

1 Conceptos básicos de técnicas de almacenamiento remoto e redundante

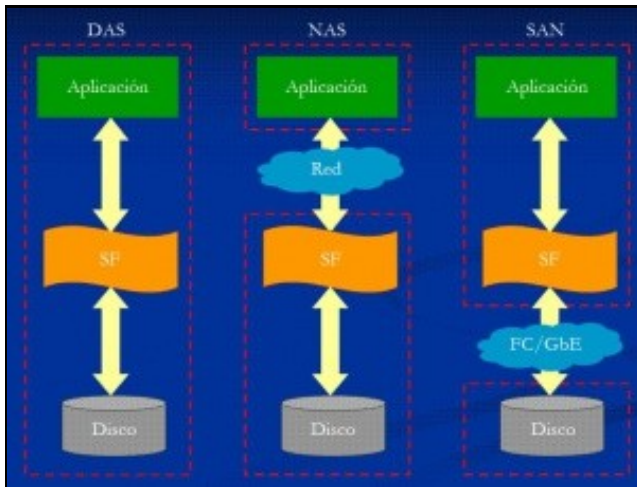
Neste apartado imos abordar os conceptos básicos e as tecnoloxías implicadas nos sistemas de almacenamento remoto e redundante.

Non se vai facer unha explicación pormenorizada de moitos destes conceptos, xa que a análise con profundidade dalgúns deles podería ser obxecto dun curso completo; o obxectivo é proporcionar unha visión xeral de todo o que hai detrás destes sistemas para que en cada momento poidamos saber que é o que estamos facendo e por que; e que alternativas poderían existir segundo as nosas necesidades.

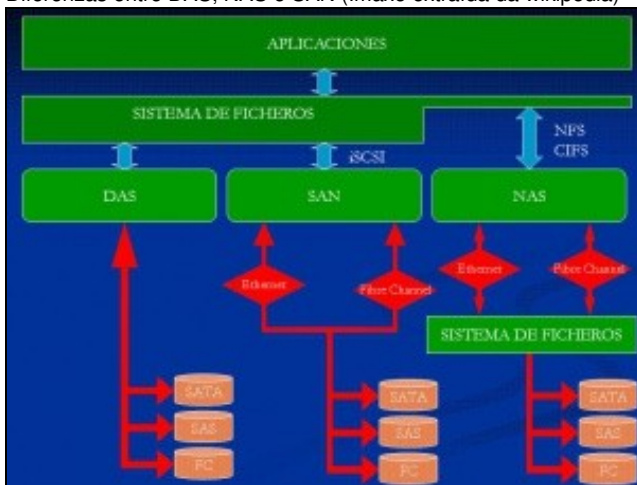
1.1 Sumario

- 1 DAS, NAS e SAN
- 2 Sistemas de ficheiros
- 3 Protocolos NAS: CIFS e NFS
- 4 RAID
 - ◆ 4.1 Niveis de RAID
 - ◇ 4.1.1 Niveis estándar
 - ◇ 4.1.2 Niveis aniñados
 - ◇ 4.1.3 Niveis propietarios
 - ◆ 4.2 Implementación dun RAID

1.2 DAS, NAS e SAN



Diferenzas entre DAS, NAS e SAN (Imaxe extraída da wikipedia)



Localización do sistema de ficheiros en cada caso (Imaxe extraída da wikipedia)

En primeiro lugar imos ver que significan estes tres acrónimos, que veñen supoñer as tres alternativas polas que podemos optar para unir nun sistema informático o subsistema de procesamento dos datos do subsistema de almacenamento dos datos:

