5 poden fer 800 IOPS.

RAID 51, cada operació escriptura son en realitat 4R+4W, podem fer 800/8

$IOPS \times 24444 \text{ discos} = 2444400 \text{ IOPS/s} \times 4 \text{ KB/IOPS} = 9777600 \text{ KBps} = 9777,6 \text{ MBps}$

Tinc que recuperar 570000000 MB i recupero 9777,6 MBps, necessito 58.297 segons (16h 11minuts)

Cabina Interactiu

261 cabines tipus 4 amb 9396 discos tipus 8 (1,92TB) en RAID51

30 TB a recuperar (1% cabina)

Els discos tipus 5 poden fer 540K R/205K W IOPS(limiten les escriptures).

RAID 51, cada operació escriptura son en realitat 4R+4W, podem fer 205k/4 escriptures

reals per disc $\times 9396 \text{ discos} = 481545000 \text{ IOPS/s} \times 4 \text{ KB/IOPS} = 1926180000 \text{ KBps} = 1926180 \text{ MBps}$

Tinc que recuperar 300000000 MB i recupero 1926180 MBps, necessito

15 segons, sempre que la xarxa/LAN pugui suportar la velocitat de transferència.

- Pèrdua/ corrupció de la totalitat de les dades

Cabina Dropbox

679 cabines tipus 5 amb 24444 discos tipus 2 (14TB) en RAID51

57000 TB a recuperar

Els discos tipus 5 poden fer 800 IOPS.

RAID 51, cada operació escriptura son en realitat 4R+4W, podem fer 800/8

$IOPS \times 24444 \text{ discos} = 2444400 \text{ IOPS/s} \times 4 \text{ KB/IOPS} = 9777600 \text{ KBps} = 9777,6 \text{ MBps}$

Tinc que recuperar 57000000000 MB i recupero 9777,6 MBps, necessito 5.829.652 segons (més de 67 dies)

Cabina Interactiu

261 cabines tipus 4 amb 9396 discos tipus 8 (1,92TB) en RAID51

3000 TB a recuperar

Els discos tipus 5 poden fer 540K R/205K W IOPS(limiten les escriptures).

RAID 51, cada operació escriptura son en realitat 4R+4W, podem fer 205k/4 escriptures

reals per disc $\times 9396 \text{ discos} = 481545000 \text{ IOPS/s} \times 4 \text{ KB/IOPS} = 1926180000 \text{ KBps} = 1926180 \text{ MBps}$

Tinc que recuperar 30000000000 MB i recupero 1926180 MBps, necessito

1.558 segons (26 minuts), sempre que la xarxa/LAN pugui suportar la velocitat de transferència.

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:

Al tenir dues línies contractades la probabilitat està entre 0,00034% i 0,00071%. Per tant, en 5 anys (43830h) hi haurà una caiguda del servei d'entre 0.15h i 0.32h que és inferior a 1h. Com tenim acordat en el SLA, hem de pagar als clients 2.000.000€ per hora de downtime. És a dir que el màxim que s'haurà de pagar als clients serà aquesta quantitat comentada.

Fallada de disc

Cabina Dropbox

Estem en RAID 51. Si falla un disc puc 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. De les fallades, 30% requereixen reconstrucció, la resta SMART.

$24444 \text{ discos} * 2,84\% \text{ fallada} * 30\% = 208.26 \%$ de probabilitat de fallar a l'any

Això vol dir que com a mínim tenim assegurat que 2 discos es trenquin cada any i un 8.26% de que hi hagi un tercer. Una còpia és molt més ràpida que la reconstrucció per tant si es necessita recuperar un disc serà molt més ràpid.

Cabina Interactiu

Estem en RAID 51. Si falla un disc puc 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. De les fallades, 30% requereixen reconstrucció, la resta SMART.

$9396 \text{ discos} * 0.45\% \text{ fallada} * 30\% = 12,68 \%$ de probabilitat de fallar a l'any

Al igual que la cabina dropbox no serà necessari una reconstrucció , ja que és més ràpid copiar-lo del 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, ...)

Tenim 120000 TB d'espai total que es el doble de les dades actuals (60000TB) per tant podem emmagatzemar un 100% més de informació, el SAN que vam muntar té de capacitat 32Gbps això ens proporciona fins a un 166% de creixement respecte al nostre consum de 12 Gbps actual. Tenim una xarxa de 30Gbps amb un 150% respecte l'ample de banda que ja consumim. La nostra única limitació es la LAN que és de 10Gbps i ens arriba just per la capacitat actual amb la implementació de la SAN però si hi ha un increment de les transaccions tindrem un coll d'ampolla i el nostre servei serà més lent.

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...

L'aspecte més important de l'empresa és la quantitat de dades que s'emmagatzemen i la velocitat en que es poden servir les peticions. Per això hem invertit una gran quantitat de diners en espai, però si l'empresa tingués una previsió de creixement major on s'hi haurien d'invertir més diners és en millorar la xarxa per aguantar la pressió de les peticions.

El que podria ser un coll d'ampolla en el creixement és la LAN que actualment és de 10Gbps i ens podria limitar el trànsit de les dades. El que hem fet ha sigut contractar una SAN que ens ha ajudat a mitigar el problema però en cas de voler invertir-hi més el següent pas hauria de ser millorar la capacitat de la LAN. Perquè pugui suportar un tràfic de dades amb una capacitat similar a la SAN o la Xarxa(WAN). Una altra cosa a tenir en compte és que la distribució dels discos està pensada per aprofitar el màxim l'espai de les cabines, per tant, en el cas de voler augmentar la capacitat d'emmagatzematge s'haurien de comprar els discos i cabines desitjats.