

Activitat EBH

Emmagatzematge, *backup* i *housing*

**Borrás i Duarte, Jorge
El Aichouni Jouied, Tarik**

CRBMIX

Data: 28/09/2023

1.-Descripció bàsica

Escenari CRBMIX

Nombre de Us	1140U
Alçada Rack (en Us)	42U
Consum	1300,kW
Sobreprovisionament d'electricitat	8%
Nombre de servidors	360
Diners Totals	€14.000.000,00
Diners gastats	€11.500.000,00

Elements de mirror i backup	
GB a emmagatzemar al backup	700000
Dies entre 2 backups	7
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)	3
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)	1

Cabina de discos	
Opció Disc principal (Entre 1 i 10)	3
Nombre de discos a comprar	101
Opció cabina de discos (Entre 1 i 6)	4
Nombre de Cabines	3
Cabina de discos 2 (cas de fer servir dos tipus)	
Opció Disc (Entre 1 i 10)	9
Nombre de discos a comprar	63
Opció cabina de discos (Entre 1 i 6)	5
Nombre de Cabines	2
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)	€236.938,18	€1.184.690,89
Empresa de Housing escollida	MOCOSA	
Cost Housing (inclou electricitat addicional)	€440.081,45	€2.200.407,27
Off-site: empresa escollida	Take the tapes and run	
Cost mirror	€0,00	€0,00
Cost backup	€47.514,29	€237.571,43
Cost Bandwidth provider	€1.512,00	€7.560,00

TAULA 4: CAPEX	Cost
Diners gastats en servers, xarxa, etc	€11.500.000,00
SAN	€1.306.314,00
Sistema emmagatzematge	€158.685,00

TAULA 5: AJUST AL PRESSUPOST	
Opex a 5 anys, total	€3.630.229,59
Capex a 5 anys, total	€12.964.999,00
Despeses totals a 5 anys	€16.595.228,59
Diferència respecte al pressupost	-€2.595.228,59

2.-Anàlisi de necessitats

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

El mínim nombre de GB que cal emmagatzemar és la suma del mínim necessari per part de la partició de la universitat i el total dels grups de recerca.

En aquest apartat només cal sumar el mínim necessari per a la universitat amb el mínim necessari per als grups de recerca:

$$400 \text{ TB (partició universitat)} + 10 \text{ (grups de recerca)} * 30 \text{ TB/grup de recerca} = 700.000 \text{ GB}$$

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

Assumint que 1 IOPS utilitza 4KBps, només cal trobar el tràfic de disc que generen tots els racks per segon i multiplicar els 2 valors:

$$(11 + 18) \text{ racks} * \frac{8 \text{ MB/s}}{1 \text{ rack}} * \frac{10^3 \text{ KB/s}}{1 \text{ MB/s}} * \frac{1 \text{ IOPS}}{4 \text{ KB/s}} = 58000 \text{ IOPS}$$

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

En aquest apartat només cal trobar el nombre de racks i multiplicar-ho per l'ample de banda amb el client que necessita cada rack (i no oblidar de passar-ho a Megabits per segon).

$$(11 + 18) \text{ racks} * \frac{4 \text{ MBps}}{\text{rack}} * \frac{8 \text{ Mbps}}{1 \text{ MBps}} = 928 \text{ Mbps}$$

2.4- Tràfic amb el disc:

En aquest apartat cal trobar el nombre de racks i multiplicar-ho per l'ample de banda amb el disc que necessita cada rack (i no oblidar de passar-ho a Megabits per segon).

$$(11 + 18) \text{ racks} * \frac{8 \text{ MBps}}{\text{rack}} * \frac{8 \text{ Mbps}}{1 \text{ MBps}} = 1856 \text{ Mbps}$$

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

L'amplada de banda mínim equival a la suma entre l'amplada de banda mínim amb el client sumant-hi l'amplada de banda mínim amb els discs:

$$928 \text{ Mbps} + 1856 \text{ Mbps} = 2784 \text{ Mbps}$$

$$2784 \text{ Mbps} * \frac{1 \text{ Gbps}}{1000 \text{ Mbps}} * \frac{1}{2 \text{ Gbps}} = 139.2 \%$$

Com que tenim una xarxa LAN de 2Gbps, no tenim suficient amb l'amplada de banda actual i tenim necessitats per sobre la capacitat actual (caldrà afegir una SAN).

