

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Soley Masats, Jordi
Vidal Gonzalez, Albert**

Escenari 020

Data: 29/09/2023

1.-Descripció bàsica

| TAULA 3: OPEX | anual | cinc anys |
|---|---------------|---------------|
| Consum energètic (hardware només) | €40.619,78 | €203.098,91 |
| Empresa de Housing escollida | Mordor | |
| Cost Housing (inclou electricitat addicional) | €118.092,97 | €590.464,84 |
| Off-site: empresa escollida | MonsoonS3 MS3 | |
| Cost mirror | €0,00 | €0,00 |
| Cost backup | €704.300,00 | €3.521.500,00 |
| Cost Bandwidth provider | €10.584,00 | €52.920,00 |

| TAULA 4: CAPEX | Cost |
|---------------------------------------|----------------|
| Diners gastats en servers, xarxa, etc | €14.000.000,00 |
| SAN | €227.706,00 |
| Sistema emmagatzematge | €386.892,00 |

| TAULA 5: AJUST AL PRESSUPOST | |
|-----------------------------------|----------------|
| Opex a 5 anys, total | €4.367.983,74 |
| Capex a 5 anys, total | €14.614.598,00 |
| Despeses totals a 5 anys | €18.982.581,74 |
| Diferència respecte al pressupost | €6.017.418,26 |

2.-Anàlisi de necessitats

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

Ens arriben 100.000.000 de vídeos a emmagatzemar amb una mida de 5MB per vídeo.

$$100.000.000 \text{ videos} * \frac{5 \text{ MB}}{\text{video}} * \frac{1 \text{ GB}}{1000 \text{ MB}} = 500.000 \text{ GB}$$

Per tant, tindriem 500.000 GB a emmagatzemar.

Suposem que els 100MB de dades (per cada vídeo analitzat) que s'envien a un altre servidor no les hem de tenir en compte ja que es un servidor extern.

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

Cada dia fem:

- ➔ 50.000 input (vídeos que arriben)
- ➔ 1.000.000 output (vídeos que analitzem)

$$50.000 \text{ videos} * \frac{5 \text{ MB}}{\text{video}} + 1.000.000 \text{ videos} * \frac{100 \text{ MB} + 5 \text{ MB}}{\text{video}} = 105.250.000 \frac{\text{MB}}{\text{dia}}$$

Com aquestes dades son en un dia i IOPS son operacions per segon:

$$105.250.000 \frac{\text{MB}}{\text{dia}} * \frac{1 \text{ dia}}{86.400 \text{ segons}} = 1218,171 \text{ MBps} = 1218171 \text{ KBps}$$

$$\frac{1218171 \text{ KBps}}{4 \text{ KB per operació IO}} = 304.542,75 \text{ IOPS}$$

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

Cada dia tenim:

- ➔ els 20GB d'informe diari enviat per l'empresa i
- ➔ 1.000.000 vídeos analitzats per 100MB de dades per vídeo.
- ➔ 50.000 vídeos que arriben per 5MB per vídeo.

$$20.000 \text{ MB} + 1.000.000 \text{ videos} * 100 \text{ MB} + 50.000 \text{ videos} * 5 \text{ MB} = 100.270.000 \frac{\text{MB}}{\text{dia}}$$

$$100.270 \frac{\text{GB}}{\text{dia}} * \frac{1 \text{ dia}}{86400 \text{ segons}} * \frac{8 \text{ Gb}}{1 \text{ GB}} = 9,28 \text{ Gbps}$$

2.4- Tràfic amb el disc:

Cada dia tenim:

- ➔ 1.000.000 vídeos analitzats per 100MB de dades per vídeo.
- ➔ 50.000 vídeos que arriben per 5MB per vídeo.

$$1.000.000 \text{ videos} * 5 \text{ MB} + 50.000 \text{ videos} * 5 \text{ MB} = 5.250.000 \frac{\text{MB}}{\text{dia}}$$

$$5.250 \frac{\text{GB}}{\text{dia}} * \frac{1 \text{ dia}}{86400 \text{ segons}} * \frac{8 \text{ Gb}}{1 \text{ GB}} = 0,486 \text{ Gbps}$$

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

Per calcular al pressió sobre la xarxa, sumem el tràfic amb el client i el tràfic amb el disc.

$$9,28 \text{ Gbps} + 0,486 \text{ Gbps} = 9,766 \text{ Gbps}$$

Si que ens arriba ja que tenim una xarxa de 10 Gbps encara que aniríem justos ja que aquest valor es una mitjana i potser el pic supera els 10 Gbps.

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 aquesta versió del problema escolliríem una cabina, ja que totes les dades són del mateix tipus i no en tenim cap que sigui més prioritària que l'altra.

Tenim 500TB (mínim) de vídeos a emmagatzemar, una pressió de 304.542 IOPS i un percentatge d'escriptura del 4,76% ($\frac{50.000}{1.0500.00} * 100$).

Pensant en un futur i en què el sistema pot fer-se més gran, hem decidit que seria òptim escollir discs que ens permetin arribar fins a uns 700TB.

