

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Garcia Estevez, Daniel
Gonzalez Valdivia, Guillem
Escenari VM1**

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	1124U
Alçada Rack (en Us)	42U
Consum	2300,8kW
Sobreprovisionament d'electricitat	10%
Nombre de servidors	540
Diners Totals	€35.000.000,00
Diners gastats	€21.500.000,00
taula 2: Elements que escolliu vosaltres	
Elements de mirror i backup	
GB a emmagatzemar al backup	2000000
Dies entre 2 backups	1
Còpies senceres a mantenir	4
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	
Opció escollida (1:MOCOSA, 2: CPDs Céspedes, 3: Mordor)	1
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	2
Segon proveïdor? (0=NO, 1=SI)	0
SAN? (0=no, 1=8Gbps, 2=16Gbps, 3=32Gbps, 4=64Gbps, 5=128Gbps)	3
Cabina de discos	
Opció Disc principal (Entre 1 i 10)	8
Nombre de discos a comprar	4032
Opció cabina de discos (Entre 1 i 6)	4
Nombre de Cabines	112
Cabina de discos 2 (cas de fer servir dos tipus)	
Opció Disc (Entre 1 i 10)	2
Nombre de discos a comprar	<u>216</u>
Opció cabina de discos (Entre 1 i 6)	5
Nombre de Cabines	6
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

TAULA 3: OPEX	anual	cinc anys
Consum energètic (hardware només)	€454.743,53	€2.273.717,65
Empresa de Housing escollida	MOCOSA	
Cost Housing (inclou electricitat addicional)	€505.423,06	€2.527.115,29
Off-site: empresa escollida	Take the tapes and run	
Cost mirror	€0,00	€0,00
Cost backup	€187.600,00	€938.000,00

Cost Bandwidth provider	€15.120,00	€75.600,00
TAULA 4: CAPEX	Cost	
Diners gastats en servers, xarxa, etc	€21.500.000,00	
SAN	€5.001.244,00	
Sistema emmagatzematge	€2.205.224,00	
TAULA 5: AJUST AL PRESSUPOST		
Opex a 5 anys, total	€5.814.432,94	
Capex a 5 anys, total	€28.706.468,00	
Despeses totals a 5 anys	€34.520.900,94	
Diferència respecte al pressupost	€479.099,06	

2.-Anàlisi de necessitats

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

No utilitzem el nombre de clients ja que en el nostre cas un client pot tenir tantes màquines virtuals com vulgui, així doncs, com la previsió de màquines virtuals es de 2000, i tenim que la capacitat de la nostra màquina virtual es de 4GB de dades en el servidor y 1 TB de dades en el servidor de disc. Tenim el següent càlcul:

$$2000 \text{ VMs} \cdot (1000 \text{ GB}) = 2\,000\,000 \text{ GB} = 2000 \text{ TB}$$

Per tant, necessitem 2000 TB d'emmagatzematge en cru, sense cap sistema RAID.

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

Considerem que les 2000 màquines virtuals que tenim de cota màxima poden fer alhora 6000 kbps de lectura i 4000 kbps d'escriptura, per tant tenim:

$$(2000 \text{ kbps} \cdot (6000 \text{ kbps} + 4000 \text{ kbps})) = 20\,000\,000 \text{ Kbps} = 2\,500\,000 \text{ KBps}$$

$$2\,500\,000 \text{ KBps} / 4 \text{ KBps} = 625\,000 \text{ IOPS}$$

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

Considerem que totes les màquines virtuals poden estar actives alhora i per tant cada una pot estar interactuant amb l'exterior, així doncs tenim que el tràfic de les VM amb l'exterior es:

$$2000 \text{ VMs} \cdot 4000 \text{ Kbps} = 8\,000\,000 \text{ Kbps} = 8 \text{ Gbps} = 1 \text{ GBps}$$

També hem de considerar el tràfic que hi ha cada vegada que una VM migra a un altre servidor físic, això implica que per moure una VM tan sols hem de moure el que ocupa una VM en el servidor, les dades de una VM es troben centralitzades a un servidor de discos. Per tant, el càlcul del tràfic que suposa la migració de VM es:

$$(5 \text{ VMs/minut} \cdot 4 \text{ GB}) \cdot (1 \text{ minut}/60 \text{ s}) = 0'33 \text{ GBps} = 2'64 \text{ Gbps}$$

Així doncs el tràfic total amb el client es:

$$0'33 \text{ GBps} + 1 \text{ GBps} = 1'33 \text{ GBps} = 10'64 \text{ Gbps}$$

2.4- Tràfic amb el disc:

El tràfic amb el disc es, el nombre de màquines virtuals que el nostra sistema suporta per la quantitat de tràfic de escriptura i lectura que fa cada maquina virtual. Obtenim els següents resultats:

$$(2000 \text{ VMs} \cdot (6000 \text{ kbps} + 4000 \text{ kbps})) = 20\,000\,000 \text{ Kbps} = 2\,500\,000 \text{ KBps} = 2'5 \text{ GBps}$$

$$= 20 \text{ Gbps}$$

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

Si fem la suma del tràfic amb el client, el tràfic entre servidors i tràfic amb el servidor de disc obtenim que la suma es mayor que 20 Gbps i per tant, sobrepassen la línia.

$$\text{Tràfic total} = 1'33 \text{ GBps} + 2'5 \text{ GBps} = 3'83 \text{ GBps} = 30'64 \text{ Gbps}$$

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

Com hem vist a les preguntes 2.1 i 2.2 tenim que per guardar les 2000 màquines virtuals necessitem com a mínim 2000 TB de dades amb una pressió de 625 000 IOPS.

Per prendre les decisió necessàries considerem els següents aspectes:

- Com que és un sistema de virtualització i que els nostres clients tenen en aquestes màquines els seus negocis creiem que és necessari que les màquines virtuals poden accedir a les dades de forma ràpida. Així mateix, veiem que fem 6000 Kbps de lectura, que representa 64,8 GB per màquina virtual (129,6 TB cada dia).
- Les dades han de poder ser fiables i segures i, per tant, volem un sistema que ens introdueixi redundància i a més en cas de fallada un ràpid accés a les dades.
- Definim que volem tindre un marge de creixement del 50%, és a dir, volem que el nostre sistema pogués créixer fins a les 3000 VMs. Fent que pel nostre escenari les dades canviïn a les següents:
 - Una capacitat de 3000 TB
 - Una pressió de 937500 IOPS

Una vegada aclarit les nostres necessitats hem decidit per la següent configuració:

- Decidim tindre un sistema RAID 51, ja que ens ofereix redundància i ràpida recuperació en cas que ens falli un disc.
- Decidim utilitzar discos SSD a causa de l'alta pressió de IOPS i que volem que el sistema sigui ràpid.
- Utilitzarem una cabina amb opció de spare disk, fer servir 2 discos amb aquesta configuració. Això és degut al fet que volem tenir una protecció addicional a fallades, volem evitar un downtime gran.
- Agafem l'opció 4 de cabina com que volem una alta capacitat de discos (36 slots) i no necessitem un SSD per fer buffering.
- La configuració del RAID serà en forma de 3 clústers RAID 51 dos de 12 discos i un de 10 discos, amb un total de 34 discos més dos de spare disc. Per tant, ens queda 14 discos de dades, 20 de redundància i 2 spare disc.
- Agafarem l'opció 8 de disc, ja que volem un SSD enterprise que ofereix protecció contra errors. A més a més, volem un IOP gran i un consum relativament baix. D'altra banda, hem de ser conscients del preu.
-

Per tant, amb la configuració detallada tenim que necessitem:

1. 112 cabines de tipus 4
2. 4032 discos de tipus 8

En la qual ens queda una capacitat de 3010,56 TB de la qual 66,43 % està ocupat permeten un creixement del 33,57 %.

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

Com hem pogut comprovar en la pregunta 2.5, el tràfic total que tenim en la nostra LAN és de 30,64 Gbps i la línia contractada és de 20 Gbps. No es pot contractar una línia amb més velocitat ens veiem obligats a contractar un servei de SAN per migrar tot el tràfic amb el servidor de disc a una xarxa SAN.

Per tant, en el nostre cas ha de ser assumible perquè si no el sistema tindrà una gran pèrdua de rendiment, però pel que fa al nostre cas es tracta d'un cost assumible.

Com veiem a la pregunta 2.4 el tràfic amb el servidor de disc és de 20 Gbps. En conseqüència, de les opcions que ens donen a escollir (8 Gbps, 16 Gbps, 32 Gbps, 64 Gbps i 128 Gbps) escollirem l'opció de 32 Gbps, la qual ens dóna 12 Gbps de maniobra, el qual suposa un marge de creixement de 1200 VM, un creixement del 60%, el qual està dintre del marge de creixement calculat amb el servidor de disc.

3.3.- Posem un *mirror*?

No, com que tenim una gran quantitat de dades (2000 TB) i una xarxa limitada, per motius de pressupost ens suposa inviable aquesta opció, encara que és una bona idea.

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

L'empresa que s'ha decidit per fer housing és Modular Containers S.A. (MOCOSA). S'ha decidit agafar aquesta pels següents motius:

- No es requereix un centre d'alta seguretat.
- Es pot acceptar certs temps de downtime a causa del baix cost del SLA
- Es tracta d'un pressupost ajustat

Per tant, aquesta empresa ens ofereix un servei barat, que s'adequa a les nostres necessitats i que ens ofereix tant un servei de monitoratge per poder canviar els spare disc com un servei de backup on-site.