3.-Decisions preses

Els detalls particulars de l'escenari CRBMIX requerien prendre un conjunt de decisions costoses perquè el producte final fos funcional i s'adequés correctament a les necessitats del client. No obstant, el pressupost de 2.500.000,00 euros (14 milions menys els 11,5 milions d'euros gastats prèviament) deixen clar que no es pot arribar a una solució satisfactòria sense superar el pressupost establert.

En comptes d'intentar realitzar una solució insuficient mantenint el mínim pressupost possible, hem decidit realitzar la mínima solució viable possible. En altres paraules, realitzar la millor solució als requisits de l'escenari balancejant el pressupost a la baixa quan fos possible. Les decisions preses entre els apartats 3.1 i 3.7 reflecteixen aquesta realitat.

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

Les necessitats del nostre cas es redueixen a la diferència entre, per una banda, el SLA dels servidors de la Universitat i els del grup de recerca i per l'altra banda els dos tipus de dades que es tracten en el nostre CPD (el 70% de les dades pràcticament no s'utilitzen, mentre l'altre 30% tenen una alta càrrega i demanda).

Degut als 2 tipus de dades que emmagatzem i que el servidor coneix quines són, hem decidit separar els requisits en dues parts:

- Una part dedicada a les dades que no s'usen sovint: amb poca capacitat d'accés, però que contingui el 70% de la capacitat d'emmagatzematge.
- Una part dedicada a les dades que s'usen amb freqüència: amb molta capacitat d'accés, però que contingui el 70% de la capacitat d'emmagatzematge.

Per la primera part hem decidit escollir únicament HDDs i per la segona SSDs. Per la partició de HDDs hem aproximat que rebria el 20% de les IOPS totals, mentre que la partició SSD rebria el 80%. Finalment hem decidit multiplicar tant la capacitat com les IOPS per un factor de 1.5, per garantir una capacitat de creixement mínima i un marge d'error suficient (l'estat inicial estaria, per tant al 66% de capacitat i al 66% d'IOPS de la màxima permesa). També hem decidit rebutjar els discs de tipus "Consumer" per la falta de confiança amb i menor vida útil d'aquests tipus de discos (sobretot perquè en la majoria dels casos no produïen solucions més barates).

Per tant, busquem: discos HDD i SSD capaç d'emmagatzemar 735 i 315 TB (en total 700 TB * 1,5) i amb una velocitat de 17400 i 69600 IOPS (58000 IOPS * 1,5) respectivament. Finalment, ens hem decidit per un sistema RAID 5 (clústers de 5 + 1 de partició), ja que és barat però ràpid i amb bona resistència davant la fallada de disc. RAID 10, 51, 6 i 61 eren massa cars en comparació mentre que RAID 1 i 10 no proporcionava suficient seguretat.

A continuació presentem els discos que hem considerat (la resta eren consumer, tenien un consum massa elevat o un preu desproporcionat):

	Disks per IOPS	Disks per Capacitat	Discs Totals	Consum Total	Preu Total
HDD Opció 2	54	65	65	507 W	28600 €
HDD Opció 3	62	89	89	845.5 W	26250 €
SSD Opció 8	3	198	198	1782 W	61752 €
SSD Opció 9	3	51	51	612 W	66435 €
SSD Opció 10	2	100	100	800 W	63000 €

Seguidament vam realitzar una exploració de les cabines. Com teníem intencions d'afegir Spare Disks vam descartar l'opció 1. Vam considerar que si tenim un sistema RAID 5 calia afegir seguretat a través d'aquesta tècnica econòmica. També vam descartar l'opció 6 per no ser compatible amb RAID 5. A continuació teniu l'exploració de com les diferents combinacions afecten el cost (inclosa l'electricitat), quantes cabines serien necessàries i l'espai lliure per cabina:

	Cost Cabina tipus 2	Nº Cab necessàries	Espai Lliure	Cost Cabina tipus 4	Nº Cab necessàries	Espai Lliure
HDD Opció 2	65.143,97 €	3	2.3	60.085,98 €	2	3.5
HDD Opció 3	69.029,40 €	4	5.5	64.497,20 €	3	6.3
SSD Opció 8	192.766,47 €	9	2	177.592,49 €	6	3
SSD Opció 9	110.271,06 €	3	7	105.213,07 €	2	10.5
SSD Opció 10	168.436,75 €	5	4	108.289,97 €	3	2.6

**** NOTA:** en aquesta taula no contemplem l'opció 3 i 5 perquè el cost és pràcticament el mateix que la 2 i 4, sent 3 i 5 una mica més cares per les caches SSD afegides.