- **DAS(Direct Attached Storage):** Esta opción é a que seguramente se teña agora mesmo no ordenador no que se está visualizando este material. Nela, o dispositivo de almacenamento está conectado directamente ao núcleo de procesamento que fai uso del (é dicir, o disco duro que está montado dentro do noso ordenador). Exemplos de tecnoloxías dentro deste grupo poden ser IDE, SATA, SCSI, SAS, USB, etc.
- **NAS(Network Attached Storage):** Neste caso, o subsistema de almacenamento comparte a través da rede os seus datos utilizando un protocolo de rede no nivel de aplicación, como pode ser CIFS, NFS, FTP, etc. Nestes protocolos, o *cliente* (o subsistema de procesamento) solicita os ficheiros ao *servidor* (o subsistema de almacenamento) para logo procesalos localmente, pero a xestión do sistema de ficheiros, permisos e usuarios faise no propio sistema de almacenamento.
- **SAN(Storage Area Network):** Este é o caso máis sofisticado, e no que se separa totalmente o subsistema de almacenamento do de procesamento. A diferenza coa alternativa anterior radica en que nesta alternativa o subsistema de almacenamento exporta directamente os volumes de almacenamento, e o equipo que accede a eles pode utilizalos como se fosen volumes locais. Ata hai pouco tempo, esta alternativa só se podería implantar usando a tecnoloxía de *canal de fibra (fiber channel)*, o cal supoñía xa meterse nunha inversión moi considerable (necesítanse interfaces de fibra tanto para os servidores como para os sistemas de almacenamento, así como switches SAN e cableado de fibra). Porén, a aparición do estándar *iSCSI*, permite implantar unha solución equivalente con Gigabit Ethernet, utilizando cable de par trenzado e switches e interfaces de rede Ethernet que podemos adquirir con un prezo moitísimo menor. Grazas ao uso de conexións a 1Gbps podemos conseguir, utilizando unha conexión específica entre os servidores e os sistemas de almacenamento (é preferible usar unha interface de rede nos servidores específica para a conexión á rede SAN, para non aumentar a latencia no acceso aos datos dos discos), un rendemento moi próximo ao do canal de fibra con moito menor custo.

1.3 Sistemas de ficheiros

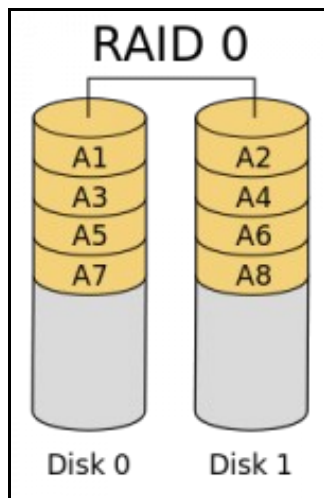
- O **sistema de ficheiros ou de arquivos** é a estrutura que podemos darlle a un espazo de almacenamento, que pode ser unha partición dun disco, un disco completo ou varias particións ou discos unidos.
- O sistema operativo é o que vai facer uso desa estrutura para poder ler e escribir os datos no espazo de almacenamento, e polo tanto ten que recoñecer esa estrutura para poder facer uso dese espazo.
- Existe un **gran número de sistemas de ficheiros** que son utilizados en distintos sistemas operativos, aínda que un S.O. pode ser capaz de recoñecer múltiples sistemas de ficheiros, ben de forma nativa ou mediante a instalación dalgún módulo de software adicional.
- Os sistemas de ficheiros máis utilizados nos sistemas Windows son **FAT** e **NTFS**, e nos sistemas GNU/Linux os máis comúns son o **estendido (ext2/ext3/ext4)**, **JFS** e **XFS**, entre outros. Ademais, Linux é capaz de recoñecer os sistemas de ficheiros propios de Windows, como FAT e NTFS.
- ¿Pero, cal deles é o mellor? Aquí está o gran problema, xa que existen moitos criterios polos que podemos comparar distintos sistemas de ficheiros: máximo tamaño dos volumes, máximo tamaño nos ficheiros, velocidade de lectura (e aquí hai que distinguir entre lectura de ficheiros grandes e pequenos), de escritura (tamén variará dependendo do tamaño dos ficheiros), consumo de RAM, flexibilidade na configuración dos volumes, etc. E o caso é que non imos atopar ningún sistema de ficheiros que "gañe" en todos os criterios, xa que en moitos casos para mellorar un aspecto sempre hai que sacrificar algún outro. Polo tanto, o que teremos que facer é buscar cal dos sistemas de ficheiros dispoñibles satisfai mellor as nosas necesidades en cada caso concreto.
- Imos describir as características xerais dos sistemas de ficheiros que utiliza **FreeNAS**, que é o sistema de almacenamento que imos utilizar:
 - ♦ **UFS (Unix File System):** É un sistema de ficheiros utilizado nos sistemas operativos derivados de **BSD** (hai que ter en conta que FreeNAS está baseada en **FreeBSD**). Admite volumes e ficheiros de ata 8ZB de tamaño, e nomes de ficheiros de ata 255 bytes. A partir da versión 9.3 de FreeNAS, elimínase este sistema de ficheiros do xestor de volumes e só se permite importar volumes xa creado por compatibilidade con anteriores versións.
 - ♦ **ZFS:** Este sistema de ficheiros foi deseñado por Sun Microsystems. Neste caso permite volumes e ficheiros de ata 16 Exabytes, pero ademais ofrece moitísimas funcionalidades para os sistemas de almacenamento, como facer instantáneas, engadir dinamicamente dispositivos de almacenamento a un volume, a creación de grupos ou *pools* de almacenamento e diversos mecanismos para comprobar a integridade dos datos almacenados, entre outras. Como contra, ten un consumo de recursos bastante elevado, xa que se recomenda ter un mínimo 8GB de RAM para obter un bo rendemento, e para discos de máis de 6TB, recoméndase un mínimo de 1GB de RAM por cada TB de disco. Así que é un sistema de ficheiros moi adecuado para sistemas de almacenamento grandes e nos que se conte con bastantes recursos, xa que se non o seu rendemento será pobre. Este é o sistema de ficheiros que se recomenda utilizar con FreeNAS.
- FreeNAS tamén soporta os sistemas de ficheiros NTFS e FAT, pero co obxectivo de utilizalos para volumes exportados por iSCSI que sexan xestionados polo sistema operativo do servidor que accede a eles.

