

Activitat Ebh

Emmagatzematge, backup i housing

Facultat d'Informàtica de Barcelona

Centres de Processament de Dades

David López

Tardor 2022

Enunciat	5
Condicions de la pràctica	5
Documentació: l'increïble full de càlcul	6
Escenaris	7
Escenari 01: NET 1	9
Escenari 02: NET2	9
Escenari 03: NETPRO	10
Escenari 04: CRB1	12
Escenari 05: CRB2	12
Escenari 06: CRBMIX	12
Escenari 07: ANA1	13
Escenari 08: ANA2	14
Escenari 09: ANA+	14
Escenari 10: VM1	15
Escenari 11: VM2	16
Escenari 12: VM+	16
Escenari 13: EC1	18
Escenari 14: EC2	18
Escenari 15: ECPRO	18
Escenari 16: VineTok1	20
Escenari 17: VineTok2	20
Escenari 18: VineTokBigData	21
Apèndix 1: Disseny d'un sistema d'emmagatzematge	22
Idees bàsiques	22
Models de disc	22
Cabines disponibles	24
Què heu de fer?	25
Apèndix 2: Housing	27
Opció 1: El Data Center en contenidors	27
Opció 2: Colocation Tier 2	28
Opció 3: Colocation Tier 3	28
Què heu de fer?	29
Apèndix 3: Backup	30

Dades off-site	30
Backup on-site	31
Empreses que ofereixen mirròr i/o backup	31
Opció 1: Microworks Apure	31
Opció 2: Monsoon S3	33
Opció 3: Transport de cintes de backup	33
Què heu de fer?	34
Apèndix 4: Consums d'ample de banda i electricitat	35
Connexió amb l'exterior	35
Consum elèctric	36
Apèndix 5: NAS o SAN?	37
NAS	37
SAN	38
Què heu de fer?	38
Apèndix 6: Cada quan falla?	39

Enunciat

El treball consisteix en dissenyar una part d'un CPD, suposant que ja tenim organitzat l'altre part (tipus i número de servidors, racks, xarxa, etcètera). Per tant partim d'una configuració ja dissenyada que anomenarem **escenari**. A partir de l'escenari que escolliu heu de preparar la vostra proposta.

Condicions de la pràctica

La pràctica la fareu, en general, per parelles. Cada parella tindrà un escenari diferent sobre el que desenvolupar el seu treball. Més endavant hi ha la descripció de tots els escenaris. El primer que heu de fer, és escollir parella i se us assignarà un escenari. Un cop assignat l'escenari heu de fer, pel lliurament 1:

- Fer els càlculs de pressió per capacitat, IOPS i per xarxa

Pel lliurament 2, haureu de:

- Escollir una arquitectura per l'emmagatzematge (vegeu apèndix 1: escollir emmagatzematge) i si voleu NAS o SAN (apèndix 5).
- Escollir un *colocation center* en funció de les vostres característiques (vegeu apèndix 2: *housing*)
- Escollir un sistema de *backup* coherent (vegeu Apèndix 3: *backup*)
- Escollir quin *bandwidth* de sortida necessiteu al vostre CPD en funció de les necessitats de l'escenari i de si teniu un *mirror* extern (veure Apèndix 5)
- Comprovar que el cost final a 5 anys CAPEX + OPEX ha d'ajustar-se al vostre pressupost.

Lliuraments:

Hi ha dos lliuraments.

Als dos lliurament s'ha de presentar un informe seguint la guia del fitxer de word "DocumentLliurament.docx" disponible amb aquest enunciat.

Primer lliurament:

Pel primer lliurament només heu d'omplir tots els subapartats de l'apartat 2 del document (la resta no). És important raonar com s'ha calculat, més que donar el número final.

Quan: **Dijous 6 d'octubre**

On: Al racó, "Espai pel primer lliurament de l'activitat EBH"

Com: Un únic fitxer PDF amb tots els subapartats de l'apartat 2 del document.

Important: el nom ha de ser :

- NumeroEscenari.NomiEscenari.CognomMembre1.CognomMembre2.pdf
- Per exemple: 03.NETPRO.Lopez.Berral.PDF
- Facilita la vida del professor si el número d'escenari el poseu amb dues xifres (a l'exemple, 03)

Abans de passar a la segona part de la pràctica, discutirem els vostres números. **Ja us explicaré com ho farem.**

Segon lliurament:

El lliurament final ha de tenir l'informe complert i el full de càlcul.

Heu de lliurar un document PDF amb un informe dels vostres resultats

Nom del fitxer a lliurar:

- NumeroEscenari.NomEscenari.CognomMembre1.CognomMembre2.pdf

El document anterior ha de ser autocontingut. Lliureu també el fitxer amb fulls de càlcul, però el professor només els consultarà en cas de tenir algun dubte; en principi es puntuarà sobre l'informe original.

Nom del fitxer:

- NumeroEscenari.NomEscenari.CognomMembre1.CognomMembre2.xlsx

Els dos fitxers finals han de lliurar-se al racó, a “Lliurament Final Activitat EBH” abans del **Dimarts 25 d'octubre a les 23:55**. Aquests documents seràn utilitzats per fer les auditories (en parlarem a un altre document de les auditories).

Consultes:

Si voleu fer consultes sobre el treball us prego que seguiu el següent protocol: envieu-me un correu a david.lopez@upc.edu amb el subject [CPD-Escenari] i el que vulgueu afegir al subject. (per exemple: [CPD-ANA2] Dubtes sobre la pressió sobre disc) Ho necessito per classificar automàticament els correus. En cursos anteriors se'm va escapar algun.

Documentació: l'increïble full de càlcul

Us adjuntem un full de càlcul per simplificar molt (però que molt) la vostra feina. Encara que us treu molta feina, heu de dedicar una estona a entendre com funciona. Aquest fitxer té 7 fulls de càlcul molt interconnectats:

- Resum: Un full on heu d'introduir les dades originals del vostre escenari (Taula 1) i dades necessàries per fer els càlculs de costos generals (Taula 2). En aquestes dues taules s'han d'omplir les cel·les grises. Les taules 3 i 4 us donaran els costos OPEX i CAPEX, i la Taula 5 l'ajust al pressupost. La diferència hauria de ser positiva, però no molt. Una diferència positiva molt gran o negativa s'ha de justificar en la documentació final.
- Cabina de discos: només conté informació per calcular els costos de les cabines de discos (la informació s'ha d'introduir des de el full resum). La informació és la que hi ha a l'apèndix 1. No cal que modifiqueu cap dada d'aquest full.
- SAN: Tenir una SAN allibera la xarxa LAN de tràfic. És una opció que us pot interessar, però és cara. Al full “SAN” trobareu un sistema simplificat per construir una SAN. Aquest és un dels fulls que menys heu de consultar, doncs es parla de xarxes a l'altra part del curs. Més informació a l'apèndix 5.

- **Electricitat:** Un full que calcula el cost anual de l'electricitat. Depèn de dos paràmetres: el sobreprovisionament d'electricitat (que ens ve donat a les condicions del nostre escenari) i la despesa calculada pel nostre hardware. Aquest depèn del hardware que ja ens ve donat per l'escenari i el que costa el hardware que afegim per l'emmagatzematge. Les dades són als fulls "Resum", "Cabina de discos" i "Backup". Aquest full no s'hauria de modificar, només és informatiu.
- **Infraestructura:** En aquest full podeu veure la comparació de les tres empreses de *housing* pel que fa a costos. Els paràmetres que influeixen estan a la taula "informació provinent d'altres fulls", i és allà on l'hauríeu de modificar les dades. Aquest full no s'hauria de modificar, només és informatiu. Fixeu-vos que no només calcula el que costa el lloguer de l'espai sinó que també afegeix el cost elèctric afegit per UPS, HVAC, etcètera.
- **Backup:** Inclou una taula que s'omple a partir de dades d'altres fulls. Aquest full no s'hauria de modificar, només és informatiu. Ens permetrà comparar, donades unes característiques, quina és l'opció més econòmica entre les que compleixen les nostres especificacions. Al final haureu d'indicar l'opció escollir a la pàgina "Resum".
- **Bandwidth Provider:** un altre full de consulta. És a la pàgina "Resum" on heu d'indicar el tipus de línia que voleu contractar i el numero de línies (vegeu apèndix 4) Aquesta pàgina permet fer el càlcul del cost i deixa el resultat a la pàgina resum.

Heu de dedicar un temps a entendre els fulls de càlcul (les implicacions de cada decisió). Aquest fulls ajuden a organitzar la feina i fer els càlculs, però no substitueixen les decisions tècniques que heu de prendre en funció del vostre escenari.

Escenaris

Plantejarem 6 entorns de treball. Hi ha més d'un escenari per cada entorn de treball, amb característiques diferents, on les diferències principals serà la grandària del problema i el pressupost disponible. A més a més, el SLA (Service-Level Agreement o contracte de nivell de servei que té cada empresa amb els seus clients) pot ser diferent.

Elements comuns:

- Tots tenen un sistema de cabines de disc (no hi ha un sistema distribuït), sigui en SAN o NAS.
- Tots necessiten un *Colocation Center* on deixar les màquines donat que no podem crear un edifici pel nostre CPD. Tenim 3 empreses on ens podem hostatjar el nostre CPD (veure Apèndix 2)
- Tots necessiten un sistema de *backup*, d'acord amb les seves necessitats (veure Apèndix 3)

Detalls generals

Un servei de Web 2.0 especialitzada en emmagatzematge de fitxers d'imatge, tipus Flickr o Picassa.

