

Boletín 1 de Ejercicios de Enxeñería de Computadores

1. Tenemos un rack de 42U. Queremos poner el mayor número de servidores posibles con 4 discos duros cada uno. Tenemos 2 opciones:

- Servidores en formato rack de 1U con capacidad para 4 discos duros.
- Servidores en formato blade con capacidad para 2 discos duros.

El chasis de blades permite incluir 14 servidores en una altura de 9U.

Para completar el número de discos de los blades disponemos además de módulos de discos duros con capacidad para 14 discos en una altura de 3U.

¿Con qué opción podremos introducir el mayor número de servidores? ¿Cuál es el número máximo de servidores en formato blade con esa configuración que se pueden incluir en el rack?

2. El consumo por servidor rack del ejercicio anterior es de 313 W. Calcular el consumo por servidor blade equivalente si tenemos los siguientes datos:

- Chasis con 14 blades consume 3248W
- Módulo con 14 discos 404 W

El KWh cuesta 0.12 euros. ¿Cuánto es el gasto en electricidad al día para cada tipo de servidor?

3. Calcular el MTBF de un sistema *Cable – Switch – Cable*.

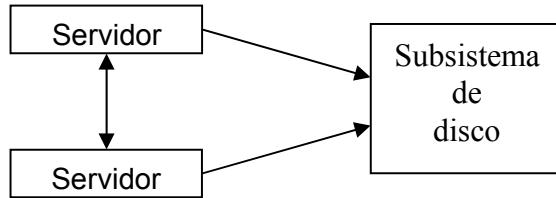
- MTBF cable = 100.000 horas
- MTBF switch = 50.000 horas.

¿Es menor el MTBF del sistema que el de algún componente individual? ¿Tiene esto sentido?

Ayuda: En un sistema de varios componentes (sin redundancia), el inverso del MTBF del sistema completo se calcula como la suma de los inversos de los MTBF de los componentes.

4. Sea el siguiente sistema con servidores clusterizados (redundantes):

- MTBF servidor = 200 horas
- MTBF cables Fibre Channel = 100,000 horas
- MTBF Subsistema de Discos = 20,000 horas



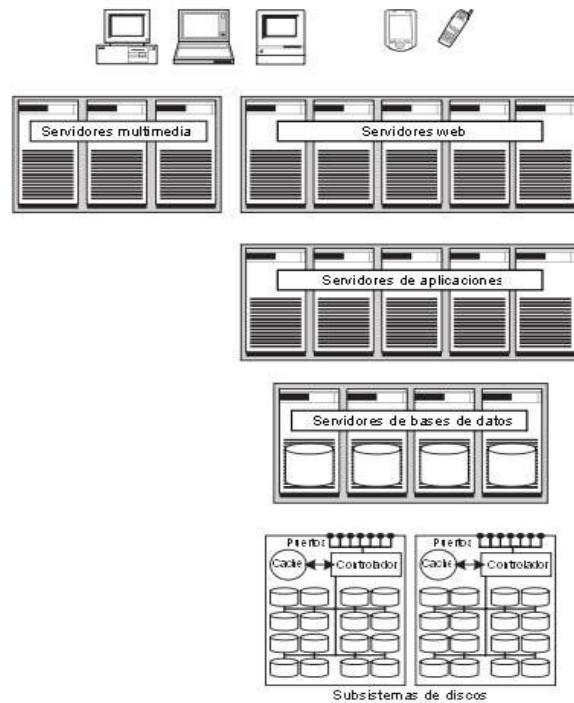
- a) Obtén que componentes tienen que fallar para que falle todo el sistema
- b) De acuerdo con el modelo anterior obtén el MTBF del sistema

Ayuda: Vamos a aproximar el MTBF de un subsistema redundante como la suma de los MTBF de los componentes. Para los elementos no redundantes se aplica la fórmula del ejercicio anterior.