1.4 Protocolos NAS: CIFS e NFS

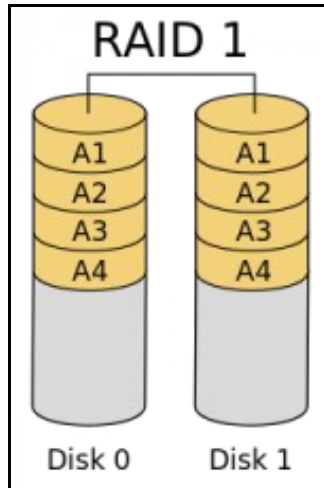
Imos describir as características máis salientables dos protocolos utilizados nos sistemas *NAS*, que son CIFS e NFS:

- **CIFS (*Common Internet File System*)** é o nome co que se coñece actualmente ao protocolo *SMB*, que é o protocolo utilizado nas redes Windows para compartir carpetas e impresoras, e comunicar os distintos equipos da rede. Polo tanto, é o protocolo que estamos utilizando cando compartimos unha carpeta nun equipo Windows (que fai o papel de servidor) e accedemos a ela dende outro equipo Windows da rede (que fai o papel de cliente); pero utilizando o módulo de *samba*, que é unha implementación deste protocolo para sistemas GNU/Linux, un equipo con Linux tamén pode usar este protocolo para acceder a unha carpeta compartida por un equipo Windows ou para compartir unha carpeta á que poderá acceder un equipo Windows. Dado que *samba* implementa os dúas partes do protocolo (cliente e servidor), incluso podemos utilizalo para que unha máquina Linux acceda a unha carpeta compartida por outra máquina Linux. Neste protocolo, cando un equipo quere acceder a un recurso compartido noutro equipo, debe autenticarse proporcionando unhas credenciais (un nome de usuario e contrasinal) que sexan válidas no equipo que comparte o recurso e teñen outorgado o permiso necesario para acceder ao recurso.
- **NFS (*Network File System*)** é o protocolo para sistemas de ficheiros en rede utilizado nos sistemas Unix e GNU/Linux. Igual que no caso de CIFS, NFS tamén divide o sistema en dúas partes: o servidor, que é o equipo que comparte ou exporta unha carpeta de datos, e o cliente, que é o equipo que accede á carpeta compartida. O gran problema deste protocolo está na seguridade, xa que no seu propio deseño ten un gran inconveniente: para autenticar o usuario que accede á carpeta compartida, o servidor confía na autenticación do equipo cliente, xa que en lugar de requirir as credenciais do usuario para comprobar a súa validez, *fíase* do identificador de usuario que lle indique o equipo cliente. Polo tanto, se o equipo cliente se ve comprometido e envía ao servidor un identificador de usuario falso, un usuario vai poder acceder a información non autorizada. Isto vai supoñer un grave problema para o uso de NFS en moitos contornos.