- Tot el CPD està orientat a una única aplicació, i les màquines són homogènies.
- S'espera una mitja de 1000 fotografies per usuari i 6 MB per fotografia (comprimida).
- El procés de servir una fotografia és el següent: el servidor rep per la xarxa una petició (HTTP, 1KB). Sol·licita la fotografia al disc, que li passa (completa) al servidor. El servidor fa un petit tractament i envia per xarxa la fotografia amb 180KB d'informació addicional.
- El procés de rebre una foto és el següent: el servidor rep una petició de 1 KB i envia una acceptació (1KB). A continuació el servidor rep per xarxa la foto (6MB), fa un tractament i la envia a disc, enviant a continuació una acceptació a l'usuari de 1KB.
- El CPD només es fa servir per l'emmagatzemament de fotos. Tot el que és la gestió d'usuaris es fa en un servidor extern del que no ens ocuparem.
- L'empresa d'aquest escenari pot oferir dos serveis diferents: un emmagatzemament de fotografies (més semblant a un servei tipus dropbox, per backup però no visualització immediata) i un sistema de visualització de fotografies. El primer és molt més barat pels clients i a nivell de funcionament son semblats. La diferència fonamental està en la disponibilitat de les dades: el servei de visualització espera una resposta ràpida, mentre que el servei tipus dropbox pot ser més lent ja que garantim tenir les fotos emmagatzemades, però no d'accés immediat - pot requerir hores recuperar-la.

Escenari 01: NET 1

Dades bàsiques

- Aquest servei espera que el 30% dels clients hagi escollit el sistema tipus dropbox i un 70% l'interactiu.
- Usuaris esperats: 600.000
- Peticions per minut: 10.000
- Percentatge de fotografies llegides / escrites: 95% / 5%
- Xarxa LAN de 10 Gbps
- Diners totals: 25.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 15.500.000 euros
- Escollits Racks de 42U
- Servers: 470 servidors de càlcul (2U) + 24 Switch (1U) + 2 Switch (4U). Total: 972Us
- Consum: 1979,8 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 12%
- SLA: Downtime màxim de 2 hores/any pel que fa al sistema interactiu. Per cada hora de downtime hem d'abonar €200.000 als clients en serveis

Escenari 02: NET2

Entorn de treball: Net

Paràmetres de l'escenari.

- Aquest servei espera que el 50% dels clients hagi escollit el sistema tipus dropbox i un 50% l'interactiu.
- Usuaris esperats: 60.000
- Peticions per minut: 3.000
- Percentatge de fotografies llegides / escrites: 95% / 5%
- Xarxa LAN de 10 Gbps
- Diners totals: 10.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 7.500.000 euros
- Escollits Racks de 42U
- Servers: 60 servidors de càlcul (2U) + 6 Switch (1U) + 2 Switch (4U). Total: 134Us
- Consum: 180,5 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 7%
- SLA: Downtime màxim de 24 hores/ any per l'interactiu. Si el sistema cau un temps oferim als clients uns diners per gastar en serveis. Per cada hora de downtime hem d'abonar €20.000 als clients.

Escenari 03: NETPRO

Entorn de treball: Net

Paràmetres de l'escenari.

- Aquest escenari és per guardar fotografies per professionals (mitjans de comunicació, per exemple) que necessiten guardar un històric de fotos molt gran, però que les que fan servir de manera interactiva són relativament poques (les properes edicions dels diaris, per exemple). En concret el 95% de les fotos es guarden al dropbox i només el 5% a l'interactiu.
- En canvi, aquest servei espera que el 10% de les transaccions vagin al dropbox i un 90% a l'interactiu.
- Usuaris esperats: 1.000
- Peticions per minut: 10.000
- Fotografies que s'espera guardar en aquest cas son 10.000.000 fotos per usuari (i no les 1000 de l'enunciat general)
- Percentatge de fotografies llegides / escrites (clients dropbox): 99% / 1%
- Percentatge de fotografies llegides / escrites (clients interactius): 60% / 40%
- Xarxa LAN de 10 Gbps
- Diners totals: 75.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 10.000.000 euros
- Escollits Racks de 42U
- Servers: 60 servidors de càlcul (2U) + 6 Switch (1U) + 2 Switch (4U). Total: 134Us
- Consum: 180,5 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 7%
- SLA: Downtime màxim de 3 hores/ any per l'interactiu. Si el sistema cau un temps oferim als clients uns diners per gastar en serveis. Per cada hora de downtime hem d'abonar €2.000.000 als clients.

Detalls generals

CBR (Centre Bàsic de Recerca) és un CPD producte de la unió de molts grups, cadascú amb les seves màquines (similar al que hi ha a l'Omega).

- Màquines heterogènies amb diferents necessitats.
- Per simplicitat, assumirem que les màquines de cada grup estan dins d'un mateix rack, de manera que si han d'intercanviar informació ho faran localment. L'intercanvi d'informació entre racks és molt inhabitual.
- N'hi ha grups que tenen molt tràfic amb l'exterior i d'altres molt poc. N'hi ha que fan molts accessos a disc i els que fan molt pocs. A més les condicions poden canviar en pocs mesos. El disseny ha de ser adaptable a aquests canvis.
- En mitja, cada rack intercanvia 4MBps de dades amb l'exterior del CPD i 8 MBps amb el sistema d'emmagatzematge. Les lectures i escriptures a disc estan sobre el 50%.
- El 70% de les dades emmagatzemades a disc no es consulten habitualment (i els servidors saben quines són). Els clients accepten que aquestes dades puguin trigar més temps en ser accedides.
- SLA (CBR 1 i CBR2): Downtime màxim de 24 hores/ any. No hi ha penalització econòmica doncs no tenim clients, però cada grup de recerca pot perdre €20.000 per hora de downtime en els seus contractes.
- SLA (CBRMIX): Pel que fa als grups de recerca, les condicions són iguals que els de CBR1 i 2, però la part de la universitat (vegeu l'enunciat) requereix un downtime màxim de 3 hores a l'any.

Escenari 04: CRB1

Entorn de treball: CRB.

- Actualment tenim 19 grups de recerca, que ocupen un total de 30 racks, amb 270 servidors que ocupen 1000 Us. Els racks són de 42 U.
- Actualment cada grup de recerca té un espai de 40 TB assignats
- La xarxa LAN és de 4 Gbps.
- Diners totals: 20.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 16.500.000 euros
- Consum: 1233 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 12%

Escenari 05: CRB2

Entorn de treball: CRB.

- Actualment tenim 10 grups de recerca, que ocupen un total de 18 racks, amb 240 servidors que ocupen 690 Us. Els racks són de 42U
- Actualment cada grup de recerca té un espai de 30 TB assignats
- La xarxa LAN és de 2 Gbps.
- Diners totals: 12.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 9.500.000 euros
- Consum: 780 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 7%

Escenari 06: CRBMIX

Entorn de treball: CRB.

- Aquest escenari és igual que el CRB2, però a més de tenir grups de recerca també tenim al mateix lloc els ordinadors de la Universitat, que comparteixen xarxa, disc, etc...

Part de la Universitat

- Els ordinadors ocupen 11 racks amb 120 servidors que ocupen 450 Us. Els racks són de 42U
- Aquesta part té assignats 400TB de dades. El tràfic i funcionament és similar al de les altres màquines

Part dels grups de recerca

- Actualment tenim 10 grups de recerca, que ocupen un total de 18 racks, amb 240 servidors que ocupen 690 Us. Els racks són de 42U
- Actualment cada grup de recerca té un espai de 30 TB assignats

Part comú

- La xarxa LAN és de 2 Gbps.
- Diners totals: 14.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 11.500.000 euros
- Consum: 1300 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 8%

Entorn ANA

Detalls generals

ANA: Empreses d'anàlisi de dades. Recullen dades que els hi proporcionen els clients i en fan anàlisis estadístics de tendències.

- Màquines homogènies, executen una aplicació.
- Estan rebent dades externes contínuament i les han de processar en temps real. Emmagatzemen moltes dades intermitges i els servidors estan intercanviant dades contínuament.
- La majoria de les dades tenen una vida molt curta als nostres servidors. Només recollim dades, les analitzem i enviem els resultats. Es guarden dades històriques, però només 10 dies. Aquestes dades històriques no s'accedeixen massa sovint però han d'estar disponibles durant 10 dies (el tràfic ja està contemplat a cada escenari). Podem considerar que cada dia un 10% de les dades acumulades ja es poden esborrar i la cosa creix en un 10% addicional.
- El resultat de les anàlisis han de donar-se en un temps molt curt. Tenim un SLA que ens fa pagar una forta suma al client si no som capaços de processar les dades a una determinada velocitat (inclús si s'espatlla algun disc). Hem de garantir el mínim de downtime possible.

Escenari 07: ANA1

Entorn de treball: ANA.

- Cada servidor té un tràfic en mitja de 1 Mbps de dades. El 75% entre servidors (LAN) i el 25% restant al disc (del qual el 70% són lectures).
- Cada servidor guarda un màxim 1TB de dades al disc centralitzat.
- De tant en quant n'hi ha pics de treball, 400 Mbps de tràfic (la distribució és la mateixa que l'anterior)
- Cada operació al disc és de 4KB
- Les dades històriques ocupen 40TB
- Tenim una LAN Ethernet de 10 Gbps
- Diners totals: 21.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 15.500.000 euros
- Escollits Racks de 42U
- Servers: 200 servidors de càlcul (4U) + 44 Switch (1U) + 2 Switch (4U). Total: 846Us
- Consum: 1812,8 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 15%
- SLA: Paguem €1.000.000 per cada hora o fracció de *downtime*.

