

RAID

Matriz redundante de discos independientes, del inglés, Redundant Array of Independent Disks.

Es un conjunto de unidades físicas de discos duros que por el sistema operativo actúan como una sola unidad lógica, es importante considerarlo como parte de una estrategia de backup.

Tipos de RAID

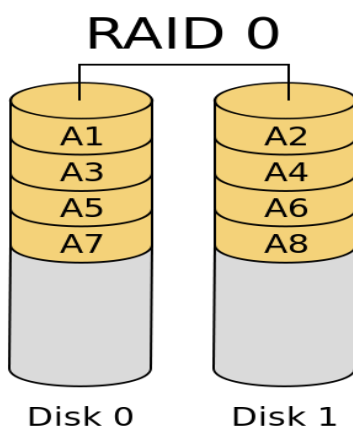
Examinaremos los distintos tipos de Raid que son ampliamente reconocidos. Cada uno presenta distintos niveles de desempeño, capacidad de almacenamiento y confiabilidad. La elección que hagamos estará determinada por el tipo de datos que deseemos resguardar, así como por la cantidad de discos duros disponibles y nuestra inversión disponible.



Raid 0

Para operar, se requieren al menos dos discos y se suma el tamaño total de todos los HDD. El **RAID 0**, también conocido como *striping*, emplea un mínimo de dos discos y distribuye los datos entre ellos.

- Proporciona un **mayor rendimiento**.
- No se aconseja para datos críticos.



El inconveniente radica en la falta de **redundancia** y **tolerancia** a fallos, lo que significa que cualquier fallo o avería en uno de los discos resultará en una pérdida completa de los datos.

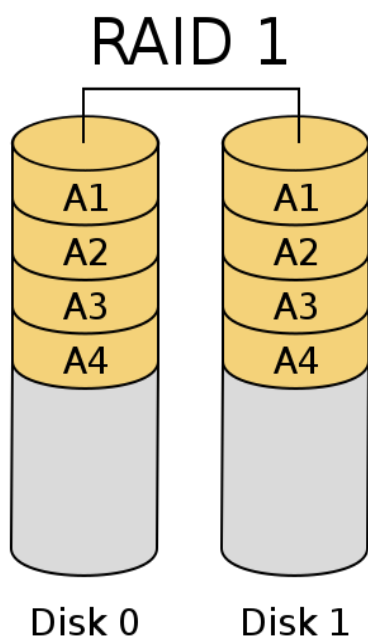
Se recomienda este tipo de configuración si priorizamos el rendimiento del sistema y el acceso rápido a la información (como en diseño gráfico, modelado en 3D y edición de video), y si contamos con un presupuesto limitado. Ofrece un rendimiento elevado, especialmente con archivos de gran tamaño.



Raid1

Para configurar RAID 1, se requieren al menos dos discos. Si hay más unidades disponibles, solo se considerará el disco de menor tamaño.

Conocido como "**espejo**" o "**mirroring**", el **RAID 1** emplea dos discos y replica todos los datos de la primera unidad de manera sincronizada en una segunda unidad de almacenamiento. De esta manera, si uno de los discos falla, el sistema seguirá operando sin problemas utilizando el segundo disco y sin pérdida de datos.



Proporciona una rápida recuperación después de un fallo de unidad, lo que lo convierte en una de las configuraciones más confiables en cuanto a **redundancia** y **tolerancia** a fallos.

Además, ofrece un mayor rendimiento de lectura/escritura, ya que permite la lectura y escritura simultáneas en ambos discos.