Vam decidir escollir l'opció 2 pels HDD i la 9 per les SSD, perquè tenien un preu menor a la resta i/o perquè comptaven amb el major nombre d'espai lliure per cabina possible. En el cas de les cabines hem escollit la 4a en el cas de les HDD i la 5a opció per les SSD. D'aquesta forma garantim tenir l'opció més econòmica mentre tenim marge de creixement per si el client decideix afegir més discos per incrementar la capacitat o la velocitat de lectura/escriptura (sobretot perquè el mateix enunciat avisa que la configuració es pot d'haver d'ajustar tant com sigui possible en el curt termini). En el cas de la cabina 5, hem escollit aquesta per les SSD per fer més ràpids els accessos i minimitzar el nombre d'accessos realitzats, allargant la vida útil en general sense afegir gaire cost (una diferència de +1000 € aproximadament).

Adicionalment, afegirem 2 Spare Disks per cabina per millorar la seguretat del RAID 5 i protegir-nos sobre fallades de disc. També afegirem un total de 6 discos SSD per emmagatzemar els snapshots que anem fent.

Per tant, comprarem un total de 3 cabines de tipus 4 i 2 de tipus 5. També comprarem 101 discos HDD de tipus 3 ($95 + 2 * 3$). Finalment també comprarem 63 discos SSD de tipus 9 ($55 + 2 + 6$).

SOLUCIÓ ALTERNATIVA

Tenint en compte la divisió entre Universitat i grups de recerca, una altra solució hauria estat formar 11 particions:

- 1 partició per la universitat amb 420 TB d'HDD i 180 TB de SSD (en total $400 \text{ TB} * 1,5$)
- 10 particions pel grup de recerca amb 31 TB d'HDD i 13,5 TB de SSD ($30 \text{ TB} * 1,5$).

Aquesta solució hauria estat més segura al ser més fàcil aïllar els errors quan es produeixin. A la vegada hauria estat més organitzada i eficient si es pogués agrupar les cabines als racks dels grups que corresponen.

Segons els nostres càlculs, seguint aquesta línia de pensament, la millor solució conté:

- 50 Discos HDD de l'opció 3, repartits en 10 cabines de tipus 1 (part recerca)
- 51 Discos HDD de l'opció 3, repartits en 2 cabines de tipus 4 (part universitària)
- 30 Discos SSD de l'opció 9, repartits en 10 cabines de tipus 4 (part recerca)
- 29 Discos SSD de l'opció 9, repartits en 1 cabina de tipus 4 (part universitària)

Aquesta solució suposava un increment de més de 100.000 €. Hi ha múltiples raons per les quals no s'ha triat aquesta solució. La primera és l'augment considerable del pressupost que ja sobrepassa el límit sense evidenciar una millora de seguretat ni velocitat ni capacitat. A més a més, aquest preu assumeix que s'ha usat RAID 5 en la partició recerca, però això no era recomanable, ja que les cabines de tipus 1 no permeten Spare Disks i per tant la seguretat hagués disminuït considerablement i/o s'hagués invertit més per recuperar-la. Aquesta solució tampoc permetia cachejar la informació de les SSD (fent encara més lent l'accés per aquests usuaris).

Finalment, aquesta solució és molt bona si es pot associar les cabines de cada grup amb els racks de cada grup, però no sabem si aquells racks tenen l'espai necessari (Només sabem que hi ha prou espai en total, però no sabem si està concentrat, per exemple, en 3 racks diferents). Per l'altra banda, cap servidor es connecta directament amb els discos del seu rack a través del TOR switch per culpa de la SAN. Aquesta redirecció del tràfic malgasta la ubicació d'aquestes cabines i augmenta el temps d'accés independentment de si estan al costat o no.