Escenari 08: ANA2

Entorn de treball: ANA.

- Cada servidor té un tràfic en mitja de 1.5 Mbps de dades. El 75% entre servidors (LAN) i el 25% restant al disc (del qual el 70% són lectures).
- Cada servidor guarda un màxim 1TB de dades al disc centralitzat.
- De tant en quant n'hi ha pics de treball, 100 Mbps de tràfic (la distribució és la mateixa que l'anterior)
- Cada operació al disc és de 4KB
- Les dades històriques ocupen 10TB
- Tenim una LAN Ethernet de 2 Gbps
- Diners totals: 10.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 7.500.000 euros
- Escollits Racks de 42U
- Servers: 22 servidors de càlcul (4U) + 6 Switch (1U) + 2 Switch (4U). Total: 96Us
- Consum: 202,8 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 7%
- SLA: paguem als clients €150.000 per hora o fracció de downtime

Escenari 09: ANA+

Entorn de treball: ANA.

- Cada servidor té un tràfic en mitja de 2 Mbps de dades. El 50% entre servidors (LAN) i el 50% restant al disc (del qual el 90% són lectures).
- Cada servidor guarda un màxim 2TB de dades al disc centralitzat.
- De tant en quant n'hi ha pics de treball, 200 Mbps de tràfic (la distribució és la mateixa que l'anterior)
- Cada operació al disc és de 4KB
- Les dades històriques ocupen 50 TB
- Tenim una LAN Ethernet de 10 Gbps
- Diners totals: 15.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 11.000.000 euros
- Escollits Racks de 42U
- Servers: 50 servidors de càlcul (4U) + 6 Switch (1U) + 2 Switch (4U). Total: 96Us
- Consum: 500 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 6%
- SLA: paguem als clients €100.000 per hora o fracció de downtime

Entorn VM

Detalls generals

CPDs que ofereixen màquines virtuals (VM) als clients. De fet, oferim un IaaS (*Infrastructure as a Service*)

- Tot i que cada client pot tenir les màquines com vulgui, el CPD és homogeni i les seves operacions també ja que parlem de màquines virtualitzades. Cada VM ocupa 4GB de disc (al propi servidor). Cada MV té 1TB de dades assignat al disc centralitzat.
- Cada VM fa 6000 Kbps de lectura i 4000 Kbps d'escriptura cap el disc.
- Cada VM realitza 4000 Kbps de comunicació amb l'exterior.
- Les VM poden migrar de servidor físic en qualsevol moment, en funció de la càrrega total del sistema, política de consolidació, la càrrega de cada server, la política de balanceig de càrrega, ... heu de tenir-lo en compte a l'hora d'aprovisionar xarxa interna.
- SLA (VM1 i VM2): els clients tenen els seus negocis en les nostres màquines virtuals (molts clients tenen més d'un VM). El nostre compromís (als dos primers escenaris) és pagar al client €250 per hora o fracció de *downtime*.
- SLA (VM+). En aquest cas hi ha dos tipus de clients, els que contracten el servei bàsic (amb un SLA com l'anterior) i els clients Premium als que hem de garantir un downtime màxim de 20 hores/any i els hi paguem €5.000 per hora o fracció de downtime.

Escenari 10: VM1

Entorn de treball: VM.

- Número de clients: 2500
- Número de VM: 2000
- Cada minut migren en mitja 5 VM.
- La LAN es Ethernet de 10 Gbps
- Diners totals: 35.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 21.500.000 euros
- Escollits Racks de 42U
- Servers: 540 servidors de càlcul (2U) + 28 Switch (1U) + 4 Switch (4U). Total: 1124 Us
- Consum: 2300,8 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 10%

Escenari 11: VM2

Entorn de treball: VM.

- Número de clients: 200
- Número de VM: 500
- Cada minut migren en mitja 5 VM.
- La LAN es Ethernet de 8 Gbps
- Diners totals: 10.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 7.500.000 euros
- Escollits Racks de 42U
- Servers: 50 servidors de càlcul (2U) + 4 Switch (1U) + 2 Switch (4U). Total: 112 Us
- Consum: 350 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 6%

Escenari 12: VM+

Entorn de treball: VM.

En aquest entorn distingim dos tipus de clients (vegeu l'escenari general VM). Hem escollit dos tipus de servers, un per el sistema estàndard i un altre pel premium

- Número de clients: 300= 200 (estàndard) i 100 (premium)
- Número de VM: 500= 250 estàndard i 250 premium
- Cada minut migren en mitja 5 VM.
- La LAN es Ethernet de 10 Gbps
- Diners totals: 15.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 8.500.000 euros
- Escollits Racks de 42U
- Servers (estàndard): 50 servidors de càlcul (2U) + 4 Switch (1U) + 2 Switch (4U). Total: 112 Us
- Servers (premium): 70 servidors de càlcul (2U) + 4 Switch (1U) + 2 Switch (4U). Total: 152 Us
- Consum: 1200 kW
- Sobreprovisionament d'electricitat (veure apèndix 4): 7%

Detalls generals

Electronic Commerce (EC): Parlem del CPD d'una empresa que ofereix venda *on-line* de productes (tipus Amazon).

- El sistema es basa en tenir una Base de Dades enorme amb descripció dels productes i temes d'stock i compra.
- Per cada connexió els usuaris solen fer diverses consultes. Cada sessió oberta dura, en mitja, 20 minuts. Hi ha usuaris que compren diversos productes i qui no compra res, però a efectes pràctics una connexió acaba (o no) amb una compra, no importa de quants productes. En mitja es venen 0,2 productes per sessió oberta.
- Tenim 3 tipus de dades:
 - Dades de client: nom, adreça d'enviament, historial de compres, dades de cobrament i facturació (i alguna més). De cada client tenim 256 KB d'informació.
 - Base de dades de productes: en primer nivell només una fotografia de baixa resolució, descripció bàsica i elements en stock. Aquesta informació és la que apareix quan fem una cerca i ens ensenya diversos productes. Si després volem la informació addicional hem de fer un segon accés a les dades del producte. Per cada producte que emmagatzemem tenim 8 KB de dades (inclou l'enllaç per trobar més dades del producte).
 - Dades dels productes: descripció, fotografies amb més resolució, opinions dels compradors, etcètera. De cada producte guardem 64 KB de dades.
- El funcionament és el següent: quan un usuari es connecta (petició HTTP 1KB) és assignat a un servidor i es quedarà en aquest servidor tota la sessió. El primer que demana el servidor són les dades del client per guardar-les localment. A continuació el client fa una cerca i es consulta la BD de productes que li ofereix les dades bàsiques de tots els productes que lliguen amb la cerca. En mitja cada usuari fa 10 consultes per sessió, amb una mitja de 400 resultats per consulta. A continuació es miren dades completes d'alguns dels productes exposats. En mitja en cada sessió un client mira 30 productes. Per cada consulta o producte s'envia al client la informació rebuda de disc, més 80 KB d'informació addicional (inclou la resposta del client). Finalment, alguns clients compren, provocant 16 KB addicionals de tràfic a la xarxa i l'escriptura dels seus 256 KB de dades al disc per cada compra.

Escenari 13: EC1

Entorn de treball: EC

- Número de sessions al dia: 600.000
- Número de productes: 20.000.000
- Número de clients: 250.000
- La LAN es Ethernet de 8 Gbps
- Diners totals: 20.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 16.500.000 euros
- Escollits Racks de 42U
- Servers: 540 servidors de càlcul (2U) + 28 Switch (1U) + 4 Switch (4U). Total: 1124 Us
- Consum: 2000 kW. Sobreprovisionament d'electricitat (veure apèndix 4): 10%
- SLA: les pèrdues estimades per hora de *downtime* és de €150.000

Escenari 14: EC2

Entorn de treball: EC

- Número de sessions al dia: 300.000
- Número de productes: 20.000.000
- Número de clients: 250.000
- La LAN es Ethernet de 2 Gbps
- Diners totals: 5.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 3.500.000 euros
- Escollits Racks de 42U
- Servers: 50 servidors de càlcul (2U) + 4 Switch (1U) + 2 Switch (4U). Total: 112 Us
- Consum: 360 kW. Sobreprovisionament d'electricitat (veure apèndix 4): 7%
- SLA: les pèrdues estimades per hora de *downtime* és de €40.000

Escenari 15: ECPRO

Entorn de treball: EC, però especialitzat: pocs productes, pocs clients però moltes operacions

- Número de sessions al dia: 2.000.000
- Número de productes: 10.000.000
- Número de clients: 20.000
- La LAN es Ethernet de 10 Gbps
- Diners totals: 10.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 7.500.000 euros
- Escollits Racks de 42U
- Servers: 50 servidors de càlcul (2U) + 4 Switch (1U) + 2 Switch (4U). Total: 112 Us
- Consum: 360 kW. Sobreprovisionament d'electricitat (veure apèndix 4): 7%
- SLA: les pèrdues estimades per hora de *downtime* és de €40.000

Detalls generals

VineTok (Inc) és una empresa dedicada a oferir vídeos curts gratuïts (a qualitat 720p 30fps, i de fins a 20 segons). Es guanya la vida a partir d'incloure publicitat abans de cada vídeo (i no, no t'ho pots saltar). Un detall important és que els vídeos que pugen els clients no es mantindran per sempre: es guardaran uns quants (els més visitats) i els més nous, la resta es va esborrant. Els més visitats d'entre els nous passen a la llista de més visitats en uns dies, esborrant algun dels més visitats. Si passat uns dies els nous no han assolit la fita de més visitats, s'esborren també. La quantitat exacta de més visitats i nous depèn de cada escenari.