5. Supongamos conjuntos de 3, 4, 5, 6, 7 y 8 discos. ¿Cuál es el porcentaje de espacio útil en cada caso si se utilizan para implementar RAID 5? ¿Y si se implementa RAID 6?
6. Supongamos conjuntos de 3, 4, 5, 6, 7 y 8 discos, configurados como JBOD. Suponiendo que el MTBF es de 150,000 horas, ¿cuál es el MTBF de cada JBOD?
7. Supongamos conjuntos de 3, 4, 5, 6, 7 y 8 discos. Con cada uno de los conjuntos se implementa un RAID 5. Suponiendo que el MTBF es de 150,000 horas, ¿cuál es el MTBF del RAID? Suponer que la probabilidad de fallo de cada disco es independiente, de tal forma que, tras fallar 1 disco, los restantes se comportan como si su MTBF partiese de cero. Suponer también que no hay regeneración del disco que ha fallado.
8. Supongamos el caso anterior. Pero ahora la probabilidad de fallo de los discos es tal que, tras fallar 1 disco, los restantes se comportan como si su MTBF estuviese corriendo desde que se constituyó el RAID original.
9. Repetir los 2 ejercicios anteriores si se implementa RAID 6.
10. Es posible combinar distintos niveles de RAID. Como ejemplo, RAID 0 y RAID 1 se pueden combinar creando RAID 10 o RAID 0+1. De igual manera, es posible combinar RAID 0 y RAID 5. Dibujar las 2 posibles variantes teniendo en cuenta que: para RAID 0 se utilizarán 3 volúmenes y que para RAID 5 se utilizarán 3+1 volúmenes. Mostrar cómo se reparte la información entre los discos físicos suponiendo bloques llamados: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q y R.
11. Disponemos de 6 switches FC de 8+8 puertos cada uno. Construir con ellos un árbol del tipo fat-tree. ¿Cuántos dispositivos podemos conectar mediante la red resultante? ¿Puede haber bloqueo en la red (*entendemos que hay bloqueo si conexiones entre dos pares de puertos deben compartir link*)? ¿Hay implementada suficiente redundancia?

12. Con los switches del ejercicio anterior construir un árbol del tipo skinny-tree que permita conectar el doble de dispositivos que en el ejercicio anterior. ¿Puede haber bloqueo en la red? ¿Hay implementada suficiente redundancia?
13. Tenemos switches de 8+8 puertos, y los utilizaremos para conectar entre si a 20 servidores con 2 HBAs cada uno y 4 SSdeDisco con 8 puertos cada uno. ¿Cómo haríamos para conectarlos si los switches están conectados en anillo? ¿Cuántos switches necesitaríamos como mínimo? Tener en cuenta que las conexiones entre switches ocupan puertos y que han de ser redundantes.
14. Repetir el ejercicio anterior, pero conectando los switches en forma de fat-tree con 3 niveles de switches.
15. Una red de almacenamiento (SAN) que debe dar servicio a 4 servidores y 4 SSdeDisco. Implementarla de forma que haya redundancia a nivel de conexiones. A continuación, añadir redundancia a nivel de switch. Finalmente, añadir redundancia a nivel de SAN. Indica que topología de red has usado.
16. Dibujar una infraestructura compuesta de un cierto número de servidores de cálculo, un servidor de ficheros y una SAN. Hacerlo de tal forma que todos los servidores de cálculo puedan acceder a la SAN y al servidor de ficheros y que éste último tome sus datos de la misma SAN.

17. En la siguiente figura se muestra una típica arquitectura de servidores de internet en 5 niveles: desde los usuarios al almacenamiento. Se han omitido las redes que los mantienen unidos. Tratar de reconstruir la red teniendo en cuenta que los distintos niveles se conectan entre ellos y que ninguno de los servidores dispone de una cantidad de almacenamiento significativa aparte de la que pueda obtener de los niveles inferiores y de los SSdD.



①

Opción 1

Como cada servidor tiene capacidad de para 4 discos:

$$\text{nº servidores} = \frac{42 \text{ U}}{1 \text{ U/servidor}} = \underline{\underline{42 \text{ servidores}}}$$

Opción 2

Para tener servidores con 4 discos cada uno necesitanse como mínimo 2 módulos de disco. De esta forma tiene un chasis de blades con 14 servidores con 2 discos cada uno, e 2 módulos de discos que aportan 28 espacios para discos (14/modulo), dando como resultado 14 servidores con 4 discos cada uno. Esta configuración ocupa 15 U. No racking caben 2 configuraciones juntas (30 U), con 28 servidores. No obstante os 12 U restantes podrían completarse con otros chasis de blades e otros módulos de discos - quedando así 42 U (más completo). Cabe destacar que no hay discos para todos los servidores (17 discos para 14 servidores) - haciendo que solo

La medida dos servidores en chasis de blades (7) tiene 4 discos. En total podríense poner 35 servidores en formato blade con 4 discos cada un no rack.

Podrígense introducir más servidores no rack en formato blade (sempre e condicón non requirén módulos de discos): 56 servidores en 36 U.

Con 4 discos cada un, podríense introducir 35 servidores en formato blade, ocupando o rack completo (42 U).

