

UT 4: Sistemas de almacenamiento de información



2ºSMR – Seguridad Informática

Almacenamiento Información

Un factor importante de Seguridad es saber dónde se puede almacenar la información.

Cuando se trata de la seguridad física de la información hay que tener en cuenta:

- **Rendimiento:** nos referimos a la capacidad de cálculo de un ordenador. Es importante sirva información/datos en menor tiempo posible.
- **Disponibilidad:** capacidad de los sistemas de estar siempre en funcionamiento. Sistema de alta disponibilidad está compuesto por sistemas redundantes o que trabajan en paralelo, de modo que cuando se produzca un fallo en el sistema principal arranque los secundarios automáticamente.
- **Accesibilidad:** se refiere a la capacidad de garantizar el acceso a personas autorizadas a los procesos, servicios y datos de los que dispone la empresa sin poner en riesgo la seguridad de la información.

Almacenamiento Información



disco duro



DVD



CD



pendrive



tarjeta SD



Memory Stick



disco duro portátil



Disquete

ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

Consiste en un conjunto de técnicas hardware y software que utilizan varios discos y se caracterizan por proporcionar:

- **Tolerancias a fallos** se aumenta la disponibilidad y hay más garantía de integridad de los datos.
- **Mayor capacidad** combinando varios discos se consigue mayor capacidad de almacenamiento que utilizando discos por separado
- **Fiabilidad** medida del tiempo de funcionamiento sin fallos.

ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID

Redundant **A**rray of **I**nexpensive **D**isk

Redundant **A**rray of **I**ndependent **D**isk

ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID Redundant Array of Independent Disks

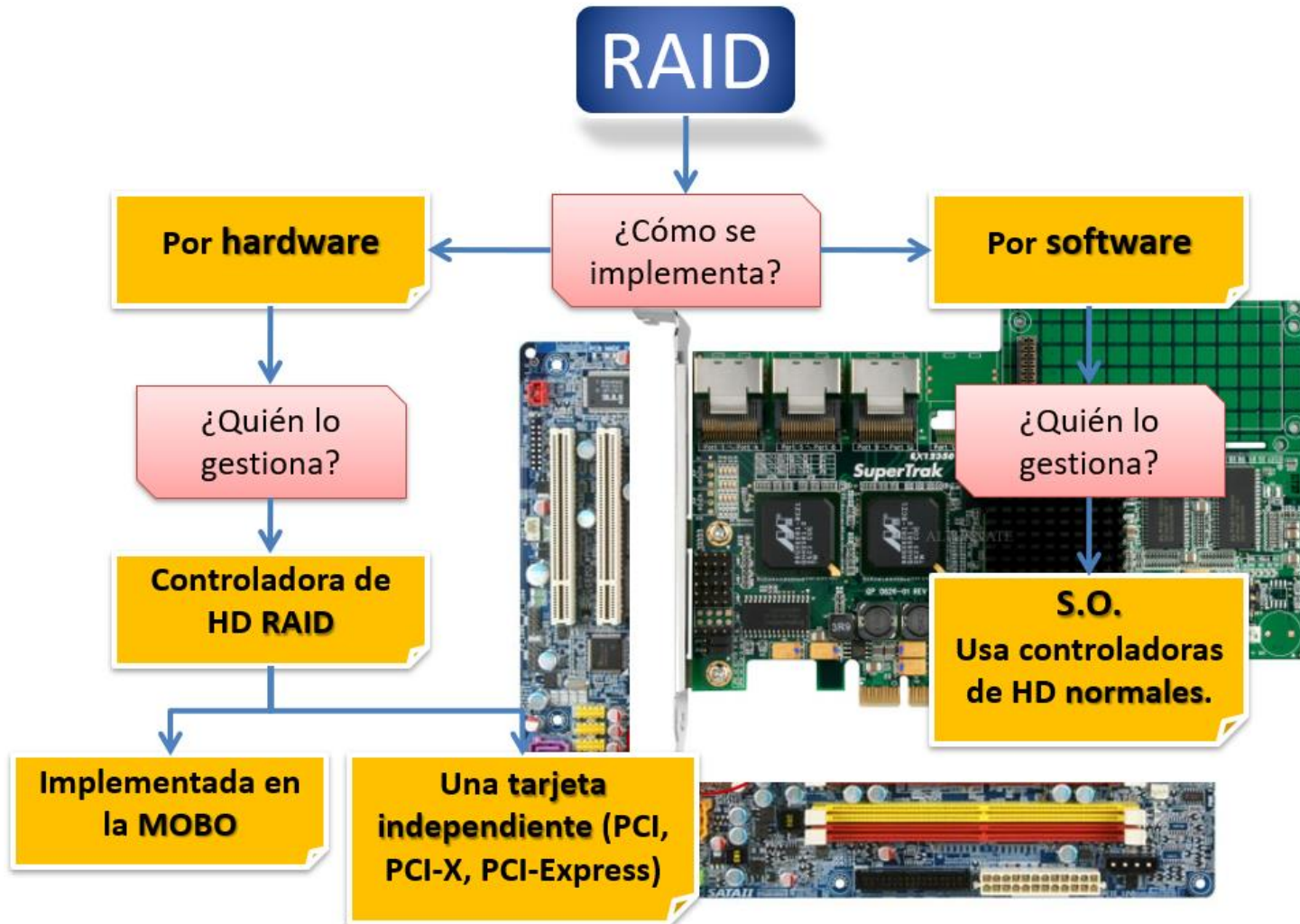
Conjunto de **varios discos duros** (un set, en inglés) actúan bajo el control de un hardware específico, de tal manera que **para sistema funcionan como una única unidad de almacenamiento.**