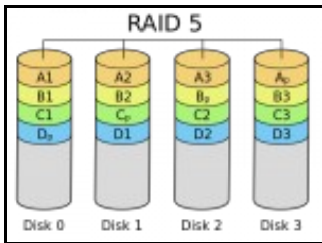
1.5 RAID



Estrutura dos datos en RAID 0



Estrutura dos datos en RAID 1



Estructura dos datos en RAID 5

Os sistemas **RAID** (*Redundant Array of Independent Disks*) son sistemas de almacenamento que permiten utilizar varios dispositivos de almacenamento formando un único volume, distribuíndo e replicando o datos entre eles. Grazas a replicación dos datos que se almacenan nos discos, os sistemas RAID permiten manter a información almacenada incluso aínda que fallen un ou varios discos do conxunto.

1.5.1 Niveis de RAID

Existen distintas variantes de RAID que se diferencian na forma en que se distribúen e replican os datos nos distintos discos do conxunto. Algunhas delas son estándares, outras se forman combinando varias estándares, e outras que son variantes propietarias que poden facer algunhas ampliacións ou melloras sobre as estándares.

1.5.1.1 Niveis estándar

Os niveis de RAID estándares máis utilizados son:

- **RAID 0:** Este nivel de RAID non é en realidade redundante, xa que os datos non se repiten en distintos discos senón que se van almacenando de forma alternativa en dous discos. O que acadamos con dous discos nun volume en RAID 0 é ter máis espazo de almacenamento (a suma do tamaño dos dous discos) e aumentar a velocidade das lecturas e escrituras, sempre que se poida facer a lectura e escritura nos dous discos de forma simultánea.
- **RAID 1:** Tamén coñecido como RAID en espello (*mirror*), o que fai é gardar cada sector de datos nos dous discos que forman o espello, de forma que os dous discos almacenan os mesmos datos. O que conseguimos así é que se algún dos discos falla mantemos toda a información, pero a costa de perder a metade do espazo de almacenamento.
- **RAID 5:** Co obxectivo de reducir o espazo de disco desperdiciado, RAID 5 agrupa un mínimo de tres discos e o que fará é distribuír os datos a gardar nun sector de cada disco (formando unha franxa ou *stripe*), pero gardando nun dos discos (que se vai cambiando en cada franxa) un sector de paridade que se obtén a partir dos sectores dos demais discos. Desta forma, se falla algún dos discos do RAID, a información pode obterse utilizando o resto dos discos do RAID, coa vantaxe de que só perdemos o tamaño de un disco do conxunto. Polo tanto o espazo desperdiciado será menor relativamente cantos máis discos teñamos no conxunto, pero tamén temos que ter en conta que o fallo de dous discos si que provocará a perda dos datos.
- **RAID 6:** É unha variante similar a RAID 5 que utiliza dous discos de paridade para poder soportar fallos de dous discos simultáneos.

1.5.1.2 Niveis aniñados

Poden combinarse distintos niveis de RAID para dar lugar a novas estruturas. Por exemplo, se unimos dous volumes RAID 0 en espello teremos un RAID 0+1 ou RAID 01, e se unimos dous volumes RAID 1 nun RAID 0 teremos un RAID 1+0 ou RAID 10. Da mesma forma, podemos ter un RAID 30, 50, 100, etc.

1.5.1.3 Niveis propietarios

Ademais dos niveis comentados, existen outros niveis definidos e utilizados por fabricantes ou produtos concretos. Por exemplo, o **RAID Z** é unha variante de RAID definido polo sistema de ficheiros **ZFS** para mellorar e optimizar o funcionamento de RAID 5 e RAID 6 sobre este sistema de ficheiros.

1.5.2 Implementación dun RAID

- Un sistema RAID pode implementarse de dúas formas:
 - ♦ **RAID por hardware:** Neste caso, o RAID é xestionado por unha controladora RAID (un hardware) que xestiona os distintos discos e ofrece de cara ao sistema operativo un único disco que engloba todo o RAID, de forma que o S.O. non saberá que realmente hai varios discos redundantes (e polo tanto o sistema operativo non terá que incorporar ningún módulo nin haberá que configurar nada para utilizar o RAID). Para configurar e xestionar o RAID faremos uso dunha aplicación propia que xestione a controladora RAID.

- ♦ **RAID por software:** Neste caso, é o sistema operativo o que configurará e xestionará o RAID, polo que non será necesario ningún hardware especial e haberá que facer a configuración no S.O. Esta será a alternativa que veremos neste curso usando FreeNAS.

-- Antonio de Andrés Lema e Carlos Carrión Álvarez --