②

Servidor en formato rack

$$\begin{aligned}
 X \text{ €/día} &= 42 \text{ servidores} \cdot 373 \text{ W/servidor} \cdot \frac{24 \text{ h}}{1 \text{ día}} \\
 &\quad \cdot \frac{1 \text{ kW}}{1000 \text{ W}} \cdot 0,12 \text{ €/kWh} \\
 &= \underline{\underline{37,86 \text{ €/día}}}
 \end{aligned}$$

Servidor en formato blade

$$\begin{aligned}
 X \text{ €/día} &= (2,5 \text{ chasis} \cdot 3248 \text{ W/chasis} + 5 \text{ módulos} \cdot 404 \text{ W/módulo}) \\
 &\quad \cdot 24 \text{ h/día} \cdot 1 \text{ kW/1000 W} \cdot 0,12 \text{ €/kWh} \\
 &\quad \underline{\underline{29,20 \text{ €/día}}}
 \end{aligned}$$

(3)

$$\frac{1}{MTTF_{sistema}} = \frac{2}{MTBF_{cable}} + \frac{1}{MTBF_{switch}}$$

$$MTBF_{sistema} = \frac{1}{\frac{2}{20000h} + \frac{1}{5000}} = \underline{\underline{25000\text{ h}}}$$

O MTBF do sistema é menor que o de qualquer dos componentes individuais. Isto é lóxico xa que se aumentamos o número de componentes também estamos aumentando a probabilidade de que algén deles sufra un erro, polo que estamos diminuindo o MTBF do sistema en consecuencia.

(4)

a) Poden darse varias situacions nos cales falle todo o sistema;

1. Falla o SSD: 20000 h

2. Falla o servidor 1 e o cable FC do servidor 2: 200200 h

3. Falla o servidor 2 e o cable FC do servidor 1: 200200 h

4. Fallan os dous cables FC: 200000 h

5. Fallan os dous servidores: 400 h

6) O MTBF serán o inverso da suma dos inversos dos MTBF de cada situación de fallo, específica:

$$MTBF_{sistema} = \frac{1}{\frac{1}{20000} + \frac{2}{100200} + \frac{1}{200000} + \frac{1}{400}} \text{ h}$$

$$= \underline{\underline{388,86 \text{ h}}}$$

(5) En RAID 5 ocúpase o espacio de 1 disco para paridades, mientras que en RAID 6 se ocupa o espacio de 2 discos. A porcentajes útil calcularíase:

$$\cdot RAID 5 \rightarrow \left(1 - \frac{1}{n}\right) \cdot 100 \%$$

$$\cdot RAID 6 \rightarrow \left(1 - \frac{2}{n}\right) \cdot 100 \%$$

Sendo n o número de discos.

Assí entón, as porcentaxes de espacio útil serían

	RAID 5	RAID 6
3 discos	66,67	/
4 discos	75	50
5 discos	80	60
6 discos	83,33	66,66
7 discos	85,71	71,43
8 discos	87,5	75

Como comentar que RAID 6 con 3 discos no es posible, ya que para xesar a perdida necesitaríanse 2 discos de datos, e neste caso só hai 1.

⑥ As estor configurados como JBOD, o MTBF do conxunto de discos calcúlase da seguinte maneira:

$$MTBF_{JBOD} = \frac{150000}{n} \text{ h}$$

sendo n o número de discos.

Así entón, os MTBFs serían:

Nº de discos	MTBF do JBOD(h)
3	50000
4	37500
5	30000
6	25000
7	21428,57
8	18750

(7)

Así estar configurados en RAID 5, para que o RAID falle tengan que fallar 2 discos, en que nese caso perderíase a información de paridad necesaria para reconstruir a información. Así pois, o MTBF do RAID sería o MTBF do conxunto sen que fallase ningun disco máis o MTBF tras fallan un disco:

$$MTBF_{RAID} = \frac{150000}{n} + \frac{150000}{n-1} \text{ h}$$

sendo n o número de discos.