3.5- Posem monitorització?

Sí, ja que necessitem que algú estigui pendent del spare disc i és una despesa que podem assolir perfectament de cara a tenir el servei més cobert.

3.6- Opció de backup?

Per tal de decidir quina opció de backup escollim he de tindre els següents punts:

- Tenim una quantitat de dades important (2000 TB sobredimensionat a 3000 TB)
- Tenim 86,4 TB d'escriptura diària:

$$\text{Escriptura diària} = 2000 \text{ VM} \cdot 4000 \text{ Kbps} \cdot 86400 \text{ s} = 691\,200\,000\,000 \text{ Kb} = 86'4 \text{ TB}$$

Com la quantitat de dades que tenim és bastant important, però a la vegada hem decidit que necessiten tindre un sistema que ens permet tindre una recuperació ràpida optem per la següent configuració de backup:

- Backup diari
 - Decidim que sigui diari com que fem escrivim 86,4 TB per dia el qual suposa el 4,32 % de les dades totals. Un percentatge prou alt per decidir fer el backup diàriament i en cas que haurem de recuperar amb el backup tan sols perdrem el 4,32 % de les dades.
- Guardarem una còpia de l'últim dia, l'última setmana, l'últim mes i l'últim any (un total de 4 còpies)
- Backup on-site, ja que a la pressió que ja existeix a la xarxa externa i alt cost d'un servei de mirròr hem decidit tenir un backup on-site.
- Shadow copy, com que no disposem del servei de mirròr i volem fer backups on site, creiem que la millor manera de poder fer backups i alhora tindre una prompta recuperació és tenir un sistema de shadow copy.

Per tal d'implementar el shadow copy necessitarem mantenir almenys una còpia, és a dir, 3010,56 TB, com que no fa falta cap mena de redundància agafem una configuració d'alta capacitat amb disc dur HDD i amb RAID 0. Per tal, que no hi hagi congestió agafarem una cabina amb SSD suport. Per tant, la configuració és la següent:

- 6 cabines de tipus 5
 - 216 discos de tipus 2
- Contractarem el servei de Take the tapes and run, ja que és el que s'acosta més a les nostres necessitats, el servei de backup al núvol suposa un cost addicional que per pressupost no podem assumir, encara faria que el bandwidth que necessiten de cara a l'exterior sigui molt més gran.

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

Com que no tenim servei de mirròr i el backup que s'ha contractat no es transporta a través de la xarxa el tràfic amb l'exterior és tan sols el tràfic que tenen les màquines amb l'exterior i el tràfic quan una màquina virtual migra a un altre servidor, que com es pot veure a la pregunta tenim que el tràfic total és de 10,64 Gbps.

Per tant, amb aquestes dades decidim tindre 2 línies agregades 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?**

Com fem un backup diari i guardem còpies de l'últim dia, l'última setmana, l'últim mes i l'últim any. La versió la pots recuperar d'aquestes 4 versions.

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

Suposem que aquest disc no s'ha pogut detectar amb la tecnologia SMART que es trencava. Com que estem en RAID 51, tenim un mirròr de les dades i, per tant, podem accedir a l'instant a les dades i puc recuperar el 100% de les dades que hem perdut. El servei no queda aturat.

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

No, ja que com hem dit tenim un RAID 51 que ens permet accedir i recuperar les dades de forma immediata, a més a més comptem amb un shadow copy de les dades, que en cas que es necessiten també podem accedir a les seves dades.

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

En el cas que caigui la línia elèctrica dependrà d'on tinguem el nostre houting. En el nostre cas no tenim més d'una línia elèctrica, el que tenim és un generador que ens assegura 48 hores de subministrament en cas de caiguda.

Llavors, en el pitjor dels casos tenim que tenim 1 hora de caiguda cada dos anys i 6 hores de caiguda cada quatre anys. Per tant, la probabilitat que tinguem una caiguda d'una hora en cinc anys és de 2,5 ($5 \text{ anys} / 2 \text{ anys} = 2,5$) la qual cosa implica, en el pitjor dels casos tindrem 3 caigudes d'1 hora en els cinc anys.

La probabilitat que tinguem una caiguda de 6 hores en cinc anys és d'1,25 ($5 \text{ anys} / 4 \text{ anys} = 1,25$), és a dir, en el pitjor dels casos tindrem 2 caigudes de 6 hores.

Així doncs, en cinc anys podem tindre 2 caigudes de 6 hores i 3 caigudes d'1 hora, un total de 15 hores sense línia elèctrica. Amb el subministrament del generador ens cobrim les caigudes de la línia elèctrica.

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

En cas que caigui la línia de xarxa no tenim una segona per tal de suplir quan la principal caigui, així doncs, si la xarxa cau, el servei caurà.