La información se distribuye o replica ofreciendo un sistema con:

- **Mayor tolerancia a fallos:** si se produce un error con RAID el sistema será capaz (en algunos casos) de recuperar la información perdida.
- **Mayor seguridad:** es más tolerante con los fallos al mantener cierta información duplicada. Se aumenta la disponibilidad y hay más garantía de integridad de los datos.
- **Mayor velocidad:** en algunos casos cierta información repetida y distribuida podrá realizar varias operaciones simultáneamente y provoca mayor velocidad.

ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID Redundant Array of Independent Disks



ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID Redundant Array of Independent Disks

La tecnología RAID (excepto en el nivel 0) protege los datos contra la caída de un disco duro.

Si se produce un fallo, RAID mantiene el servidor activo y en funcionamiento hasta que se sustituye la unidad defectuosa.

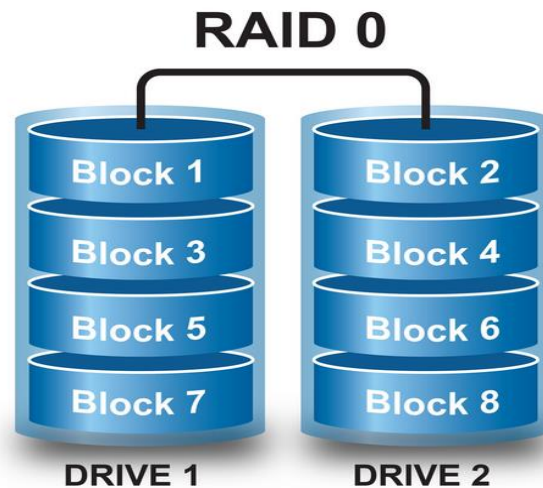
Todos los sistemas RAID suponen la pérdida de parte de la capacidad de almacenamiento de los discos para almacenar los datos de paridad y conseguir la redundancia.

RAID ofrece varias opciones, llamadas **niveles RAID**. Cada nivel de RAID ofrece un equilibrio distinto entre **tolerancia a fallos (redundancia)**, **rendimiento y coste**, diseñadas para satisfacer las diferentes necesidades de almacenamiento de los usuarios

ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-0 Redundant Array of Independent Disks

Se necesitan dos discos. Agrupamos discos para tener un disco más grande y más rápido:



En RAID-0 se utilizan dos técnicas:

- Striping
- Spanning

ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-0 Disk Striping

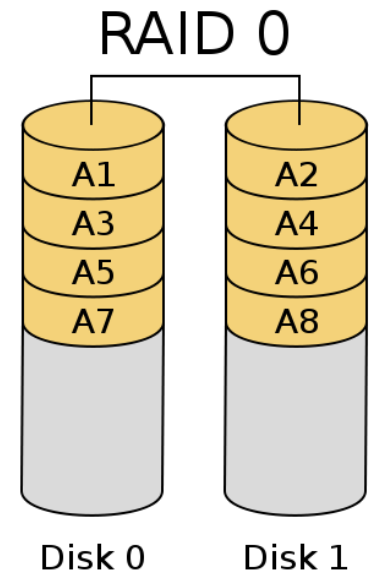
Datos se distribuyen de forma equilibrada entre los dos discos.

Los bloques de la unidad A se almacenan de forma alternativa entre los discos-0 y 1 de forma que los bloques impares de la unidad se almacenan en el disco-0 y los bloques pares en el disco-1.