- Per poder treballar amb VineTok has d'estar donat d'alta i instal·lar-te una app, per accedir per mitjà d'aquesta (això és per simplificar l'escenari)
- El primer que fa l'usuari és fer un login a l'app (al seu dispositiu), i aquesta envia al nostre CPD una sol·licitud de connexió amb identificació de l'usuari (això ocupa 4KB)
- Un cop feta la sol·licitud, es crea una màquina virtual (MV) per atendre aquesta petició. La MV s'assigna a un servidor i en cap cas migrarà de servidor. El codi de la MV està emmagatzemat al disc local de cada servidor, i per tant posar en marxa la MV no tindrà tràfic associat.
- La MV assignada a la petició farà una mitjana de 10 peticions a disc sobre l'historial del client(5 peticions) i el que tenim en oferta per fer-li recomanacions (5 peticions més). Per cadascuna de les peticions a disc la MV rebrà 128KB de dades.
- Per cada client emmagatzemen 4GB de dades (no totes es consulten cada vegada)
- La MV calcularà, en funció d'aquesta informació, les recomanacions i les enviarà a l'app del client (512KB, incloent publicitat)
- El client pot demanar de pujar un vídeo propi a VineTok, o descarregar-se diversos vídeos per visualitzar. Assumirem que si demana pujar un vídeo no farà cap altre operació, mentre que si vol veure vídeos, en mitjana veurà 5 vídeos per sessió (sense abandonar la sessió i per tant utilitzant la mateixa MV)
- Si decideix pujar un vídeo, l'app l'enviarà per la xarxa. Assumirem que la mida mitjana del vídeo és de 20MB. La MV virtual farà un procés i ho comprimirà i emmagatzemarà al sistema de disc (ocupant 5MB)
- La visualització implica, per cada vídeo, una petició de l'app al servidor de 4KB, el servidor sol·licitarà a disc el vídeo (el vídeo que passa de disc al servidor ocupa 5MB), el servidor el descomprimirà i afegirà la publicitat que toqui, enviant finalment el vídeo modificat del servidor a l'app per streaming (20MB).
- Si l'usuari es desconnecta, la MV mor.
- Tot el tema de publicitat i cobrar pels anuncis no ho considerem com a part de l'enunciat.
- La situació de partida i creixement en vídeos depèn de cada escenari.

Escenari 16: VineTok1

Entorn de treball: VineTok

- Quants vídeos es guarden per molt visitats: 1.000.000
- Número de clients: 30.000
- Número de sessions al dia: 1.000.000
- Sessions que són de pujar vídeos: 10%
- Dies mínim que es guarden els vídeos nous (abans d'esborrar o ser considerats molt visitats): 10
- La LAN es Ethernet de 10 Gbps
- Diners totals: 30.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 24.000.000 euros
- Escollits Racks de 42U
- Servers: 22 servidors de càlcul (2U) + 6 Switch (1U) + 2 Switch (4U). Total: 96 Us
- Consum: 202,9 kW.
- Sobreprovisionament d'electricitat (veure apèndix 4): 10%
- SLA: no hi ha (servei gratuït)

Escenari 17: VineTok2

Entorn de treball: VineTok

- Quants vídeos es guarden per molt visitats: 2.500.000
- Número de clients: 500.000
- Número de sessions al dia: 1.000.000
- Sessions que són de pujar vídeos: 15%
- Dies mínim que es guarden els vídeos nous (abans d'esborrar o ser considerats molt visitats): 5
- La LAN es Ethernet de 10 Gbps
- Diners totals: 30.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 21.000.000 euros
- Escollits Racks de 42U
- Servers: 60 servidors de càlcul (2U) + 6 Switch (1U) + 2 Switch (4U). Total: 134 Us
- Consum: 180 kW.
- Sobreprovisionament d'electricitat (veure apèndix 4): 7%
- SLA: no hi ha (servei gratuït)

Escenari 18: VineTokBigData

Entorn de treball: VineTok

Aquest escenari és una mica diferent: hi ha un únic client (la nostra empresa) que ja té un CPD creat (com els dels escenaris anterior). El CPD d'aquest escenari és per fer anàlisi de vídeos per descobrir tendències. Per tant les condicions són diferents: per un banda, els servidors reben vídeos i els emmagatzemen directament. Per altre, analitzen vídeos emmagatzemats i generen un informe diari de tendències. Cada dia esborren tants vídeos de l'històric com vídeos nous arriben (esborren els més vells, per cercar tendències actuals).

- Quants vídeos es guarden: 100.000.000
- Vídeos que arriben cada día: 50.000
- Els vídeos ja arriben compactats (5MB per vídeo quan arriba, que es grava directament)
- Els servidors analitzen diàriament 1.000.000 de vídeos dels emmagatzemats al sistema. Per cada vídeo analitzat, el servidor que l'ha analitzat envia 80MB de dades a un altre servidor (centralitzat, que va generant l'informe diari)
- Informe diari generat (enviat a l'empresa): ocupa 20GB
- La LAN es Ethernet de 10 Gbps
- Diners totals: 25.000.000 euros
- Gastat en Racks, servers, cables, switchos, etc...: 14.000.000 euros
- Escollits Racks de 42U
- Servers: 22 nodes (2U) + 6 Switch (1U) + 2 Switch (4U). Total: 96 Us
- Consum: 200 kW. Sobreprovisionament d'electricitat (veure apèndix 4): 6%
- SLA: no hi ha (servei propi)

Apèndix 1: Disseny d'un sistema d'emmagatzematge

Idees bàsiques

El disseny depèn d'una sèrie de paràmetres que heu de tenir en compte. En primer lloc, cal saber la mida de les dades i la pressió de lectura/escriptura sobre els discos, així com el percentatge habitual d'operacions que són escriptures.

- Bytes a emmagatzemar (*raw*)
- Pressió sobre els discos (IOPS)
- % d'escriptures

Aquestes dades s'han de deduir de cada escenari. En funció d'aquests paràmetres podem calcular el nombre de discos MÍNIMS requerits per IOPS i per capacitat. Això no vol dir que han de ser els discos que comprem definitivament, doncs no és qüestió d'aprovisionar pel màxim actual, sinó pensant en el futur. Pareu atenció, perquè la sobreprovisió és especialment important per l'informe final i les recomanacions als inversors.

El següent pas és pensar quin tipus de raid vull. Puc dissenyar des d'un JBOD o un RAID 0, fins a un RAID 61. Per fer una primer aproximació us oferim el full de càlcul Discosrequerits.xlsx. Si obriu el full trobareu a la cantonada superior esquerra una taula que heu d'omplir amb les IOPS requerides, el percentatge d'operacions d'escriptura (entre 0 i 1) i la mida de les dades. Aquestes són les dades que heu d'omplir. La taula superior dret indica les operacions requerides i discos mínims per cada tipus de RAID que contemplarem (RAID 0, 10 -que a efectes pràctics és igual que el 01-, 5, 51, 6 i 61). La resta de taules del full de càlcul indiquen per cadascun dels models de disc que contemplem, quants discos d'aquell tipus farien falta per donar suport a la càrrega màxima requerida.

Models de disc

A l'hora de dissenyar una cabina de discos podeu escollir entre molts models al mercat, i l'elecció pot variar en unes setmanes si surten nous models o varien els preus. Per simplificar aquesta activitat us oferim 10 models de disc perquè a la pràctica escolliu entre ells, sense tenir que buscar al mercat nous models.

Oferim una cosa variada: 5 discos magnètics (HDD) i 5 d'estat sòlid (SSD) alguns de tipus *Consumer* i altres *Enterprise*. Les opcions cobreixen un ampli espectre de capacitat, velocitat i preu.

Les 5 primeres opcions són discos HDD i les 5 següents SSD. Les característiques dels discos les podeu trobar a les taules que hi ha a continuació (la primera per HDD i la segona per SSD).

Opcions HDD:

- Opció 1: Seagate Barracuda ST8000DM0004
- Opció 2: Toshiba MG07ACA14TA
- Opció 3: Seagate ST10000NM009G
- Opció 4: HPE 765466-B21
- Opció 5: HPE EG002400JWJNN

Opcions SSD:

- Opció 6: Samsung 860 EVO
- Opció 7: Intel Optane H10
- Opció 8: Kingston SEDC100M
- Opció 9: WD Gold S768T1D0D
- Opció 10: WD Ultrastar DC SN640

Opció	1	2	3	4	5
Model	Seagate Barracuda ST8000DM0004	Toshiba MG07ACA14TA	Seagate ST10000NM009G	HPE 765466-B21	HPE EG002400JWJNN
Tipus	Consumer	Enterprise	Enterprise	Enterprise	Enterprise
Capacitat (TB)	8	14	10	2	2,4
Consum (W)	6.8	7.8	9.5	7	7.1
Preu (€)	235	520	350	250	360
IOPS R/W	640	800	710	3360	5210
RPM	5400	7200	7200	10000	10000
€/GB	0,029375	0,037142857	0,035	0,125	0,15

Especificacions de les opcions 1-5: discos HDD

Opció	6	7	8	9	10
Model	Samsung 860 EVO	Intel Optane H10	Kingston SEDC100M	WD Gold S768T1D0 D	WD Ultrastar DC SN640
Tipus	Consumer	Consumer	Enterprise	Enterprise	Enterprise
Capacitat (TB)	2	1	1,92	7,68	3,8
Consum (W)	2.2	5,8	9	12	8
Preu (€)	310	195	372	1545	750
IOPS R/W	90k / 10k	330K /250k	540K /205K	467k/ 65K	511K / 82K
Tecnologia	3D QLC NAND	3D QLC NAND	3D TLC NAND	3D TLC NAND	3D TLC NAND
€/ GB	0,155	0,195	0,19375	0,201171875	0,197368421

Especificacions dels discos 6-10

Al full de càlcul teniu, per cada una de les 10 opcions, les dades bàsiques entrades (taula esquerra) i els càlculs següents:

- Nombre de discos requerits degut a la pressió de les operacions de disc (IOPS)
- Nombre de discos requerits degut a la mida de les dades (capacitat)
- Nombre de discos mínim requerits (el màxim dels dos valors anteriors)
- Capacitat de disc de la solució (nativa, sense tenir en compte replicacions degut a RAID)
- Consum del sistema de discos
- Preu en euros DELS DISCOS de la solució.