Tenim que la probabilitat que caigui la xarxa una hora en 5 anys és de 3,33 ($5 \text{ anys} / 1,5 \text{ anys} = 3,33$), és a dir, en el pitjor dels casos tindrem 4 caigudes d'una hora en 5 anys.

També ens diuen que la probabilitat que caigui la xarxa 3 hores en 5 anys és d'1,67 (5 anys / 3 anys = 1,67), és a dir, en el pitjor dels casos tindrem 2 caigudes de 3 hores.

I finalment ens diuen que la probabilitat que caigui la xarxa 9 hores cada 5 anys és de 0,83 (5 anys / 6 anys = 0,83), és a dir, en el pitjor dels casos tindrem 1 caiguda de 9 hores.

Per tant, en 5 anys podem tindre 4 caigudes d'una hora, 2 caigudes de 3 hores i una caiguda de 9 hores, que suposa un total de 19 hores en 5 anys.

- **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**

Tenim que per cada cabina tenim 14 discos de dades, que suposa 26,88 TB per cabina de disc. Tenim una ocupació del 66,43 % en dades. Per tant, tenim un volum de dades de 26,88 TB · 112 cabines · 0,6643 = 1 999,92 TB

Tenim que les discos poden fer 540 000 / 205 000 IOPS (limiten les escriptures) i sabem que en un sistema RAID 51 per cada escriptura hem de fer 4 lectures i 4 escriptures, es a dir, podem fer 205 000 IOPS / 4 = 51 250 IOPS.

Per tant, 51 250 IOPS per disc · 112 cabines · 34 discos per cabina · 4 KB = 780 640 000 KBps = 780,64 GBps = 6245,12 Gbps

En el nostre cas disposem d'un SAN de 32 Gbps que ens limita, en conseqüència, farem a 32 Gbps la recuperació.

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

En cas de perdre l'1 % de les dades haurem de recuperar 1 999,92 TB · 0,01 = 19,9992 TB a una velocitat de 32 Gbps suposa 80 ms de recuperació (19,9992 TB · 0,004 TBps = 0,0799968 segons = 80 ms) i un cost de 0,0056 € (0,000022 hores · 250 €/hora = 0,0056 €)

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

En cas de perdre el 100% de les dades haurem de recuperar 1 999,92 TB · 1 = 1 999,92 TB a una velocitat de 32 Gbps suposa 7,99968 segons (1 999,92 TB · 0,004 TBps = 7,99968 segons) i un cost de 0,55 € (0,0022 hores · 250 €/hora = 0,55 €).

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:

En l'únic cas on el servei s'aturaria seria en el cas d'una caiguda de la xarxa, on hem calculat que en el pitjor dels casos en 5 anys tindrem 19 hores on no hi hauria servei. Complint amb el SLA tindrem que el cost d'aquestes caigudes seria:

$$\text{Costos xarxa SLA} = 19 \text{ hores} \cdot 250 \text{ €} = 4.750 \text{ € cada 5 anys}$$

Per tal, d'evitar aquest cost hauríem d'invertir en un housing que ens doni un segon proveïdor de xarxa, que en aquest cas seria canviar a l'empresa Mordor Colocation Center, amb un increment del pressupost de 853.942,39 € més en 5 anys. I la línia addicional ens suposaria un 40% del cost de la principal, la qual cosa implica 30.240 € més en 5 anys. Un total de 884.182,39 € més d'inversió en 5 anys.

Com veiem, en aquest cas, tan sols recomanaria fer aquesta inversió per motius de qualitat, és a dir, si volem garantir el nostre client que no tindrà cap downtime.

Fallada de disc

Com hem mencionat abans, en cas de fallada d'un disc, no tindrem cap downtime i, per tant, no suposa cap cost que ens falli un disc, tan sols, el cost de comprar un de nou.

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, ...)

Com hem mencionat abans s'ha fet un sobre aprovisionament de tal manera que puguem tindre un 50% més de màquines virtuals (són unes 1000 VM). Per tant, tal com s'ha dissenyat podem augmentar perfectament un 20% en el volum de tot el negoci, ja que en volum de dades encara podem créixer un 30 %. Tant en SAN com en bandwidth tenim molt marge de creixement.

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...

El que valdria la pena invertir és en més seguretat pel CPD, ja sigui seguretat física com més redundància sigui amb línies elèctriques o de xarxes per evitar un downtime que empitjoraria l'experiència de l'usuari.

També caldria invertir en la LAN, ja que encara que amb el SAN podem continuar donant servei, la LAN ja es queda curta amb tot el tràfic del servidor, per tant, recomanaríem una xarxa LAN amb més velocitat o si no n'hi ha, augmentar la velocitat del SAN i la xarxa amb l'exterior.

Així mateix, creiem que un sistema de mirror ajudaria a millorar els serveis en termes de seguretat amb les dades i menys manteniment per la nostra part amb la shadow copy.