Esta técnica **favorece la velocidad**: se pueden leer y escribir datos almacenados en dos discos diferentes simultáneamente (los discos deben estar gestionados por controladoras diferentes).

El **RAID-0 no incluye ninguna información redundante**.

Si se produce un fallo en cualquier disco provoca la pérdida de información en dicha unidad.

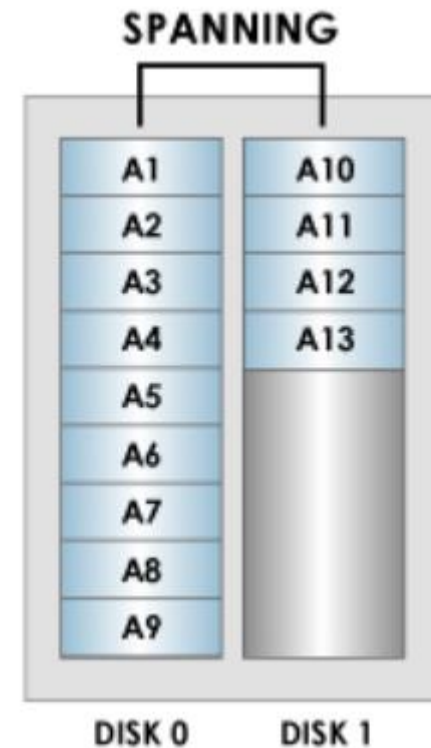


ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-0 Disk Spanning

Esta técnica es más rápida que la anterior.

Los bloques de la unidad A se escriben en el primer disco hasta que se llena y entonces pasa al siguiente y así sucesivamente.



ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-1 Mirroring

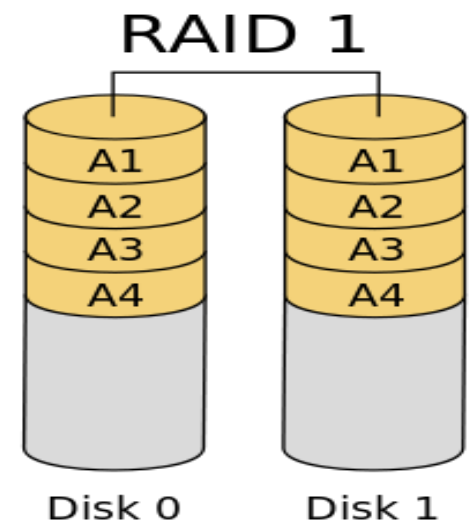
Consiste en mantener una copia idéntica de la información de un disco en otro.

Usuario sólo ve una unidad pero físicamente esta unidad está siendo almacenada de forma idéntica en dos o más discos.

Los bloques de datos se escriben en los discos a la vez.

Si se produce un fallo en un disco podría seguir funcionando sobre un solo disco, mientras se sustituye el disco dañado.

Se reduce a la mitad el espacio disponible (si disponemos de dos discos de 1TB y montamos un RAID-1 esta unidad tendrá un espacio de 1TB)



ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

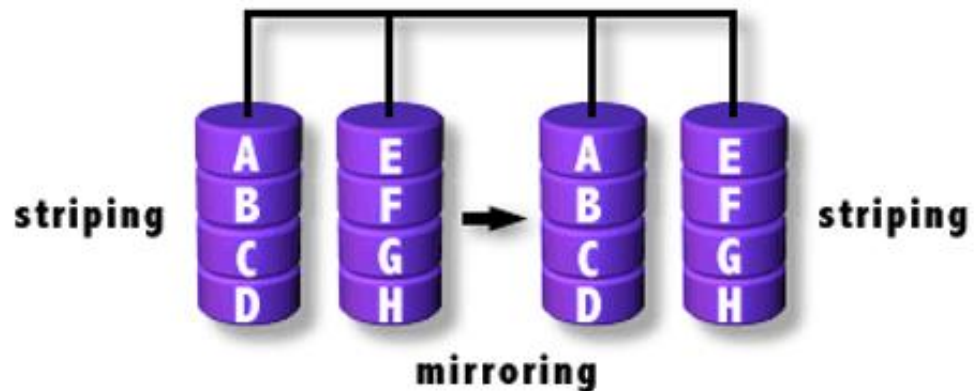
RAID 0+1/ RAID 0/1 o RAID 10

Combinación de los arrays anteriores que proporciona velocidad y tolerancia al fallo simultáneamente. RAID 0+1 fracciona los datos para mejorar el rendimiento, pero también utiliza un conjunto de discos duplicados para conseguir redundancia de datos.