Así entón, os MTBFs serían:

Nº de discos	MTBF do RAID(h)
3	725000
4	87500
5	67500
6	55000
7	46428,57
8	40178,57

8) Agora ao non ser os fallos entre discos independentes, o MTBF do RAID seria:

$$MTBF_{RAID} = 2 \cdot \frac{150000}{n} \text{ h}$$

Así entón, a taboa que durá:

Nº de discos	MTBF do RAID(h)
3	100000
4	75000
5	60000
6	50000
7	4285714
8	37500

9) Exercicio 7

Ag ser RAID 6, para que o RAID falle teñen que fallar 3 discos, ao estar a paridade duplicada. Ag ser os fallos dos discos independentes, o MTBF é:

$$MTBF_{RAID} = \frac{150000}{n} + \frac{150000}{n-1} + \frac{150000}{n-2} \text{ h}$$

Así entrén, a tu bocan serviría:

Nº de discos	MTBF do RAID(h)
3	<u> </u>
4	262500
5	117500
6	92500
7	76428,57
8	65178,57

Exercicio 8

Así serían algunos los fallos dependentes, o MTBF serían:

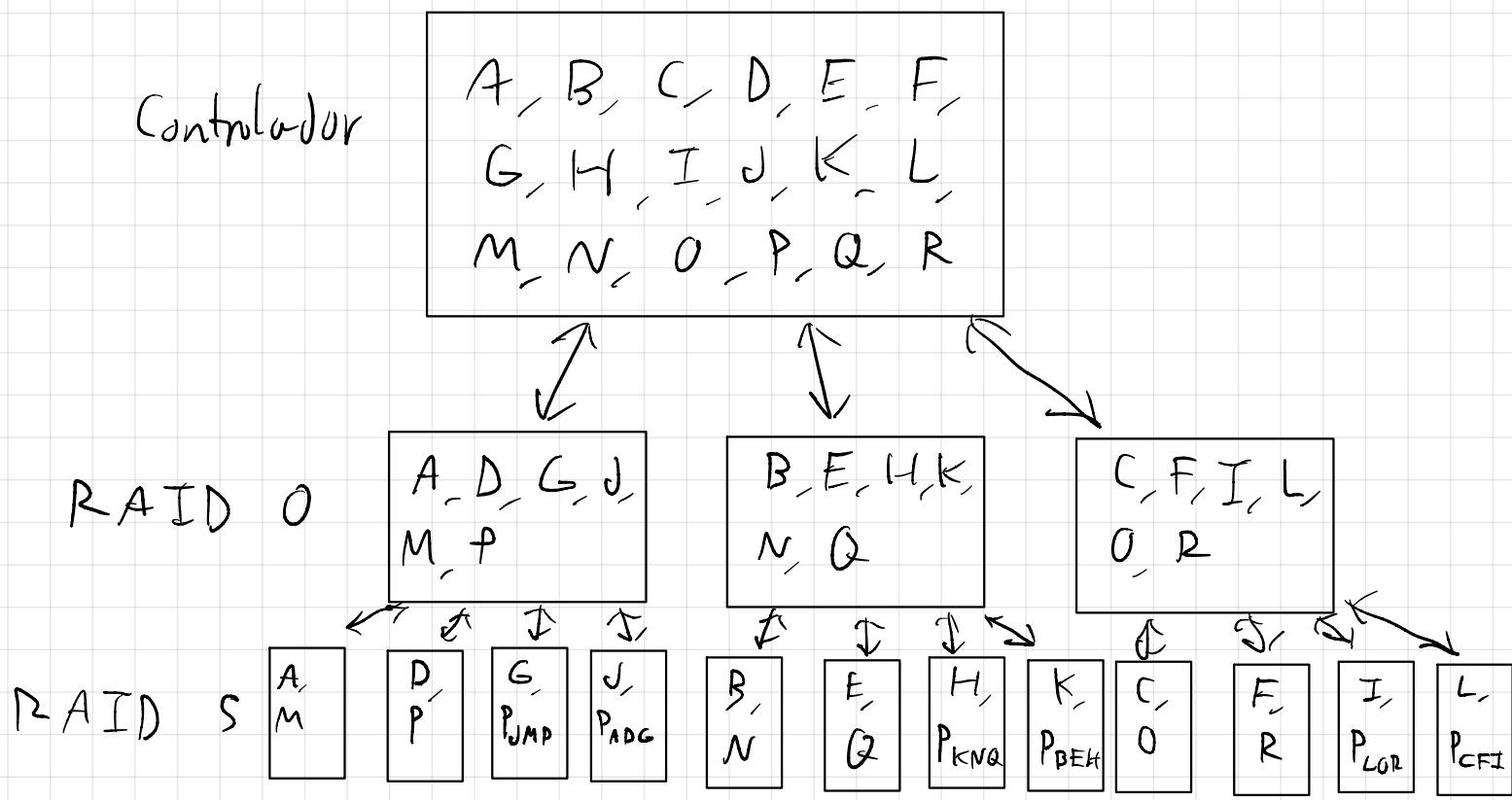
$$MTBF_{RAID} = 3 \cdot \frac{150000}{n} \text{ h}$$