Això us ha de servir per orientar-vos a prendre una decisió de quants discos compreu, de quin model i com els organitzeu (quin RAID feu servir).

A més a més, heu de instal·lar aquests discos en una cabina. Novament per simplificar hem escollit unes poques cabines a escollir.

Cabines disponibles

Hem escollit 6 cabines de disc (a partir de <http://www.thinkmate.com/system/stx-en-xe24-2460>).

En principi totes elles tenen fonts d'alimentació redundants i totes tenen els processadors, memòria ECC, chipset, Boot drive (un SSD) i dues controladores de xarxa PCIe SAS 12Gbps, excepte la solució 6 que és molt més senzilla ja que només és un JBOD.

Per simplificar, les badies indicades poden tenir discos de Form Factor 2,5" i 3,5".

- El model 1 està pensat per tenir pocs discos (màxim de 12), tant HDD com SSD. Admet JBOD, RAID 0, 10, 5, 51, però no 6 ni 61. Ocupa dues Us
- Els models 2 i 3 son bàsicament el mateix, amb 24 badies i 4 Us de mida, admeten tots tipus de RAID (però no JBOD, és massa car per ser un JBOD). La única diferència entre els dos és que el model 3 inclou dos SSD Enterprise (que no ocupen badies i que pugen el preu 500€) per fer de cache d'operacions d'E/S al sistema de disc. Aquesta solució és recomanable en cabines de discos HDD on hi ha ràfegues de moltes lectures/escriptures en breus espais de temps. Es poden posar en cache 100.000 IOPS, tenint en compte que els discos són de 500GB cadascú.
- Els models 4 i 5 són similars als 2 i 3, però amb més badies (fins a 36). Novament la diferència entre tots dos és que el model 5 té dos SSD de cache.
- El model 6 està pensat només per ser un JBOD i té 24 badies.

A continuació teniu una taula-resum:

Model	1	2	3	4	5	6
Badies	12	24	24	36	36	24
SSD suport	0	0	2	0	2	0
JBOD	SI	NO	NO	NO	NO	SI
RAID 0	SI	SI	SI	SI	SI	NO
RAID 10	SI	SI	SI	SI	SI	NO
RAID 5, 51	SI	SI	SI	SI	SI	NO
RAID 6, 61	NO	SI	SI	SI	SI	NO
# Us	2	4	4	4	4	4
Admet Spare disk?	NO	Si, fins a 4	Si, fins a 4	Si, fins a 4	Si, fins a 4	NO
Consum (Watts)	800	1200	1216	1300	1316	1000
Preu (en €)	3400	4600	5100	5000	5500	2400

Què heu de fer?

El que heu de pensar és si teniu més d'un tipus de dades per necessitat, seguretat, etcètera. És possible que al vostre escenari tingueu dades que convé tenir en ràpid accés i altre no tant, o molt replicades i d'altres no tant. El primer que heu de fer és tenir clar si teniu més d'un grup i, per cada grup, calcular la quantitat mínima d'espai i IOPS que necessito. El full de càlcul Discosrequerits.xlsx un ajudarà a decidir un model de disc en funció de les necessitats (o més d'un possible model), i la quantitat de discos que comprareu, amb la sobreprevisió necessària (dissenyeu per possibles ampliacions).

El següent és decidir el model de cabina de discos que millor s'adapta i la quantitat. Podeu tenir diverses cabines si teniu molts discos o si teniu models separats. Per exemple, podeu tenir un parell de cabines de JBOD amb HDD per un tipus de dades massiu i una cabina amb SSD en RAID 51 per altres dades... depèn del vostre escenari. A més és més fàcil organitzar el *backup* / replicació per cabines.

També haureu de calcular com feu els *clusters* de discos si organitzeu RAID. No té lògica tenir una cabina amb 36 badies, assumir un RAID 5 i pensar que tens 35 discos i 1 de paritat. El més habitual és fer *clusters* de 5-8 discos (més el *mirror*). Per exemple, suposem que volem una cabina 4 en RAID 51 amb dos discos *spare disk*. Tenim 36 badies, dues ocupades per *spare disk*, queden 34 discos per dades. Puc fer 3 *clusters* de 12, 12 i 10 discos. Els de 12 discos tenen la següent arquitectura: (5 discos de dades + 1 degut al RAID 5) x 2 degut a RAID 1, i el cluster de 10 discos (4 discos de dades + 1 degut RAID 5) x 2 per RAID 1. En total tenim 36 discos, però la capacitat real és de 14 per dades, 20 per redundància (RAID) i 2 *spare*.

A partir d'aquí, i segons les necessitats de l'escenari heu de calcular número de cabines de cada tipus (màxim 3 tipus diferents) i número de discos de cada tipus (màxim 3 tipus diferents). En l'exemple anterior seria 1 cabina tipus 4 i 36 discos tipus X (el que sigui).

Tota aquesta informació s'ha d'indicar al full de calcul principal, full resum, taula 2. Cal omplir la informació referent a la cabina de discos principal (sempre tindreu al menys una) i podeu posar fins a dos tipus més de cabines. De cadascuna heu d'escollir el tipus de disc (entre 1 i 10) i de cabina (entre 1 i 6) i el nombre de cabines principals i secundàries (podeu tenir més d'una cabina de cada tipus). En aquest cas no admetem més de tres tipus de cabines (no acostuma a ser necessari tenir-ne més). Si no feu servir algun tipus de cabina, n'hi ha prou amb que poseu a zero el nombre de discos i cabines a la taula 2.

ATENCIÓ: Això no comprova coherència, que és feina vostra. Per exemple, si poseu 40 discos en una cabina on hi ha 24 badies, el full de càlcul no es queixarà.

Apèndix 2: Housing

Ja que no podem (per temps ni per costos) construir un edifici, hem decidit utilitzar el sistema de *Colocation*. És a dir, llogarem un espai on posar el nostre CPD. Per fer-ho tenim 3 possibilitats. Fem notar que els preus (per simplicitat, l'iva està inclòs), encara que realistes, són inventats.

Opció 1: El *Data Center* en contenidors

Empresa (fictícia): Modular Containers S.A. (MOCOSA)

Ubicació: Cerdanyola del Vallès.

- El sistema funciona per mitjà del lloguer de contenidors dins un parc de contenidors. Els elements del contenidor tenen una vida útil de 5 anys.
- S'ha de llogar un nombre sencer de contenidors. El preu per contenidor és de 115.000 euros/any. Inclou neteja tècnica dels contenidors.
- Cada contenidor pot contenir fins a 19 racks de fins a 48 U (per servers, *switchos* i emmagatzematge). Té protecció d'exterior antihumitat, fum, aigua, incendis, impulsos electromagnètics i radiació solar. El que no té és cap Tier Certificat.
- L'electricitat consumida es paga a part, per consum i potència contractada (vegeu apèndix 4). En aquest cas hem de pagar el nostre UPS i la nostra refrigeració. Cada contenidor admet un màxim de 700kw (input), el que inclou la despesa de UPS i refrigeració. Ens asseguruen un PUE de 1,3, així que hem de calcular de consumir un 30% addicional per UPS, refrigeració, etcètera.
- Seguretat: hi ha un sistema antiincendis per reducció d'oxigen, càmeres interiors (CCTV) monitoritzables exteriorment. Sistema de seguretat d'entrada per lectors de bandes magnètiques, sensors biomètrics (empremta digital) i teclat alfanumèric. El parc té, inclòs en el lloguer, un sistema de seguretat de manera que només entren al parc les persones autoritzades, però la seguretat del parc no té accés a els sensors de dins el contenidor, ni al CCTV, ni d'entrada al contenidor a menys que ho contractis a part (Sistema de monitorització i vigilància 24/7: +2,000 euros/any per contenidor, inclou canvi de *spare disc* a la cabina de discos si n'hi ha)).
- Redundància: tots els sistemes dins el contenidor tenen redundància N+1. El parc té un sistema d'alimentació dièsel garanteix 48 hores de subministrament pic. El parc només té una línia contractada d'electricitat i una de dades, compartides per tots els contenidors.
- *Backup* i *mirror*: podem fer servir qualsevol combinació de les presentades a l'apèndix 3. Si volem incloure un sistema de *backup on-site* i volem que ho gestionin ens cobraran €1.500 al mes per canviar cintes, guardar-les allà i si cal fer les còpies per enviar-les al magatzem *off-site*.

Opció 2: Colocation Tier 2

Empresa (fictícia): CPDs Céspedes S.L.