Requiere un mínimo de cuatro unidades y sólo dos de ellas se utilizan para el almacenamiento de datos. Las unidades se deben añadir en pares cuando se aumenta la capacidad, lo que multiplica por dos los costes de almacenamiento.

Este nivel de RAID es el más rápido, el más seguro, pero por contra el más costoso de implementar.

RAID 0+1



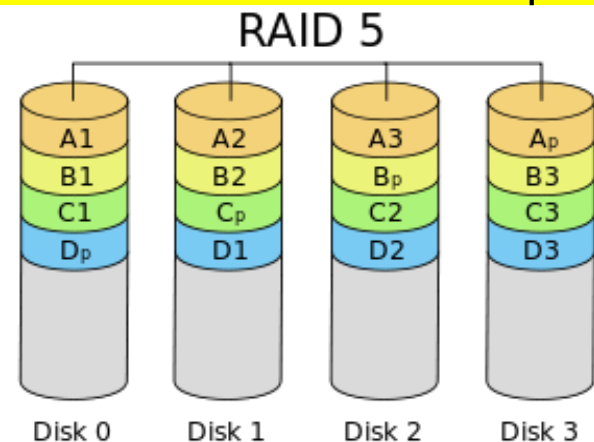
ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-5

Consiste en calcular para cada dato su **paridad** de manera que se pueda recuperar el dato en caso de perder cualquier disco. Una vez obtenida la paridad se hace striping para repartir el dato y su paridad por los disco conectados en RAID.

Gracias al striping se consigue mejor rendimiento que el disco individual y gracias a la paridad estamos más seguros que en RAID-0. A cambio sacrificamos la capacidad de un disco.

Los bloques de datos que se almacenan en la unidad, la información redundante de dichos bloques se distribuyen cíclicamente entre todos los discos que forman el volumen RAID-5.



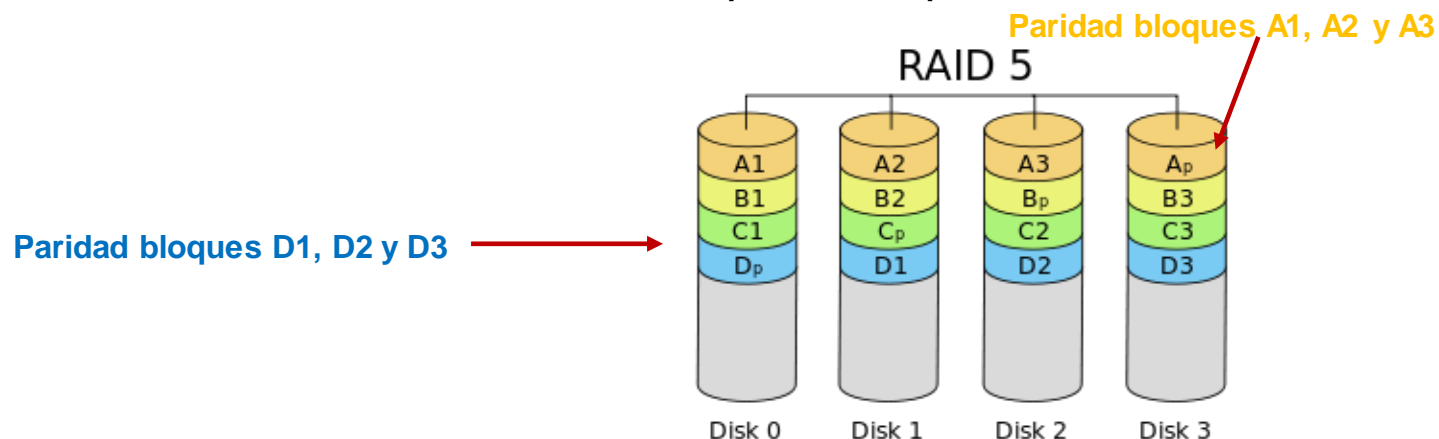
ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-5

Aplicamos RAID-5 sobre 4 discos los bloques de datos se colocan sobre 3 discos, dejando un hueco en cada línea que ira rotando de forma cíclica. En el hueco se colocara un bloque de paridad que cada vez se colocara en un disco.

Cálculo paridad:

- **Paridad par** (paridad 0): si número total de unos es impar el bit paridad es 1, cuando el numero total de unos es par el bit paridad es 0.
 - **Hacer ejercicio de paridad.**
- **Paridad impar** (paridad 1): si número total de unos es par el bit paridad es 1, cuando el numero total de unos es impar el bit paridad es 0.