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

Tal com es va calcular en l'apartat 2.5, tenim una pressió sobre la xarxa local de 2784Mbps, però tenim instal·lada una LAN de 2 Gbps (o 2000Mbps). Això comporta que necessitem fer ús d'una SAN per poder funcionar amb les condicions mínimes, ja que la demanda actual sobre la xarxa LAN (sense la SAN) sobrepassa la seva pròpia capacitat i provocaria problemes de rendiment considerables.

Afegir una xarxa SAN bé a costa d'un preu elevat, en el nostre cas encara pitjor al tenir molts servidors actius (360 servidors). Hem escollit l'opció 1 de 8Gbps, perquè ens assegura un bon servei i marge de creixement suficient al menor preu possible.

Cal comentar que la SAN treballaria a una capacitat mitjana d'23,2%, i evidentment, és molt més del que realment necessitem. Desafortunadament, no existeix cap alternativa que s'ajusti millor al nostre cas. Aquesta decisió, malgrat ser necessària, ens condiciona negativament en la tria de components, ja que el cost total ascendeix a 1.306.314 € que suposa més de la meitat del pressupost disponible.

3.3.- Posem un *mirror*?

Inicialment, la idea de posar un *mirror* per les dades del grup d'Universitats va ser ben valorada, perquè s'adequa perfectament a les condicions establertes en el SLA d'aquest grup (garantir un downtime molt petit). Això més endavant es va desestimar pel seu preu que no és assumible i hem optat per opcions molt més econòmiques per adaptar-nos a les condicions establertes i tenint en compte el reduït pressupost.

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

Principalment, ens hem fixat en l'empresa de *housing* que ofereixi una solució barata, però amb la capacitat suficient per suportar les condicions d'uptime que necessitem, donar bon servei de seguretat (no volem ser vulnerables al robatori d'informació important que pot fer perdre moltes hores de recerca).

Amb els models que tenim a disposició no es pot aconseguir, ja que principalment tenim la restricció del SLA d'universitats que fa que hàgim d'utilitzar l'opció més cara i, per tant, gastariem al voltant de 50.000€ addicionals per any.

Si no tenim aquesta restricció en compte, la millor opció seria la de MOCOSA, ja que proporciona condicions de seguretat molt bones que ens assegurin la integritat de les dades, a més d'una reducció de cost considerable. La única pega és que no té cap certificat de uptime pel que correm el risc de no seguir el SLA.

Per últim també disposa de la possibilitat de fer la gestió dels backups a un cost més econòmic cosa que ens pot interessar en el cas de ser viable a posteriori.

3.5- Posem monitorització?

Com els nostres requisits necessiten una qualitat de servei raonablement elevada, ens convé fer servir monitorització, però també hem de tenir en compte el preu. Tenim equips diferents amb usos diferents i alta possibilitat de canvi i per aquest motiu, és de vital importància tenir un seguiment de l'estat de la infraestructura on s'hostegen les nostres dades.

Pel que fa a les opcions que tenim, sembla que l'empresa de housing ens proporciona monitorització a un preu molt econòmic en comparació amb altres solucions (2000€ per container). Després de fer l'anàlisi, ens hem decantat per comprar-ho, ja que proporciona una gran funcionalitat a un preu molt assequible.

3.6- Opció de backup?

Hem plantejat diferents solucions de backup per poder fer front a les dos grans limitacions: el QoS (Quality of Service) i el preu. La base comuna és la d'emmagatzemar tota la informació, fer backups cada 7 dies, mantenir 4 còpies seguint el model, setmana, mes, trimestre, any.

Primera solució:

Com que tenim definits diferents tipus de servei per a diferents grups, aquesta solució tracta d'oferir un recurs personalitzat per cada grup, de forma que es pugui adequar a les seves necessitats. D'aquí n'hem extret tres tipus de backups:

- Backup off-site de totes les dades: Fer un backup off-site de totes les dades (700 TB) amb cintes, ja que proporcionen la relació, quantitat de dades/preu més rendible
- Backup in-site de les dades pels grups de recerca: Aquesta informació (300 TB), emmagatzemada en cintes, considerem que ens ofereix un rendiment suficient, pel fet que, tot i que són cintes i la seva recuperació és més pesada, considerem que amb els spare disk + monitorització + RAID 5 cobrim les fallades més comunes i amb les cintes, els casos més extrems.
- Backup mirror de les dades pel grup d'Universitat: Inicialment hem contemplat l'opció de fer servir mirrors perquè proporcionen una recuperació de les dades molt eficient però el preu és massa elevat.