Utilitzarem discs Enterprise WD Gold S768T1D0D els quals tenen una capacitat de 7680GB i són SSD. Usarem una estratègia de RAID 10, per tant, en total necessitem 184 discs. Aquesta configuració tindrà un cost de 284.280 euros. Amb la configuració de RAID 10 tenim parelles les quals tenen les dades replicades, en conseqüència, només hi haurà perdudes importants si fallen els dos elements d'una parella. Pensem que és rellevant aquest factor per no perdre dades.

Per guardar els discs farem servir les cabines 5. Afegirem Spare Disks per tal de donar més seguretat. Distribuirem de la següent manera:

- ➔ 6 cabines amb 30 discs (15 + repliques) i 2 spare disk.
- ➔ 1 cabina amb 4 discs (2 + repliques) i 2 spare disk.

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

Pensem que posar un SAN seria la millor opció. Tenim un tràfic total de 9,766 Gbps mentre que la nostra xarxa és de 10 Gbps. Aquesta xarxa no ens dona marge de millora i en el nostre sistema hem mirat a futur i hem deixat opció de poder emmagatzemar fins a 700TB de dades. Suposaria un creixement del tràfic total del 140% és a dir, 13,66 Gbps.

Amb aquests números, la xarxa actual se'ns queda petita i podríem usar un SAN. Per tant, faríem servir el de 16 Gbps que és el que més s'ajusta al nostre sistema.

3.3.- Posem un *mirror*?

Pensem que posar un *mirror* no ens surt a compte. Al nostre sistema hem plantejat una solució amb RAID 10 on hi ha una còpia de les dades a dos discs. A part, tenim dos Spare Disks a cada cabina i no tenim SLA amb cap client. Per tant, considerem que no ens surt a compte tenir un *mirror*.

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

Com a empresa de *housing* hem escollit la tercera opció (Mordor Colocation Center). S'ajusta al nostre pressupost i ens ofereix millor seguretat. Això és important, ja que el nostre model de negoci es basa en les dades i generar informes. També ofereix monitorització del hardware que facilita la monitorització. Compta com a factors a favor tenir un bon PUE i una replicació de 2N, el qual ens va bé per tal de protegir-nos davant de fallades. Finalment, tenim dues línies de xarxa i d'electricitat.

3.5- Posem monitorització?

Si ja que amb el servei de *housing* que hem escollit ve inclosa en el preu. A part, en vindrà bé per monitoritzar els discs, ja que, en cas de fallada, puguin ser reemplaçats per nous. Comptar amb dos Spare Disks també ens incita a usar monitorització.

3.6- Opció de backup?

Amb el nostre sistema, escollim l'opció de fer back-ups on-site sense *mirror*. Farem servir Secure Storage System i realitzarem còpies cada 5 dies (2,5% de canvis). Això ens permet tenir un bon marge per si es perden les dades. Mantindrem al sistema còpia dels últims dos backups és a dir, còpia dels últims 10 dies.

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

Ens seria òptim contractar una línia per fer backups, ja que tenim moltes dades a guardar. Per tal de fer un backup o per recuperar les dades necessitem un *bandwidth* alt. Hem pensat que la millor opció seria contractar els 10 Gbps que ens ofereix l'enunciat.

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?**

Podem recuperar la versió de fins a 10 dies enrere ja que fem un *backup* cada 5 dies i al sistema mantenim fins a dues versions.

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

Si un disc es trenca, les dades no es perden gràcies al sistema de RAID 10 i es trigarà a recuperar-se el temps que es tardi a habilitar els discs sobrants que hi ha a cabina. Abans de que es trenqui un disc, si *spare disk* detecta que està a punt de fallar (el 70% dels casos) el negoci no s'ha d'aturar, es podrà seguir treballant mentre es fa la copia de les dades entre el disc a punt de fallar amb un dels sobrants que tenim a cabina.

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

Hi podria haver problemes de servei en el cas que fallin dos discs amb les mateixes dades, en aquell cas s'hauria d'aturar per a reconstruir les dades en un nou disc.

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

El servei de subministrament elèctric que tenim amb Mordor Colocation Center ens garanteix un 99,982% de *uptime* i ens ofereix dues línies d'entrada d'electricitat, per tant si una fallés ens entraria l'energia per l'altre.

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

En cas que caigui la línia de xarxa, com que tenim dues línies contractades, al caure una, es passaria l'entrada a l'altre.

- **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

Creiem que amb 10Gbps tenim suficient per a el restabliment de dades.

Totalitat de les dades:

Tenim un total de 500.000GB de dades.

Per a obtenir el temps de recuperació dividim la quantitat de dades entre l'ample de banda que tenim per a recuperació.

$$\begin{aligned} \text{Total de dades} &= 500.000\text{GB} = 500.000\text{GB} * \frac{8 \text{ bits}}{1 \text{ Byte}} = 4.000.000\text{Gb} \\ \text{Temps de recuperació} &= \frac{4.000.000\text{Gb}}{10 \text{ Gb}} = 400.000 \text{ segons} = 4,629 \text{ dies} \end{aligned}$$

Ens calen 4,629 dies per a recuperar el total de les dades.