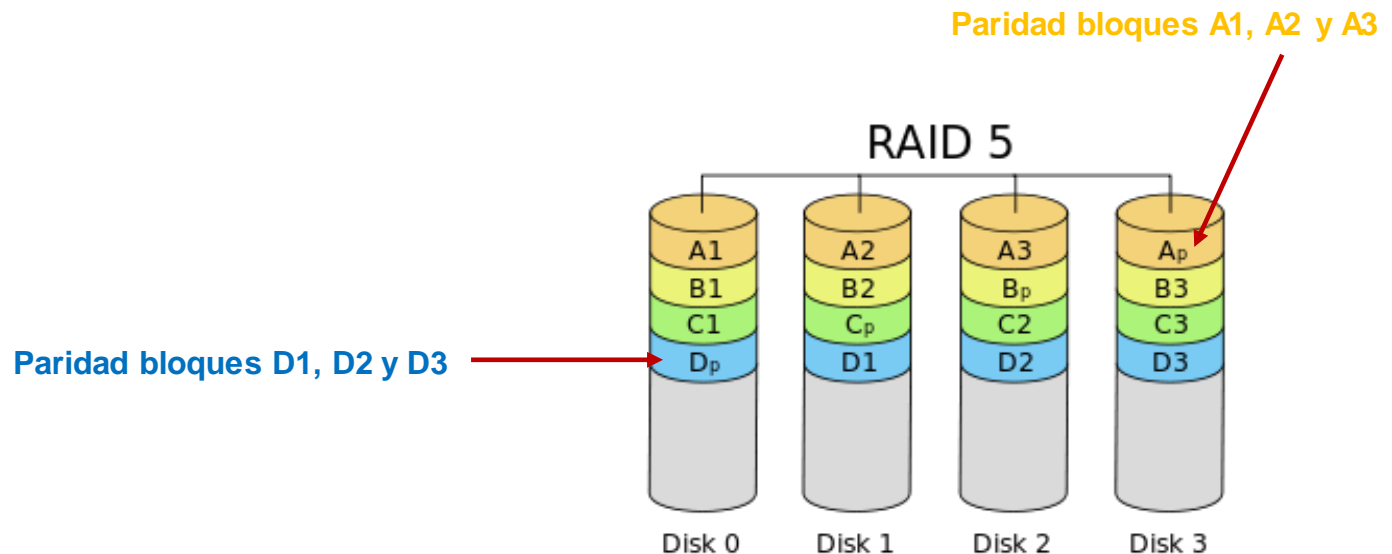
ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-5

Se “gasta” en paridades el tamaño equivalente a uno de los discos.

Si se crea un conjunto de RAID 5 con cuatro discos de 500GB, el almacenamiento se comporta como un único disco de 1500GB.

RAID 5 puede utilizarse con cualquier número de discos a partir de 3, pero con menos de 4 se gasta bastante espacio en paridades, y con más de 4 se va aumentando el riesgo de que dos discos fallen simultáneamente.