Segona solució:

Aquesta segona solució és la que finalment hem escollit i dona més prioritat al cost, ja que en el cas dels backups podem arribar a estalviar molts diners. En aquest cas tractem de mantenir la mateixa distribució entre grups però amb modificacions per reduir els costos:

- Backup off-site de totes les dades: Mantenim el mateix enfocament que anteriorment.
- Backup in-site de totes les dades: Ara en comptes de dividir entre grups de recerca o Universitat, incloem tot com a backup en cinta in-site.
- Snapshots diaris: Snapshots diaris de tots els discos. La idea és emmagatzemar 3 dies de snapshots diaris per assegurar un uptime el més gran possible (sobretot perquè en cas). Cada snapshot pesa uns 10,0224 TB per la qual cosa necessitem uns 50 TB extra d'emmagatzament.

$$58000 \text{ IOPS} * 0.5 \text{ OP d'escriptura} * \frac{4 \text{ KB}}{1 \text{ s}} * \frac{1 \text{ TB}}{1.000.000.000 \text{ KB}} * \frac{3600 * 24 \text{ s}}{1 \text{ dia}} =$$

$$= 10.0224 \text{ TB / dia de dades modificades en la partició d'Universitat}$$

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

A causa de les decisions de l'apartat 3.3 i 3.6 hem decidit no contractar cap servei per tenir mirrors i/o backups externs que requereixin cap comunicació extra amb l'exterior. Per tal de mantenir el pressupost al mínim, hem decidit contractar el servei backups on-site amb cintes magnètiques i el transport d'aquestes físicament a través de l'empresa "Take the Tapes and Run".

En l'apartat 2.4 es va calcular que el nostre CPD requereix un bandwidth de 928 Mbps per comunicar-se amb l'exterior. Com no tenim cap altra comunicació, hem decidit contractar l'opció 3 amb dues línies agregades per tenir un bandwidth total de 2000 Mbps. Malgrat que amb una matemàticament en teníem prou, considerem que és massa justa (sobretot si tenim en compte que durant el dia el tràfic pot superar el doble de la nit, superant fàcilment el límit 1 Gbps), així que hem contractat 2 Gbps per tenir uptime garantit inclús si en el futur augmenta la necessitat de comunicar-se amb l'exterior.

4.-Recomanacions als inversors

4.1.- Anàlisi de Riscos (*Risk Analysis*)

- **Hi ha pèrdua d'un fitxer (per error o corrupció). De quan puc recuperar versions?**

Com s'està usant RAID 5, en cas de corrupció es pot recuperar l'arxiu mitjançant la reconstrucció del disc sense haver d'aturar el CPD. No es perdria informació, però sí que es trigarien en mitjana 36 o 24,58 hores en funció de si és tracta d'un disc HDD o SSD.

En cas de ser per error o simplement no es pugui reconstruir, no queda cap altre remei que recuperar l'arxiu de l'últim backup realitzat. Si es recupera per l'snapshot, llavors es perdrien com a màxim 24 hores (fem snapshots cada 24 hores). En cas contrari s'ha de recuperar de l'últim backup de cintes realitzat i en el pitjor dels casos seria la versió de fa 7 dies.

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

Si es trenca un disc, hi ha un 70% de possibilitat que es detecti amb SMART i s'utilitzi un spare disk (no es perden dades ni el negoci s'ha d'aturar). Altrament, com s'està usant RAID 5, es pot reconstruir el disc sense haver d'aturar el negoci i es trigaria entre 36 i 24,58 hores en recuperar el contingut (si parlem d'un HDD o SSD). Cal remarcar que si es tracta d'un HDD, la cabina afectada veurà una reducció del 75% a la velocitat mínima esperada.