Así entrén, a tu bocan serviría:

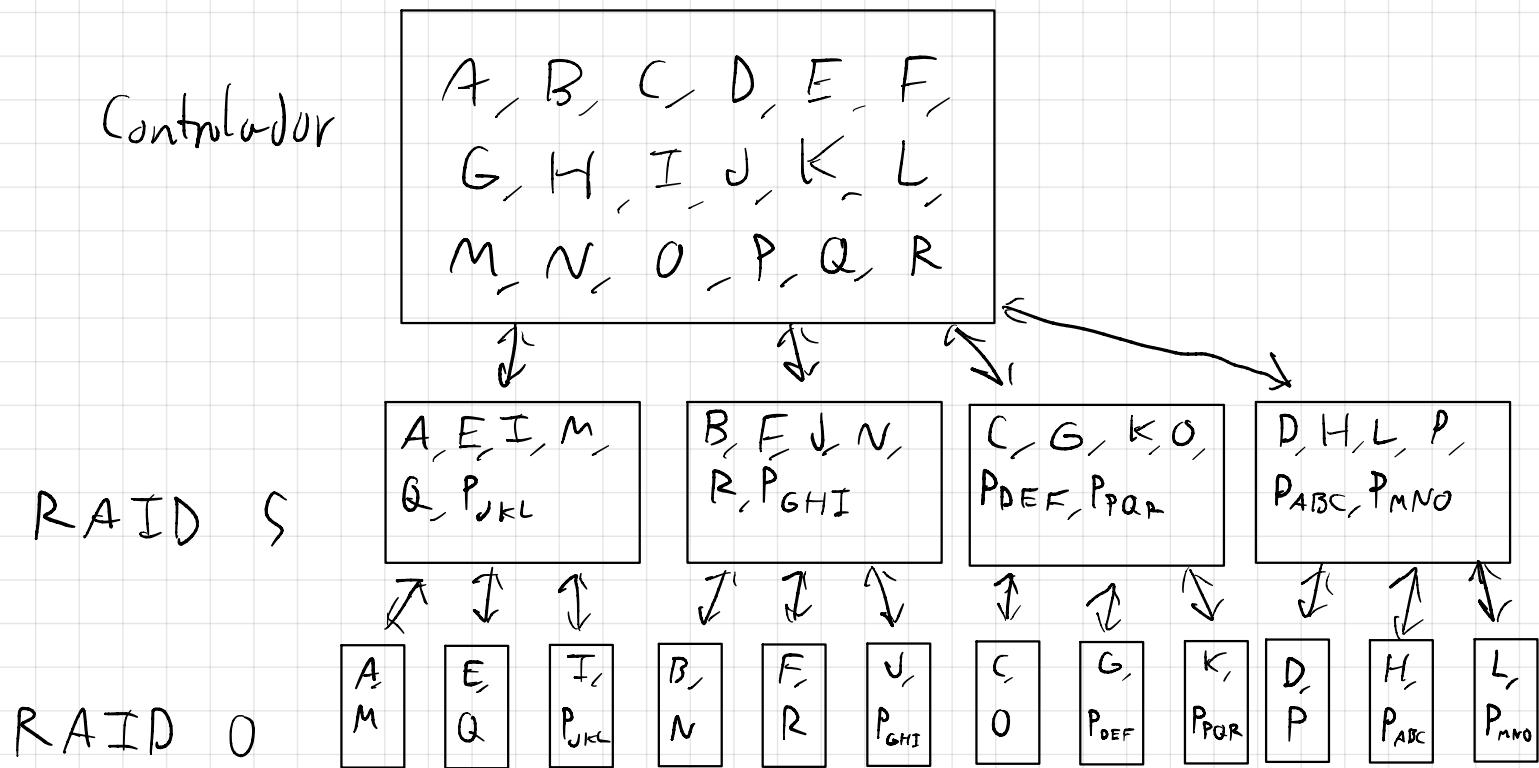
Nº de discos	MTBF do RAID(h)
3	<u> </u>
4	112500
5	90000
6	75000
7	64285,71
8	56250