Ubicació: Barcelona, a la Diagonal, prop de Francesc Macià.

- Es lloga una gàbia dins un data center existent. El preu és de 9000 euros/any per cada rack de fins a 48U (no importa si està ple o buit). S'han d'incloure servers, *switchos* i emmagatzematge.
- L'electricitat consumida es paga a part, per consum i potència contractada (vegeu apèndix 4), més un 40% adicional del consum en concepte de UPS i refrigeració. El Data Center té un PUE de 1,4. El lloguer inclou també la neteja tècnica i la recàrrega del sistema antiincendis, si aquest es disparés.
- L'empresa té redundància N+1 en UPS i HVAC i un sistema dièsel alternatiu. La seva certificació garanteix un 99,749% de *uptime* (és a dir, un màxim de 22 hores de *downtime* l'any).
- No hi ha la possibilitat de llogar una segona línia d'accés a la xarxa.
- Podem fer servir qualsevol sistema de *backup* i *mirroring* dels que hi ha a l'apèndix 3. Si volem un sistema de *backup on-site* i que ens ho gestionin, ens cobraran €2000 al més (inclou muntar, activar i desmuntar les cintes).
- Despeses addicionals:
 - Sistema de monitorització i anàlisi: +900 euros/any per *rack*. Inclou canviar hardware (com *spare disk*)

Opció 3: Colocation Tier 3

Empresa (fictícia): Mordor Colocation Center

Ubicació: Un edifici del 22@ (ubicació secreta)

- Es lloga espai dins un data center existent (gàbies). El preu és de 14000 euros/any per cada rack de fins a 48U (no importa si està ple o buit). S'han d'incloure servers, *switchos* i emmagatzematge.
- Hi ha el millor en seguretat (*mantraps*, vigilància, identificació redundat per password i empremta digital, ...). La monitorització està inclosa en el preu, el que inclou el canvi de hardware si s'escau (com *spare disc*).
- L'electricitat consumida es paga a part, per consum i potència contractada (vegeu document a part) més un 15% adicional del consum en concepte de UPS i refrigeració. El Data Center té un PUE de 1,15. El lloguer inclou també la neteja tècnica i la recàrrega del sistema antiincendis, si aquest es disparés.
- L'empresa té una replicació 2*N de tots els elements de HVAC i UPS. També té un generador dièsel amb capacitat per aguantar la potència pic durant 72 hores, que a més pot fer re-ompliment en calent. Tenen dues línies d'entrada d'electricitat i dues línies de connexió de xarxa. La seva certificació garanteix un 99,982% de *uptime* (és a dir, un màxim de 1,6 hores de *downtime* a l'any)
- El cost del lloguer ja inclou la connexió a dues línies d'electricitat (de fet, això és transparent a l'usuari). El servei de xarxa es paga a part, i heu d'escollir ample de banda

voleu contractar, i si voleu accés a la segona línia (vegeu apèndix 4). ATENCIÓ: només aquest *colocation* permet tenir duplicades les línies de connexió a la xarxa.

- La gestió local del *backup on site* no està inclosa, però per €3.500 al mes s'encarreguen de tot, incloent tenir una còpia de les cintes off-site i, si ha algun problema fer la recuperació.

Què heu de fer?

Amb l'ajut del full de càlcul "infraestructura" heu d'escollir quina empresa s'ajusta més a les vostres necessitats. Heu de fer-lo en funció de les característiques del vostre CPD, i heu de justificar-lo més enllà del cost (pot haver opcions més cares però més ajustades al que necessiteu en qüestions de seguretat o *downtime*). El full "infraestructura" només us donarà el cost de cada opció, per tal que no us passeu.

L'elecció l'heu d'incloure al full "Resum", amb altres paràmetres ja calculats com cost de l'energia.

En resum, heu d'indicar al full "Resum" l'opció escollida, si voleu gestió local de *backup on site* i si voleu monitorització (si s'escau, perquè amb la opció 3 ja està inclosa).

Insistim en que el full de càlcul és una eina que us traurà un munt de feina, però primer l'heu de controlar.

Apèndix 3: *Backup*

Seguint les indicacions de l'Apèndix 1 heu d'haver escollit tipus de cabina, nombre i la mida dels discos, el tipus (HDD o SSD), la configuració (tipus de RAID, amb /sense *spare disk*), etcètera.

Ara toca organitzar la part de *backup*.

Dades *off-site*

Sigui quin sigui l'escenari en que esteu, necessiteu un sistema *off-site*. Per fer això, podeu tenir un servidor de *backup* intern de cintes i moure-les físicament a un centre extern, o tenir el sistema extern per xarxa. Entre les coses que he de decidir hi ha:

- Cada quan es fa un *backup* (per simplificar l'escenari, no distingirem complert / incremental). Per fer *backups* en cinta heu d'escollir un sistema de cinta i calcular el cost de comprar-lo i aprovisionar-lo de cintes per cinc anys (el número de cintes depèn del número de *backups* que guardeu i la mida de cada un, vegeu el següent punt).
- Quins *backups* es guarden. Encara que només cal un *backup* complert per recuperar, és tradicional guardar algun més per recuperar dades esborrades per error. Un sistema tradicional per gent que fa un *backup* complert a la setmana és guardar l'últim de l'any , l'últim del trimestre i l'últim del mes. Per exemple, si avui és dimecres 15 de maig i els *backups* es guarden setmanalment la nit de dissabte a diumenge, ara tindria 4 còpies guardades: la de l'última setmana (12 de maig), l'últim del mes anterior (27 d'abril), l'últim del trimestre anterior (31 de març) i l'últim de l'any anterior (30 de desembre). En cas de fer un *backup* diari, es guardaria també el del dia en comptes del setmanal, però també l'últim de la setmana anterior.
- Heu de decidir si contracteu un *mirror* extern de les dades per recuperació ràpida. Un *mirror* redueix el *Recovery Point Objective*, però no elimina la necessitat de *backups* per recuperació de fitxers antics.
- Heu de decidir si teniu un *backup* extern amb transport de cintes o via xarxa.
- Si decidiu tenir un sistema de *backup* extern sense *mirror*, heu de tenir un *backup on-site* obligatòriament.

Per simplificar, en aquesta pràctica tenim empreses que ens lloguen els seus serveis. Mirant els preus i els paràmetres dels vostres escenaris heu de decidir si voleu duplicar dades, enviar-les encriptades i comprimides, etc...

Backup on-site

Per no complicar més la pràctica, només heu de decidir si voleu un sistema de *backup* de cintes *on site* o no, en funció de la resta d'eleccions pel que fa al *backup*. El full de càlcul ja calcula el model de robot i cost de les cintes en funció dels altres paràmetres. A mode informatiu, us direm que estareu escollint entre 4 models de biblioteca de cintes:

Opció	Nombre de cintes	Cost (Euros)	Consum (Watts)	Nombre d'Us
1	24	2500	168	2
2	48	3800	312	4
3	96	12000	560	8
4	512	70000	1500	40

Empreses que ofereixen *mirror* i/o *backup*

Opció 1: Microworks Apure

Empresa (fictícia): Microworks Apure.

Productes

Aquesta empresa ofereix dos productes:

- Replicació de dades fent un *mirror* de les nostres dades. Enviem totes les escriptures a disc per la xarxa i l'empresa manté una còpia de les nostres dades actualitzades. Aquesta còpia només es pot utilitzar per restaurar les dades (no per consultar-les). La recuperació es fa via xarxa.
- Sistema de *backup*. Aquesta empresa també ofereix un servei de *backup* per recuperació total del sistema o de fitxers determinats (per exemple, recuperació de fitxers antics o esborrats per error). Les recuperacions es fan també via xarxa i les podem gestionar des de la consola del CPD.

Podem contractar només el sistema de *backup*, o bé el sistema de *mirror* i el de *backup*.

Mirror + backup

- Si contractem el sistema de *mirror* i el de *backup*, aleshores utilitzarem el *mirror* cas de tenir que fer una recuperació total del sistema, i el de *backup* per recuperació de fitxers antics. Per tant no caldrà comprar una unitat de cintes ni les pròpies cintes. Això si: has de

comptar de tenir prou ample de banda amb l'exterior per poder enviar l'informació de cada escriptura de disc al nostre proveïdor del *mirror*.

- En cas de necessitar una recuperació total, al recuperar del *mirror* les dades que s'hauran perdut són molt poques (les transaccions s'envien a disc i per la xarxa al *mirror* al mateix temps), per tant considerem un RPO de segons. El temps de recuperació depèn de la quantitat de dades i l'ample de banda contractat amb Internet. L'empresa t'envia les teves dades a tota velocitat però la recuperació té un cost addicional: tenir la recuperació a màxima velocitat costa 0,1 euros per GB recuperat.
- El proveïdor fa un *backup* complet congelant les seves dades en un moment donat i guardant-les en cinta (es guarden totes les modificacions que arriben per la xarxa en discos auxiliars mentre es fa el *backup* i després s'actualitza el *mirror* amb les transaccions pendents).
- La recuperació de fitxers antics és gratuïta, ja que es programa i es fa aprofitant quan el sistema no està a càrrega màxima.