Si per casualitat es trenquen 2 discos dintre d'algun dels clústers RAID 5 llavors s'hauria d'aturar el CPD i fer una restauració sencera des de l'últim backup. Si es pot recuperar per part d'snapshot, es podria recuperar en 13,35 hores. Altrament s'ha de recuperar del sistema de cintes de backup in-site i es trigarien 140 hores.

$$(710 * 0,2 + 65000 * 0,8) IOPS * \frac{4 KB/s}{1 IOPS} * \frac{1 TB}{1.000.000.000 KB} * \frac{1 h}{3600 s} = 0,751 TB/h \text{ de recuperació}$$
$$10.0224 TB \text{ per snapshot} * \frac{1 h}{0,751 TB} = 13.34 \text{ hores de recuperació per snapshot}$$

$$700 TB \text{ de backup cintes} * \frac{5 TB}{1 h} = 140 h \text{ de recuperació per cintes}$$

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

Si falla un disc, no s'ha detectat amb SMART (només el 30% dels casos) i s'ha de reconstruir, durant aquest temps la cabina funcionarà a la meitat dels IOPS. Si això es produís a una cabina SSD (més probable que a una cabina HDD) es reduiria la velocitat de disc, però encara seria suficient. En canvi, si això succeís a una cabina HDD els usuaris es veuran afectats i la cabina funcionaria al 75% de la velocitat esperada.

Si fallessin dos o més discos d'un mateix sistema raid llavors s'ha d'aturar el CPD i recuperar el backup sencer des del sistema de cintes o de l'últim snapshot (més detalls a l'apartat 4.1.2).

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

L'empresa d'Housing contractada, "MOCOSA" compta amb un sistema d'alimentació que garanteix el subministrament pic fins a 48h després d'un tall en la línia elèctrica. Considerant que en el pitjor dels casos, com a màxim cau una línia per 6 hores, mai afectarà els nostres serveis.

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

En cas d'una caiguda de la línia de xarxa no tenim cap manera de continuar oferint els nostres serveis i caldria esperar que la línia es restaurés per seguir operant. Això pot succeir menys d'1 hora cada 18 mesos, entre 1 i 3 cada 3 anys i entre 3 i 9 hores un cop cada 6 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:**

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

Considerant que es guarden com a mínim 700 TB de dades, l'1% d'aquestes són 7 TB. En el millor dels casos, aquests 7 TB s'han perdut d'un únic d'un mateix disc HDD (tenen un espai útil de $10 * 0,2 = 8$ TB) i es podria reconstruir sense costos ni sense haver d'aturar els servidors. El procés duraria 36 hores. Similarment, si aquestes dades són de diferents discos en costers RAID diferents, es poden reconstruir gratuïtament i sense aturar el CPD en un període de 36 o 24.567 hores (HDD o SSD).

En el cas que les dades afectin múltiples discs d'un mateix sistema RAID, no es pot fer res més que recuperar l'última versió prèvia guardada en el backup. Si els errors estan continguts únicament en els clústers RAID i es pot recuperar la informació dels snapshots llavors es pot restaurar el sistema en 13,34 hores i al preu de 20.000 € / hora de downtime per la part de recerca es perdrien 266.800 € (per la part d'universitat no ens cobren res extra). Es perdrien 24h de feina com a molt.

Si únicament es pot recuperar la informació del sistema de backup de cintes on-site, on es perdran les operacions realitzades fins fa 7 dies com a màxim. Aquest procés durarà 140 hores i degut a la penalització per part dels grups de recerca costaria 280.000 €.

***Nota: Els detalls sobre els càlculs dels temps de recuperació es troben a l'apartat 4.1.2.*

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

Una pèrdua total de les dades no té cap altre remei que restaurar-les d'un full backup. De nou, s'haurà d'aturar el CPD i de recuperar les dades del sistema snapshots o del de cintes de backups on-site. Aquest procés durarà 13,34 hores amb un cost de 266.800 € o bé 2140 hores amb un cost de 2.800.000 € respectivament.

- **Es produeix un incendi.**

L'empresa de Housing contractada, "MOCOSA" compta amb un sistema antiincendis per reducció d'oxigen. En cas que es detectés foc llavors s'activaria aquest sistema i el nostre CPD quedaria protegit.