10

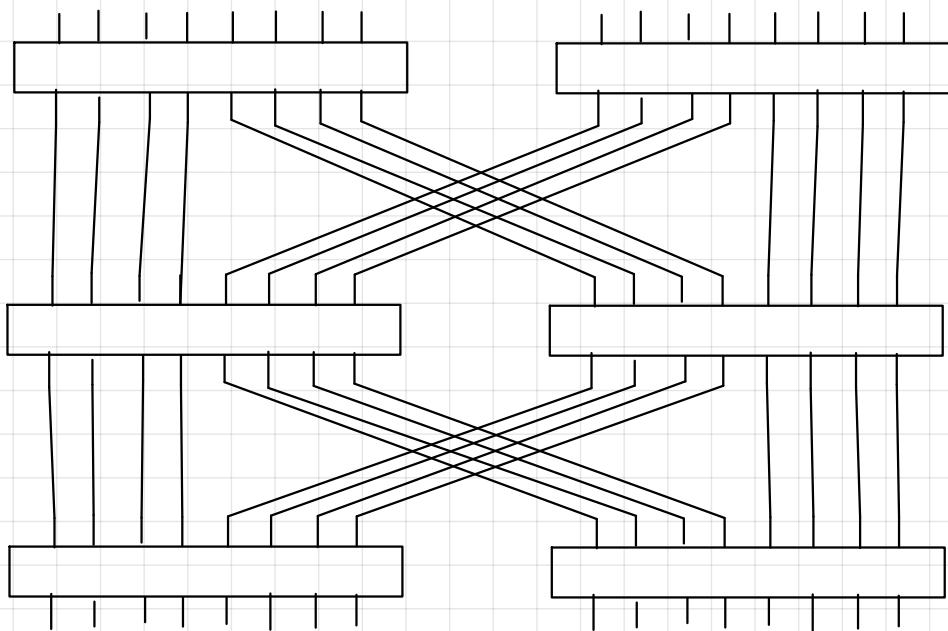
RAID 50



RAID 0 + 5

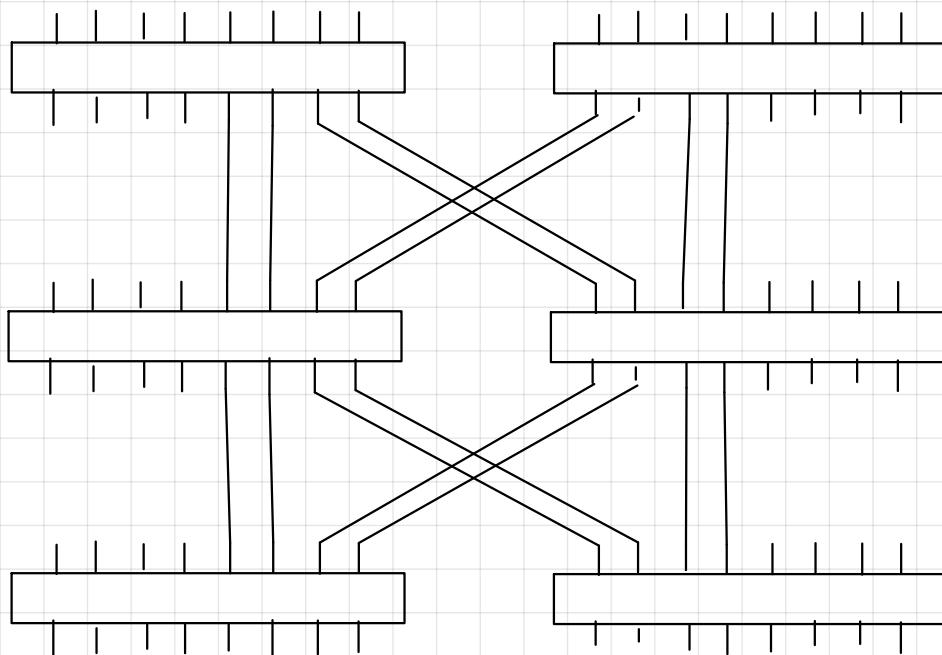


11



Mediante a rede resultante pódense conectar 32 dispositivos. Non pode haber bloqueos de rede xa que todas as conexións entre portos ferán un camiño libre sempre (se falla algúñ link, si que aparecerá bloqueo). Poderíase dicir que si que hai suficiente redundancia, pois aínda que falte algúñ link ou switch intermedio existen camiños que permiten o correcto funcionamento.

12



Desta forma poderían conectarse 6 dispositivos. Non obstante, agora pode haber bloqueos na rede xa que as conexións non son as suficientes para permitir a todos os dispositivos comunicarse á vez. Agora taménouco hai redundancia suficiente xa que calquera switch que falle provocaría a perda de conexión de dispositivos da rede.