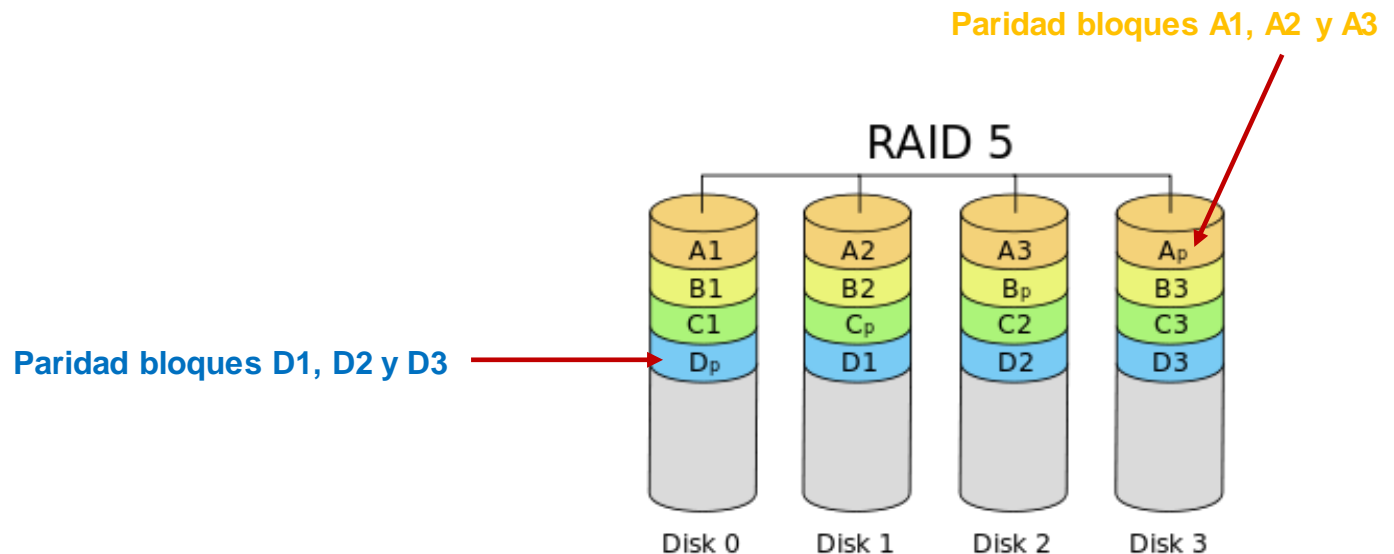


ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID-5

Ofrece tolerancia a fallos, pero además, optimiza la capacidad del sistema permitiendo una utilización de hasta el 80% de la capacidad del conjunto de discos.

Esto lo consigue mediante el almacenamiento alternativo por bloques en todos los discos del conjunto, y el cálculo de información de paridad



ALMACENAMIENTO REDUNDANTE Y DISTRIBUIDO

RAID

Con RAID los datos siguen siendo vulnerables a una amplia variedad de riesgos aparte del fallo físico de disco:

- RAID no impedirá que un virus destruya los datos,
- RAID no impedirá que los datos se corrompan o que sufran la modificación o borrado accidental por parte del usuario
- RAID no impedirá que un fallo físico en otro componente del sistema afecten a los datos.
- RAID no facilita el traslado de los discos a un sistema nuevo. Cuando se usa un solo disco, es relativamente fácil trasladar el disco a un sistema nuevo: basta con conectarlo, si cuenta con la misma interfaz. Con los discos de un conjunto RAID no es tan sencillo.

Cluster de Servidores

Conjunto de servidores que se instalan para trabajar en red.

Los servidores se unen con una red de alta velocidad, funcionan como un único equipo, no tienen por qué tener el mismo sistema operativo y no es necesario que tengan el mismo hardware.

¿Qué servicios ofrecen?

- Alta disponibilidad.
- Alto rendimiento.
- Balanceo de cargas.
- Escalabilidad.

Cluster de Servidores

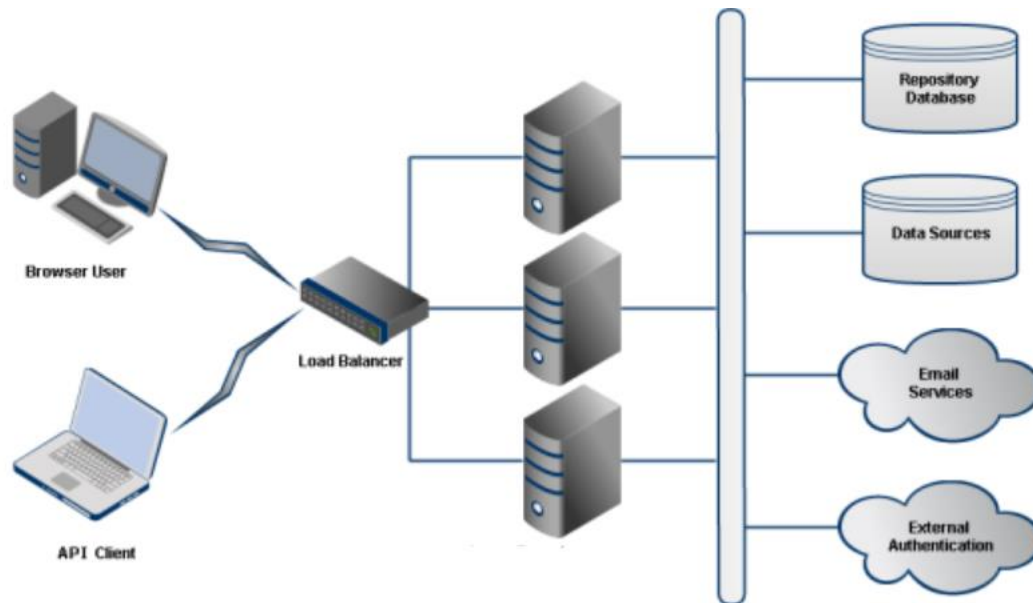
Clasificación Clusters:

- **Cluster alto rendimiento:** tareas que requieren gran capacidad de cálculo y gran cantidad de memoria.
- **Cluster de alta disponibilidad:**
 - ✓ Dotar de disponibilidad y confidencialidad a los servicios que ofrece.
 - ✓ Utilizar hardware duplicado para no tener un único punto de fallo.
 - ✓ Si se produce una avería siempre tendrá que tener otro que lo reemplace.
 - ✓ Tiene que garantizar la disponibilidad de los datos.
- **Cluster de alta eficiencia:**
 - ✓ Ejecutar el mayor número de tareas en el menor tiempo posible.