- **Smart Detecta un disc que comença a fallar**

En el cas que Smart detectes algun disc que comences a fallar, llavors es copiaria el contingut a un dels dos Spare Disks instal·lats en totes les cabines. El disc que havia començat a fallar seria aïllat i es deixaria d'utilitzar, a l'espera de ser substituït. Tot això ocurriria sense que el funcionament del CPD es veies afectat.

- **Es trenca algun component dels sistemes HVAC o UPS.**

Gràcies a l'empresa de Housing contractada, "MOCOSA" comptem amb una redundància d'N+1 dels sistemes d'HVAC i UPS. En cas que es trenqués un component qualsevol d'aquests sistemes es faria servir el de backup sense que el servei es veies afectat.

- **Es produeix algun desastre que destrueix el Colocation Center del nostre CPD.**

Si el Colocation Center de "Mocosa" quedés destruït, el més probable és que el nostre CPD també quedés fora de funcionament i irreparable. No obstant, gràcies als backups off-site seríem capaços de recuperar la informació de fa 7 dies en el pitjor dels casos.

4.2.- Anàlisi de l'impacte al negoci (*Business Impact Analysis*)

Caiguda de la xarxa elèctrica i/o incendi i Fallada d'algun component HVAC o UPS:

Com hem mencionat en els apartats 4.1.4, 4.1.7 i 4.1.9, aquestes situacions estan cobertes per l'empresa de Housing contractada, "MOCOSA". No tindrien cap impacte sobre el negoci (ni a nivell econòmic ni a nivell de servei).

Caiguda de la xarxa de dades:

Com hem mencionat en l'apartat 4.1.5, no tenim cap cobertura sobre aquest escenari. Degut a la part de grups de recerca perdríem 200.000 € per cada hora que la xarxa de dades estigui caiguda. En mitjana, la xarxa està caiguda 10,83 hores a l'any, costant 216.666 € a l'any (Assumint que la xarxa de dades cau menys d'1 hora cada 18 mesos, entre 1 i 3 cada 3 anys i entre 3 i 9 hores un cop cada 6 anys).

Amb l'empresa de Hosing "MOCOSA" contractada no hi ha cap manera de anticipar aquest escenari. Caldria canviar a "MORDOR" i contractar una segona línia de backup (454054,00 € extres anualment).

Fallada de disc

Donada una fallada de disc detectada amb el monitoratge SMART (70% de probabilitat), es substituiria el disc immediatament sense cap repercussió sobre el negoci. En el cas contrari tampoc suposaria cap despesa econòmica gràcies a RAID 5. S'hauria de reconstruir el disc i durant les properes 24,57 o 36 hores es reduiria la velocitat d'accés depenent si ha fallat un SSD o HDD mentre es realitza aquest procés. En el cas d'HDD la velocitat d'accés a la cabina afectada es reduiria un 25%.

En mitjana fallaran 12,65 discos HDD opció 2 i 0,344 discos SSD opció 9 al llarg dels 5 anys. Arrodonint cap amunt, en total substituir aquests discos suposaria 4550 € i 1545 € respectivament.

Recuperació per Snapshot

Tal com s'ha calculat en l'apartat 4.1.2, en el cas que haguessin de recuperar totes les dades des de l'últim snapshot realitzat, s'haurien d'aturar totes les màquines i iniciar el procés de restauració. Es perdrien fins a un màxim de 24 hores d'operacions i el procés en total trigaria 13,34 hores. Degut a la penalització de 200.000 € / hora del grup de recerca en cas de downtime ens costaria 266.800 €.

Per millorar aquest preu hauríem de contractar un millor proveïdor de backups que minimitzi el downtime per sota de 13,34 hores, col·locant un mirròr per exemple (154.000 € anualment).

Recuperació via Backup in-site o off-site

Tal com s'ha calculat en l'apartat 4.1.2, en el cas que haguessin de recuperar totes les dades des de l'últim backup guardat en el sistema de cintes, s'haurien d'aturar totes les màquines i iniciar el procés de restauració. Es perdrien fins a un màxim de 7 dies d'operacions i el procés en total trigaria 140 hores. Degut a la penalització de 200.000 € / hora del grup de recerca en cas de downtime ens costaria 2.800.000 €.