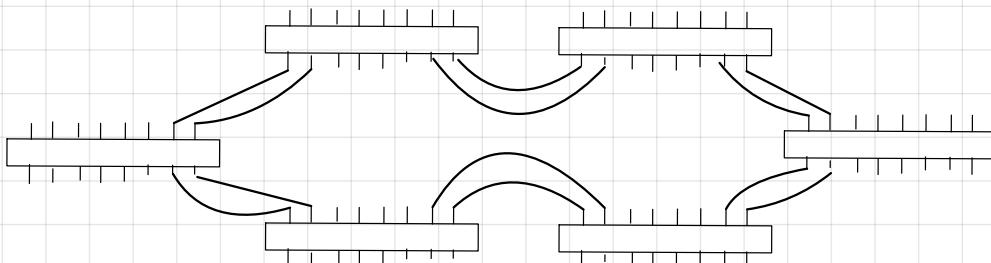
13

Necesitarse un total de 72 portas:

- 40 para servidores
- 32 para SSDs

Así estar a rede en anel, cada switch terá 4 ISLs, polo cal lle quedan 72 portas para dispositivos. Son necesarios 6 ($\frac{72}{12}$) switches como mínimo.

Os switches quedarián conectados así:



O ideal sería que cada servidor se conectase a 2 switches distintos (1 HBA a cada switch) e cada SSD se conectase aos 6 switches (2 portas a 2 switches e 1 puerto a cada un dos 4 switches restantes).

14

Así ser un fat tree, os switches das capas superior e inferior só poderán usar o 50% dos seus portas para dispositivos. Así pois, necesitarse 9 ($\frac{72}{8}$) switches entre as capas superior e inferior.

to ter que haber os mismos switches en cada capa, este valor será o inteiro imediatamente maior a metade dos switches necessarios: $5 \left(\lceil \frac{9}{2} \rceil\right)$

Os switches que darán conectados de forma que cada switch da capa superior/inferior se conectarán aos switches da capa intermediaria com doble conexión a 3 deses switches e con conexión simple aos 2 restantes. Isto é debido a que o número de portas(8) non é múltiplo do número de switches por capa(5). O ideal sería que as conexións dobles se repartisen de forma que a redundancia fose o máis igual posible para todos os switches e conexións.

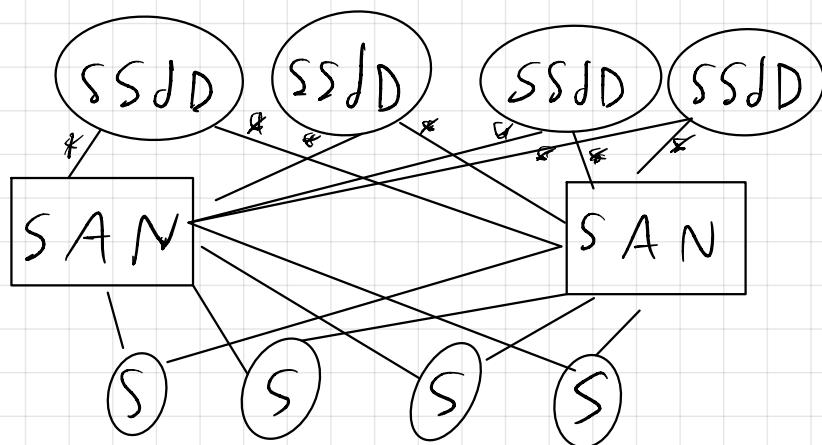
En contras as conexións cos servidores e SSJDs, como no exercecicio anterior o ideal servía repartir as conexións entre switches: cada servidor a 2 switches e cada SSJD, a 8. Sobrará 8 portas, xa que en total hai 80(2.5.8).

15

Supóngase que los servidores tienen 2 HBAs y los SSDPs, 8 portos. Así pues, necesitarse 40 portos en total.