Cluster de Servidores

Componentes Cluster:

- **Nodos:** es cualquier maquina servidor u ordenador.
- **Sistema operativo:** SO multiproceso y multiusuario.
- **Conexión de red:** nodos deben estar conectados a red alta velocidad
- **Middleware:** es un software que esta entre el sistema operativo y las aplicaciones. Con este tipo de software aparenta que estamos ante un súper ordenador y permite optimizar el uso del sistema.
- **Sistemas de almacenamiento:** son discos duros de los propios equipos y esquemas de almacenamiento NAS y SAN

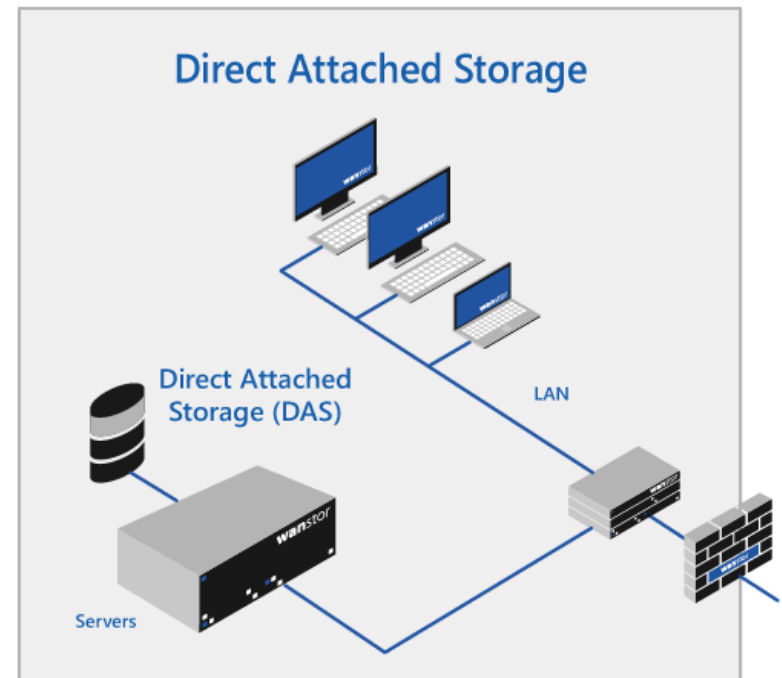


Arquitecturas de Almacenamiento

DAS - Direct Attached Storage

Método tradicional de almacenamiento y el más sencillo. Consiste en conectar el dispositivo de almacenamiento directamente al servidor o estación de trabajo, es decir, físicamente conectado al dispositivo que hace uso de él.

Un sistema DAS habilita capacidad extra de almacenamiento a un servidor, mientras mantiene alto ancho de banda y tasas de acceso.