Si per qualsevol raó s'hagués de recuperar del backup guardat off-site, s'hauria de pagar 240 € extres de la tarifa de transport.

De nou, per millorar aquest preu hauríem de contractar un millor proveïdor de backups: per exemple la segona opció (diferència de 684.584,12 anualment) i/o col·locant un mirròr (154.000 € anualment).

4.3.- Anàlisi del marge de Creixement

Un dels principals aspectes d'aquest cas és la seva heterogeneïtat i dinamisme, per la qual cosa el sistema pot demanar més recursos depenent del moment. El nostre disseny ha tingut en compte aquest concepte, però sempre amb la limitació de preu que se'ns ha imposat. D'aquesta manera, tenim que podem suportar:

- Un creixement en capacitat del 57,84% . Actualment, el nostre sistema pot emmagatzemar un total de 1097, 92 TB (dels 700 TB sol·licitats). A més a més, les nostres cabines compten amb 16 badies buides on si poden col·locar entre 122,88 i 160 TB addicionals (en funció de si són l'opció 3 HDD o bé l'opció 9 SSD).
- Un creixement en xarxa força elevat, principalment per l'ús de SAN que fa que tinguem un creixement de LAN del $2000 \text{ MB} / 928 \text{ MB} = 2,15 \rightarrow 115\%$, i un creixement en xarxa de disc SAN del $8000 \text{ MB} / 1856 \text{ MB} = 4,3 \rightarrow 330\%$ sense tenir en compte operacions relacionades amb backups. En l'àmbit de discos a hores d'ara estem al 55% de capacitat màxima en la part d'HDDs i del 5% a la part de SSDs (els quals es podrien incrementar afegint discos a les badies buides).
- Un creixement insuficient a nivell de servidors, ja que només ens sobren 18 Us, però sí que s'hi podrien col·locar un parell de cabines de discos més.

(1218 Us totals - 1140 Us servidors - 60 Us nostres= 18 U restants).

Per tant, si tenim un increment del 20% en el volum de negoci (que en aquest cas serien més grups de recerca o més demanda en general) podem suportar l'increment tenint en compte que res s'espalla. No obstant qualsevol creixement afectaria negativament els processos de recuperació dades i la restauració de backups (els quals ja són prou lents).

4.4.- Inversions més urgents

Donat el CPD resultant és possible que no hàgim 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 d'en què valdria la pena invertir per millorar rendiment, seguretat o...

Donat aquest cas, podem observar com el nostre disseny ha superat el pressupost inicial, per la qual cosa la inversió més urgent vindria per cobrir aquestes despeses necessàries per poder tenir un sistema mínim funcional. El nostre disseny en aquest cas ha sigut amb la intenció de proporcionar un producte mínim viable que s'adeqüi a les especificacions del cas, és a dir, que pugui funcionar en casos normals sense gaires desviacions del que s'ha proposat. Les grans despeses han sigut donades per l'ús d'equipament (com la SAN) que incrementa molt en cost, però que considerem que és essencial per obtenir un mínim funcionament.

Un cop hem cobert aquesta despesa inicial, hi ha aspectes que es poden millorar molt més per aconseguir un rendiment òptim. Aquestes passen sobretot en afegir redundància i un sistema de recuperació de dades més robust i ràpid. Trobem dins d'aquesta llista, la redundància de bandwidth provider, una millor distribució de RAID per poder suportar fallades amb un millor response-time (per exemple passar de RAID 5 a RAID 10 o 51) i evitar downtime, l'ús de mirror per agilitzar el procés de recuperació de dades en casos on necessitem recuperar moltes dades, millorar la capacitat de xarxa LAN. També seria recomanable la distribució proposada a la solució alternativa de discos i cabines per millorar la seguretat i l'organització, l'adequat-se millor a les necessitats de cada grup (posats a canviar els clústers RAID, aquesta distribució és més apropiada que l'actual).

Finalment, pel que fa a Infraestructura, considerem també que, encara que no tan urgent com les millores especificades anteriorment, s'hauria de plantejar contractar l'empresa de housing MORDOR (que té certificat d'uptime) per assegurar que els recursos físics són tractats de la millor forma per adaptar-nos a les especificacions dels SLA.