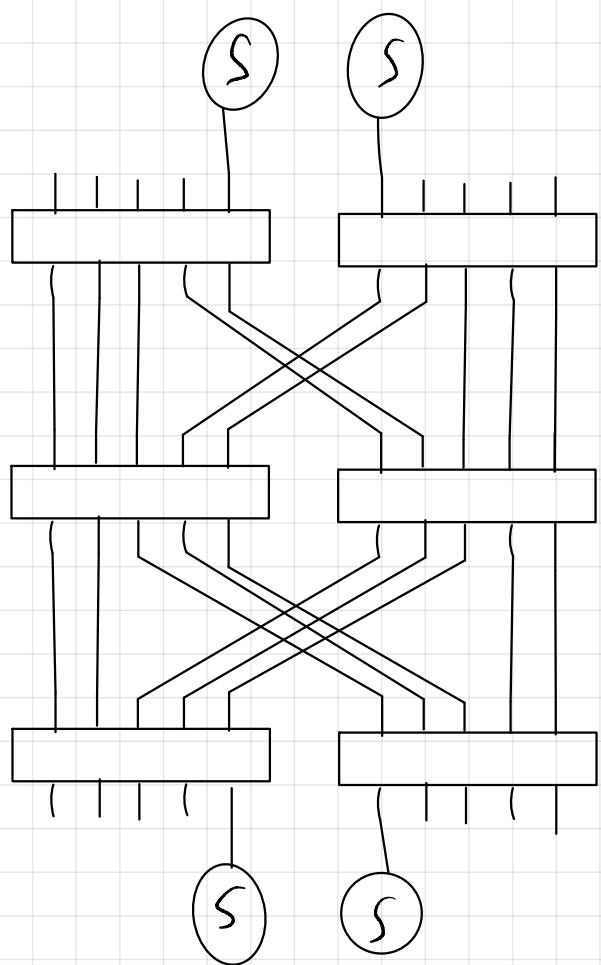
Como hay que asegurar a redundancia a nivel de conexión, switch e SAN, decidirse crear una red de tipo fat tree con 2 switches S15 en cada capa. A cada red conectarse 1 HBA por servidor y 4 portos por SSDP.

A rede sería que daría así:



• Cada conexión de SSDP a SAN equivale a 4 portos.

Dentro de cada bloco SAN tiene o seguinte:



Os portos restantes
son para os SSDs.
Cada un dar 4 portos
de cada SSD conecta-
se a un switch
distinto, de forma
que estean conectados
a un porto de
cada un dos 4
switches das capas
superior e inferior.

16)

?

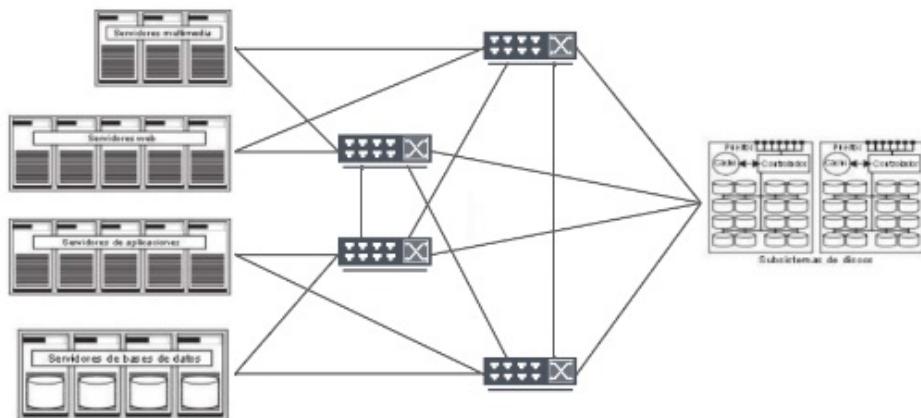
17

Se ha decidido diseñar una SAN con topología de malla de cuatro switches. Cada servidor se conecta a dos switches, y los SSdD se conectan a los cuatro switches. De esta forma se logra:

- Redundancia a nivel de conexión: si falla una conexión, bien sea entre un servidor y un switch o entre switches, queda la otra para garantizar la disponibilidad.
- Redundancia a nivel de switch: si falla un switch de la SAN, queda otro para garantizar la disponibilidad de los dispositivos que había conectados a este.
- Alta disponibilidad de los SSdD: como los servidores no disponen de una cantidad de almacenamiento significativa, es importante garantizar una alta disponibilidad de los subsistemas de discos, ya que sin estos la aplicación dejaría de funcionar. Esto se consigue conectándolos a cuatro switches en lugar de a dos, logrando así una mayor redundancia.

Se ha decidido no implementar redundancia a nivel de red por ser demasiado costoso para los requisitos de disponibilidad del sistema, que se entiende que consiste en una aplicación normal y corriente y, por tanto, no requiere una muy alta disponibilidad.

Así entonces, la SAN sería de la siguiente forma:



Como protocolo de la red, suponiendo que los servidores y subsistemas de discos están en distintos puntos geográficos, se podría usar iSCSI.