Només backup

- Si no tens contractat el *mirror*, aquest sistema està pensat per tenir un *backup on-site* i un altre *off-site*. Per tant has de comprar un sistema de cinta amb aprovisionament de cintes i programar quan vols fer els *backups*.
- Quan s'acaba de fer el backup al propi CPD, es transmet (xifrat) a l'empresa que l'emmagatzema *off-site*. Al ser una còpia del *backup* no té problemes de temps, pel que assumirem que es mira l'utilització de la connexió amb l'exterior i s'aprofita quan hi ha disponibilitat. Per tant no afegim pressió a les necessitats de connexió amb l'exterior.
- En cas de necessitar una recuperació total, al recuperar d'una còpia de *backup* les dades que s'hauran perdut depenen la data de l'últim *backup*. El temps de recuperació depèn de la quantitat de dades i l'ample de banda contractat amb Internet. L'empresa t'envia les teves dades a tota velocitat però la recuperació té un cost addicional: tenir la recuperació a màxima velocitat costa 0,15 euros per GB recuperat.
- La recuperació de fitxers antics és gratuïta, ja que es programa i es fa aprofitant quan el sistema no està a càrrega màxima.

Preus

Magatzematge *mirror*. Com és habitual en aquest tipus de negoci, es factura mensualment depenent de la quantitat de dades que guardis, d'acord amb la següent llista:

- First 1 TB / Month euros (e) 0.09 per GB
- Next 49 TB / Month e 0.075 per GB
- Next 450 TB / Month e 0.065 per GB
- Next 4,500 TB / Month e 0.055 per GB
- Over 5,000 TB / Month e 0.05 per GB

Backup sense mirror contractat: El proveïdor et factura el cost de mantenir les cintes de *backup*; el preu és de 0.02 euros per GB i còpia, més 100 euros mensuals per cada còpia a mantenir. Així, per exemple, suposem que fem un backup setmanal de 2 TB i mantenim l'últim *backup* anual, trimestral, i mensual, més el de l'última setmana, el que són són 4 còpies. Assumim una mitja de 4,3 setmanes / mes (52 setmanes / 12 mesos):

El cost mensual seria: 1 còpia/setmana * 4,3 setmanes/mes * 2048 GB/còpia * 0,02 euros/GB + 4 còpies guardades x 100 euros al mes per còpia = 575,128 euros/mes.

Backup amb *mirror* contractat. És bàsicament el mateix, però al tenir un *mirror* contractat et fan un preu especial: 0.011 euros per GB i còpia, més 100 euros mensuals per cada còpia a mantenir.

Opció 2: Monsoon S3

Empresa (fictícia): Monsoon S3 (*Secure Storage System*).

Ofereix el mateix que l'opció 1(per tant els detalls son iguals, consulteu-los a les pàgines anteriors) però té altres preus:

Preus

Magatzematge *mirror*. Es factura mensualment depenent de la quantitat de dades que guardis, d'acord amb la següent llista:

First 1 TB / Month e 0.095 per GB

Next 49 TB / Month e 0.078 per GB

Next 450 TB / Month e 0.07 per GB

Next 4,500 TB / Month e 0.065 per GB

Over 5,000 TB / Month e 0.055 per GB

Backup sense *mirror* contractat: El proveïdor et factura el cost de mantenir les cintes de *backup*; el preu és de 0.019 euros per GB i còpia, més 90 euros mensuals per cada còpia a mantenir.

Backup amb *mirror* contractat. És bàsicament el mateix, però al tenir un *mirror* contractat et fan un preu especial: 0.009 euros per GB i còpia, més 90 euros mensuals per cada còpia a mantenir.

Recuperació d'una còpia sencera. En cas de tenir contractat *mirror*, 0,05 euros per GB recuperat; en cas de no tenir-lo contractat, 0,07 euros per GB recuperat.

Opció 3: Transport de cintes de *backup*

Empresa (fictícia): *Take the tapes and run*.

Ofereix serveis de transport de cintes de *backup*, amb emmagatzematge a les seves instal·lacions.

- Els backups es poden ordenar de manera remota o presencial, però els fa la gent de la nostra empresa. L'empresa *Take the Tapes and Run* desplaça personal seu a la nostra empresa cada cop que s'ha fet un *backup*, i és aquest personal el que desmunta les cintes, les substitueix per les que es gravaran al proper *backup* i finalment s'emporta les cintes per guardar-les a la seva empresa.

- A les instal·lacions de l'empresa es guarden les cintes en armaris ignífugs, amb control de temperatura, humitat i escut per radiacions solars i magnètiques.
- La recuperació total o parcial (fitxers antics o esborrats), si es fa d'aquestes cintes requereixen el desplaçament del personal i el muntatge. La recuperació (la feina d'operació) és cosa nostra.
- És feina nostra comprar la màquina de cintes per fer el *backup* i les cintes necessàries per guardar les còpies que considerem necessari. És elecció nostra quantes còpies velles en guardem i cada quant fem un *backup*.
- Igualment, i si no volem que recuperacions parcials (fitxers vells o esborrats) impliquin el cost de desplaçament, podem tenir un segon joc de cintes i fer còpies del *backup*, de manera que tenim una còpia a la nostra empresa i una còpia *off-site* per si de cas.

Preus:

- Transport + personal especialitzat: 240 euros / backup
- Emmagatzematge a les instal·lacions de la pròpia empresa: 100 euros/any per pack de 20 cintes.

Què heu de fer?

Recordeu que heu de contractar un *backup off-site*.

Heu d'escollir si voleu o no un *mirror*, i amb quina empresa el contracteu.

Si voleu un *mirror*, no cal tenir *backup on-site*.

Haureu d'omplir al full "Resum" l'empresa contractada, si voleu o no *mirror* més d'altres paràmetres com nombre de còpies que emmagatzemem del *backup* i cada quan es fa un *backup* complert. El full de càlcul "Backup" us indicarà els costos en funció d'aquests paràmetres (a introduir al full "Resum"), però l'elecció depèn de molt més que del seu cost; l'anàlisi de riscos que heu de fer i l'impacte de negoci serà molt més determinant: dades que es poden perdre, opcions de recuperació, temps que trigo en recuperar una còpia...

Apèndix 4: Consums d'ample de banda i electricitat

Connexió amb l'exterior

El *Network Service Provider* (NSP) dels nostres *colocation services* ha muntat una xarxa 100G a partir d'equips de diversos fabricants¹, el que significa que usant la tecnologia Dense Wavelength Division Multiplexing (DWDM), pot transportar 100Gbps per cada longitud d'ona que viatja per la fibra. En funció del fabricant, cada equip permet transportar entre 88 i 96 longituds d'ona (canals) per cada fibra, i per tant, l'ample de banda que pot oferir per fibra és de 8.8Tbps a 9Tbps. Per tant, tenim disponible un ample de banda molt gran.

Dades rellevants per a l'activitat 1

Internament, les empreses de *housing* que us hem proposat, posen a la nostra disposició, per un preu detallat a la següent taula, la interconnexió cap a l'exterior a través d'interfícies Ethernet.

Opció	Velocitat de sortida	Tipus de connector que té muntat el centre de colocation	Preu mensual (€)
1	10Mbps	1GbE RJ45	0,63
2	100Mbps	1GbE RJ45	6,3
3	1Gbps	1GbE RJ45	63
4	10Gbps	SFP-10G-SR	630
5	100Gbps	CFP-100G-SR10	6300

Per fer-lo bé, nosaltres hauríem de tenir en el nostres switches d'agregació el connector indicat a la taula, i ells s'encarreguen de fer-nos arribar el cable i donar connectivitat IP cap a l'exterior. Aquesta part del disseny del CPD, però, la fareu a la propera activitat.

Per tant, l'únic que heu de calcular és quin ample de banda necessiteu. L'ample de banda necessari ve donat per les especificacions del vostre escenari, més el que es necessiti per mantenir la comunicació necessària amb el *mirror* (si és que en teniu un). Tingueu en compte que és possible agregar el tràfic a través de diferents canals dels llistats (per exemple aconseguir 200Gbps de sortida a partir de dos connector de 100Gbps cadascun) i,

¹ Els equips muntats són una combinació de Cisco ONS 15454 Trunk Cards, Nokia Siemes Networks hiT 7300, Fujitsu FLASHWAVE 9500 i Alcatel-Lucent 1830 Photonic Service Switches. Més informació a la segona part del curs.

òbviament, pagant el cost de totes les línies individuals que haguem usat. Heu d'incloure al full "resum" el tipus de canal contractat i el número de línies que llogueu. El full de càlcul "Bandwidth provider" és un auxiliar per fer els càlculs.

L'empresa Mordor Colocation Center (i només aquesta) ofereix l'oportunitat de tenir contracte amb dos proveïdors diferents. En aquest cas, i per simplicitat, considerarem que el proveïdor secundari ens cobra un 40% del preu de la línia (doncs només està allà per si cau la principal). Per exemple, si necessito un ample de banda de 2Gbps tindria que agafar 2 línies agregades de 1Gbps cadascuna. Si agafo l'opció de replicació el preu pujarà un 40% respecte al preu de les dues línies sense replicació. Si al full "Resum" indiqueu el tipus de línia, el nombre de línies i si voleu accés al segon proveïdor, el full de càlcul ja actualitza costos sense que tingueu que fer res més.

Un problema addicional és el temps de recuperació de les dades. Si hem de recuperar un *backup* complet, l'ample de banda també marcarà el temps de recuperació, i ja que no podem tornar a explotació fins que hem recuperat les nostres dades, també influeix en el temps en que estem en *downtime*.