En quant al cost per GB és de 0,07€. Això suposa un cost total de:

$$Cost\ total = 500.000GB * \frac{0,07€}{GB} = 35.000€$$

1% de les dades:

Tenint en compte que necessitem 4,629 dies per a recuperar la totalitat de les dades, per a recuperar el 1% de les dades necessitem:

$$Temps\ de\ recuperació = 4,629\ dies * 0,01 = 1,1111\ hores$$

En quant al cost, sabent que el cost total és de 35.000€, per el 1% de les dades:

$$Cost\ total\ 1\% \text{ dades} = 35.000€ * 0,01 = 350€$$

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:

Segons l'enunciat la línia sol caurà menys d'una hora cada 18 mesos, d'entre 1 i 3 hores cada 3 anys i de més de 3 hores (amb un màxim de 9 hores) un cop cada 6 anys. La probabilitat de caiguda mensual és de:

$$Prob.\ de\ caiguda\ mensual = 1h/18mesos + (2h \pm 1h)/36mesos + (6 \pm 3h)/72mesos$$

$$Prob.\ de\ caiguda\ mensual = 14 \pm \frac{5h}{72mesos} = 14 \pm \frac{5h}{52600h}$$

La línia està caiguda entre 9 i 19 hores de cada 72 mesos. Per tant el percentatge de downtime és de:

$$Downtime\ (\%) = \left[\frac{9}{52600} * 100, \frac{19}{52600} * 100 \right]$$

$$Downtime\ (\%) = [0'017\%, 0'03\%]$$

En el cas de dues línies, la % de no estar disponibles les dues alhora:

$$Downtime\ (\%) = [0'017^2, 0'03^2]$$

$$Downtime\ (\%) = [0'000289, 0'0013]$$

Al no tenir cap SLA, no tindríem cap penalització degut a la caiguda de xarxa de dades.

Fallada de disc

El nostre sistema consta d'un RAID 10, per tant si falla un disc es pot copiar de l'altre disc que conté les mateixes dades, no caldria reconstruir-ne les dades.

Amb SMART podem detectar en el 70% dels casos quan un disc està a punt de fallar i evitar que ho faci clonant el disc en un disc lliure (*spare disk*) destinat a aquests casos.

En total tenim 184 discs amb una probabilitat de fallada de 0,45% anual per disc, ja que els discs WD Gold S768T1D0D són SSD Enterprise MLC-TLC. Gràcies al sistema mencionat anteriorment, només el 30% de fallades necessiten una reconstrucció de disc.

$$184 \text{ discs} * 0,45\% \frac{\text{fallada}}{\text{any}} * 30\% \text{ reconstruir} = 24,84\% \text{ de probabilitat de fallada en un any}$$

Això es tradueix en la probabilitat de que hi hagi una fallada de disc en que s'hagi de reconstruir d'un 99,36% en 4 anys. No és gaire alta i no tindria un gran impacte en l'empresa en el nostre sistema de discs, la còpia d'un disc a un altre serà força ràpida.

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, ...)

En quant al possible creixement en la nostra empresa podria ser que s'haguessin d'emmagatzemar i analitzar més vídeos, com generar informes d'aquests amb més freqüència o en més profunditat.

Un increment del 20% de volum de negoci, en la capacitat dels vídeos passaria de 500TB a 600TB, on en tenim 700TB disponibles, no suposaria cap problema.

Respecte el IOPS i la xarxa, en tenim més de les que requerim, per tant tampoc no serà un impediment per créixer. En cas d'un creixement massiu, gràcies a tenir suficient pressupost podríem afrontar-lo ampliant el sistema.

Creiem que estem preparats per aquest possible creixement en aquests diferents àmbits del sistema que en cas no estar-ho ens obligarien a modificar la infraestructura havent d'invertir una gran quantitat de diners.

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...

Aquests són aspectes en els que apostaríem en l'inversió:

- ➔ Augment de la capacitat d'emmagatzematge de vídeos i processar-los, que implicaria tenir molts més discs tenint en compte també els discs per rèpliques i cabines extres necessàries...
- ➔ Canviar l'estratègia de RAID per millorar-lo a un RAID 51 per exemple, sabent que implica tenir més discs per mantenir les mateixes dades.
- ➔ Augmentar l'ample de banda de SAN per a suportar un major tràfic en la xarxa (sistema més escalable).
- ➔ Augmentar el *bandwidth* del backup en cas que volguem reduir el temps de transmissió d'aquest.

Considerant el nostre escenari, hem disposat d'un pressupost elevat per al disseny del CPD. Podríem haver adquirit tots els serveis disponibles i s'hauria ajustat al pressupost però hem intentat fer el disseny de manera eficient intentant maximitzar el major benefici possible a l'empresa. És per això que no hem realitzat un mirroring en el *backup*, com també no tenir un major ample de banda en el SAN i pel *backup*.