Arquitecturas de Almacenamiento

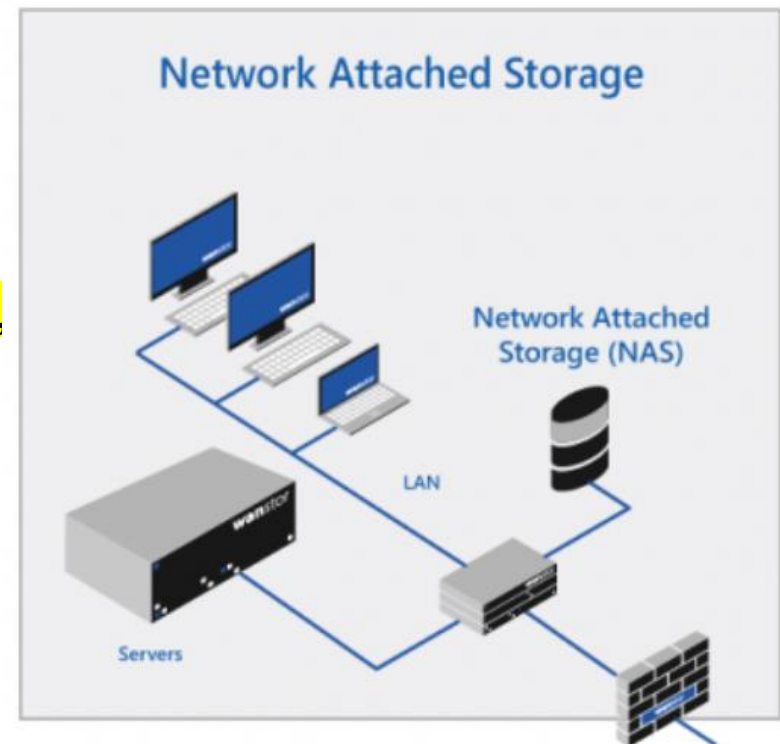
NAS - Network Attached Storage

Almacenamiento conectado en red.

Las aplicaciones hacen las peticiones de datos a los sistemas de ficheros de manera remota mediante protocolos de red TCP/IP. El usuario solicita al servidor un fichero completo y cuando lo recibe, lo maneja localmente.

El uso de NAS permite, con un bajo coste, realizar balanceo de carga y tolerancia a fallos. Es una buena solución para una LAN de tamaño pequeño o medio.

Los sistemas NAS suelen estar compuestos por uno o más dispositivos que se disponen en RAID, para aumentar capacidad, eficiencia y tolerancia a fallos.



Arquitecturas de Almacenamiento

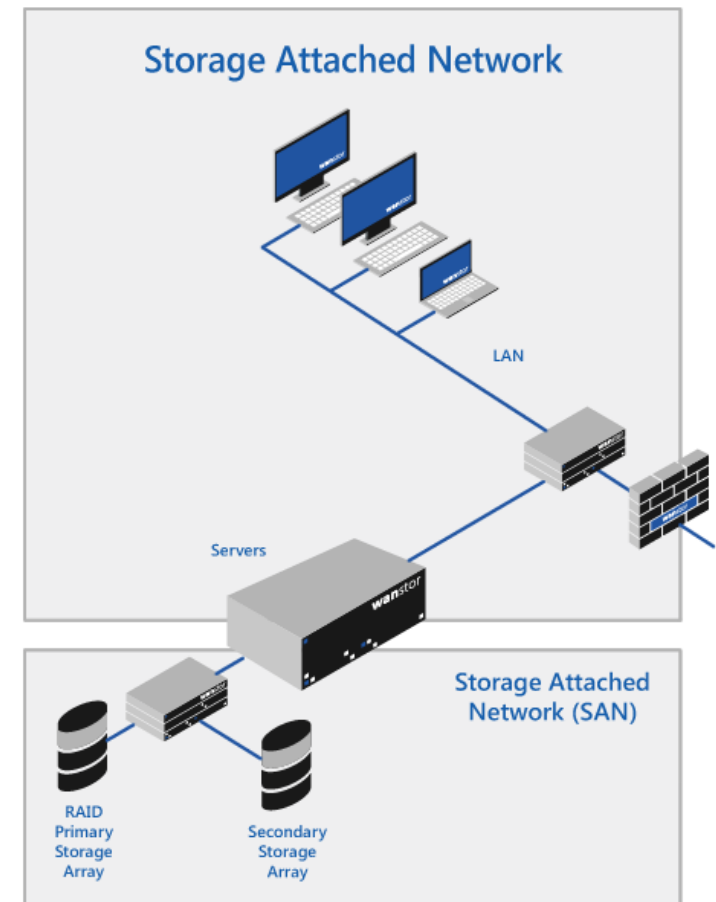
SAN – Storage Area Network

Dispositivos de almacenamiento se encuentran conectados a una red de alta velocidad directamente y resuelven las peticiones que se le realizan.

La infraestructura necesaria y el coste la hace viable solo en grandes organizaciones.

El funcionamiento esencial es que se hacen peticiones de datos al servidor y se encarga de obtener esos datos en el disco.

Se utiliza para grandes volúmenes de datos.



Arquitecturas de Almacenamiento

Tanto en DAS como en SAN las aplicaciones hacen peticiones de datos al sistema de ficheros directamente. La diferencia reside en la manera en la que dicho sistema de ficheros obtiene los datos requeridos del almacenamiento.

En una DAS, el almacenamiento es local al sistema de ficheros, mientras que en una SAN, el almacenamiento es remoto.

En la tecnología NAS, las aplicaciones hacen las peticiones de datos a los sistemas de ficheros de manera remota.