Consum elèctric

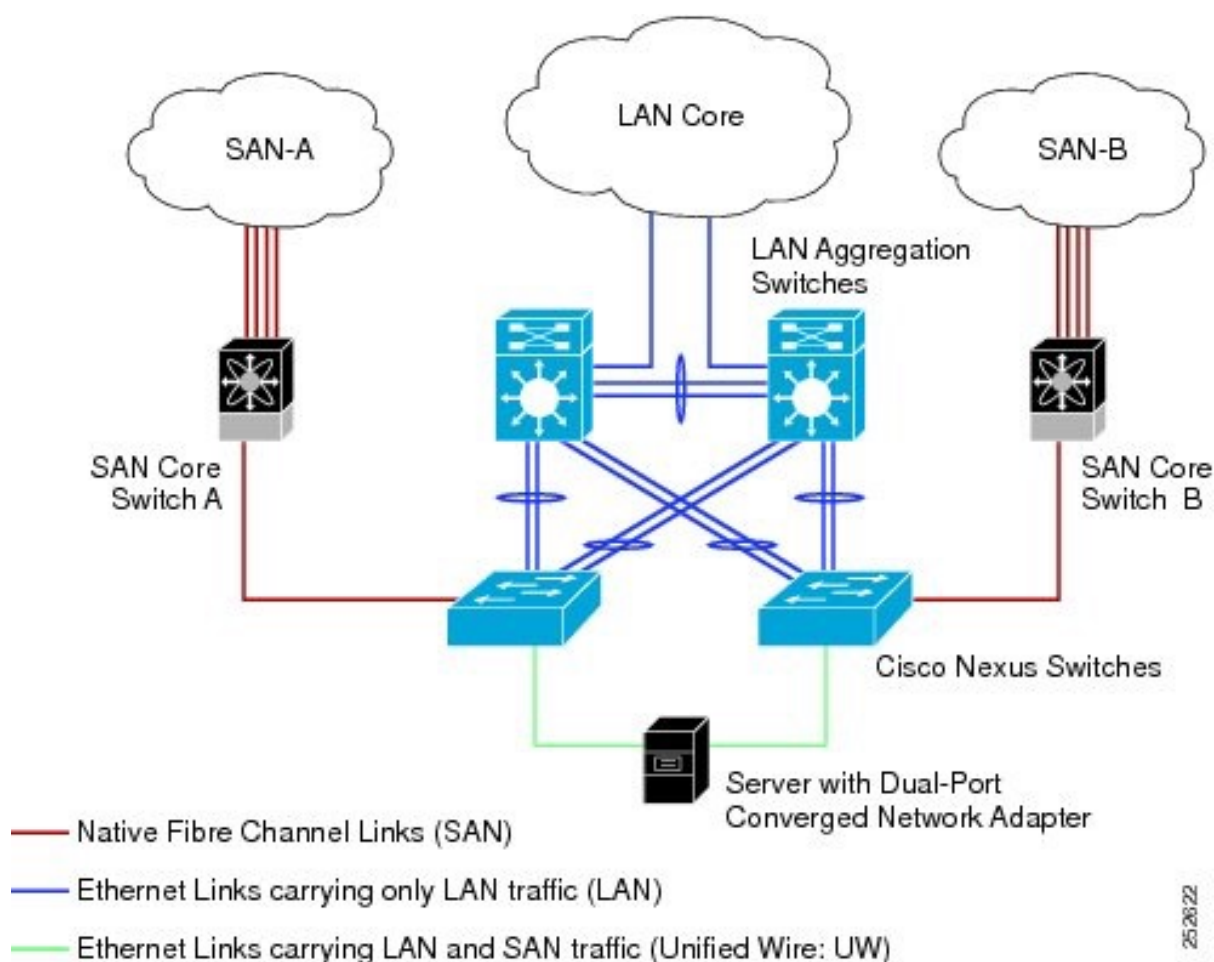
La factura de la llum no és fàcil de calcular. A Espanya si tens un contracte de línia d'alta tensió (20kV) en el model tarifari 6.1A (l'habitual en l'entorn en que ens movem), l'electricitat té diferents preus per períodes. Hi ha 6 períodes, i quin et toca en un moment donat depèn de si estem en cap de setmana o no, de l'hora del dia i del mes en que estem (per allò de les hores de llum i tal). Hem simplificat el nostre model assumint un consum continu i oferint-vos un full de càlcul per calcular els resultats. El full de càlcul es diu "Electricitat". Els preus són els vàlids per 2013 però no pensem actualitzar aquest fitxer cada any. En aquest fitxer només varien dues caselles:

- Els kW consumits pel vostre equipament: ve donat pel vostre escenari (full escenari), sumant-li el que consumeixin els elements que afegiu per muntar el sistema de emmagatzematge i *backup* (full cabinet de discos).
- Quin sobreprovisionament voleu contractar: si espereu un consum determinat, és habitual contractar més potència per seguretat. Aquest número afecta poc al resultat final, però ens vindrà donat pel nostre escenari.

El cost final (anual) apareix a la taula final. Es pot veure a **Cost Total Energia Anual** i és el que es computarà calculant amb el PUE de cada opció de *Housing*.

Apèndix 5: NAS o SAN?

L'elecció de si feu servir NAS o SAN depèn de les característiques del vostre escenari. Bàsicament de la pressió que afegeix sobre la xarxa Ethernet l'accés a disc, més les avantatges de fer servir SAN. En els dos cassos tindrem una xarxa similar a la d'aquesta solució de Cisco:



NAS

En el cas d'escollir NAS només teniu la LAN (part blava) i no la part verda (connexió FC amb els racks de disc), ni la SAN (vermella). Els discos estarien directament connectats a la LAN. Per tant, si escolliu un NAS, no heu d'afegir cap cost ni econòmic ni en consum al model del vostre escenari.

SAN

En el cas de SAN, necessitem la xarxa addicional (vermella), més una connexió LAN/discos (verda) que requereix CNAs (*Converged Network Adapter*), un *Network Interface Card* (NIC) que conté tant connexions per Fiber Channel (HBA: *Host Bus Adapter*) com TCP/IP Ethernet NICs. En particular, al nostre exemple usariem dos Nexus que porten 32 ports Ethernet, i als que se'ls hi pot muntar una expansió amb 8 ports Fiber Channel a 1/2/4/8 Gbps.

Per tant, en el cas de necessitar una SAN, aleshores s'afegiran:

- Dos Nexus
- Un HBA per cada servidor.
- Dos switch (redundància) en cada rack de servidors.
- Els SFPs (*Small Form-factor Pluggable*) necessaris per muntar l'expansió de Fiber Channel
- Els cables necessaris
- El switch CORE (replicat)

Si voleu veure el que costa la SAN, només cal que indiqueu a la casella SAN del full "Resum" el que voleu. **Un 0 si no volem SAN, un 1 si volem una SAN de 8Gbps, un 2 si volem una SAN de 16Gbps, un 3 si volem una SAN de 32Gbps, un 4 per una SAN de 64Gbps i un 5 per una SAN de 128 Gbps. Per simplicitat només admetem aquestes opcions.**

Què heu de fer?

Primer decidir si pel tràfic que teniu amb el sistema d'emmagatzematge i el tipus d'informació que manegueu val la pena tenir una SAN. Si no val la pena, poseu al full "Resum" un zero a l'opció SAN.

Si penseu que val la pena, aleshores podeu activar l'opció SAN (un 1, 2, 3, 4 o 5 a "Resum", depenent de l'ample de banda de la SAN que voleu) i mirar quan cost afegeix. Tenir SAN afegeix dos costos: incrementa el consum (això es suma automàticament al full "Electricitat", no cal que feu res més) i augmenta el CAPEX (també automàticament).

Si els costos de la SAN superen el que podeu pagar, podeu reduir el cost d'altres coses, o bé renunciar a la SAN, però si renuncieu heu de calcular l'impacte que pot tenir en el rendiment final.

Apèndix 6: Cada quan falla?

Per presentar als inversors la seguretat estudiada, necessitem números de la fiabilitat dels diversos elements. Els números que hi ha a continuació són aproximats, però al temps molt realistes pel que fa als que es fan servir pels professionals.

- Suposant un funcionament 24/7, probabilitat anual que falli un disc és:

SSD Consumer	1,6 (% anual)
SSD Enterprise QLC	0.97 (% anual)
SSD Enterprise MLC-TLC	0.45 (% anual)
HDD Consumer	2,84 (% anual)
HDD Enterprise <10000 rpm	2,84 (% anual)
HDD Enterprise >= 10000 rpm	2,84 (% anual)

- Possibilitat de detectar que un disc està a punt de fallar usant SMART: 70% de les fallades son previsibles i podem aprofitar l'*spare disc* (si en posem)
- Si falla un disc en RAID 5 o 6 i s'ha de reconstruir es pot fer en calent (no cal aturar la màquina) però durant el temps de reconstrucció els discos d'aquella cabina van a la meitat d'IOPS. El temps de reconstrucció habitual és de 4 hores per TB de disc.
- Caiguda de la xarxa elèctrica: calculem una aturada de menys d'una hora un cop cada dos anys, una de entre 1 i 6 hores cada 4 anys. Les línies d'alta tensió industrials estan preparades per garantir un màxim de 6 hores per reconnexió.
- Caiguda de la xarxa de dades: calculem una caiguda de 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.
- El temps de recuperació de *backup* depèn de les IOPS dels discos on són les dades a recuperar i de l'ample de banda de la xarxa contractada per la qual recuperem dades del *mirror/ backup*. En cas de recuperació d'un sistema *on-site*, els sistemes de cintes de *backup* que hem incorporat a l'escenari poden recuperar fins a 5TB/hora (compte que són un munt de IOPS, el límit segurament estarà en els vostres discos)