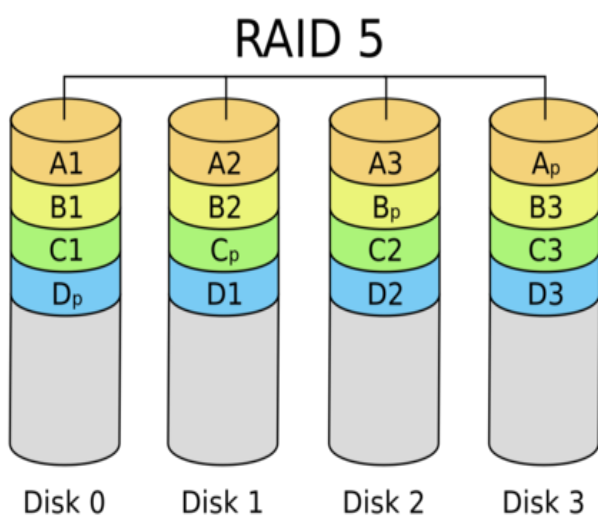
La desventaja es que se necesitan dos discos para obtener la misma capacidad de almacenamiento efectivo (por ejemplo, 2x2TB, que equivale a 2 TB en total), y puede haber una carga adicional en el disco.

Se recomienda para servidores de nivel básico que solo dispongan de dos discos duros. Es ideal para bases de datos pequeñas u otras aplicaciones con requisitos de capacidad modestos pero necesidad de una completa redundancia de datos.



Raid 5

El RAID 5 requiere un mínimo de tres discos (permitiendo la pérdida de un disco sin perder datos). El espacio disponible en un RAID 5 será de $n-1$, donde n es el número total de discos en el RAID. Por ejemplo, si utilizamos cinco discos de 1TB cada uno, tendremos $5 \text{ discos} - 1 = 4 \text{ discos}$, lo que equivale a 4TB disponibles.



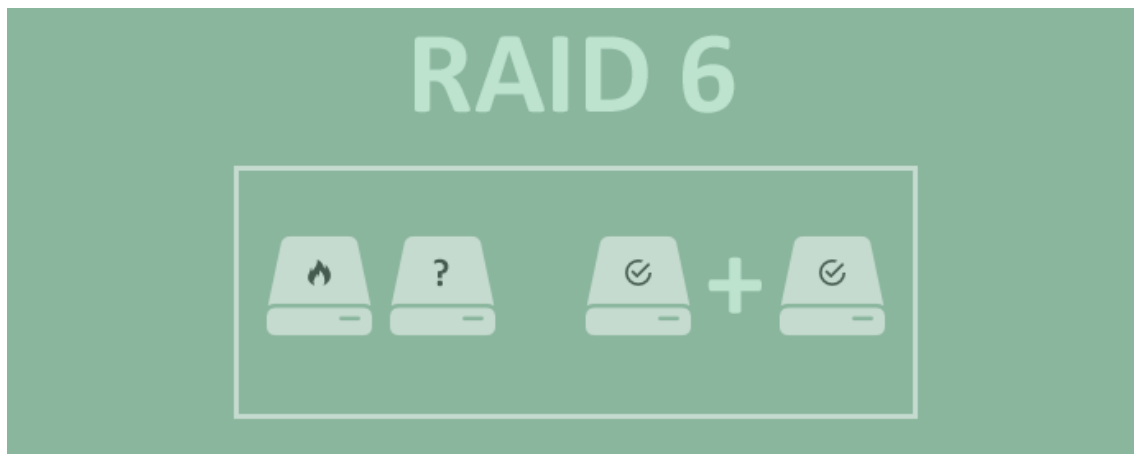
Es comúnmente utilizado en servidores debido a que combina la velocidad y rendimiento del RAID 0 (alta eficiencia en el uso del espacio, alto rendimiento en lectura y escritura) con la seguridad del RAID 1 frente a la pérdida de datos.

El RAID 5 utiliza la **paridad** para recuperar los datos en caso de fallo de un disco. Los datos se distribuyen en bloques en los diferentes discos, de modo que si uno falla, esa parte de los datos se reconstruye utilizando la información almacenada en los demás discos, permitiendo al usuario continuar con su trabajo, aunque a una velocidad reducida.

¿Cuáles son sus desventajas?

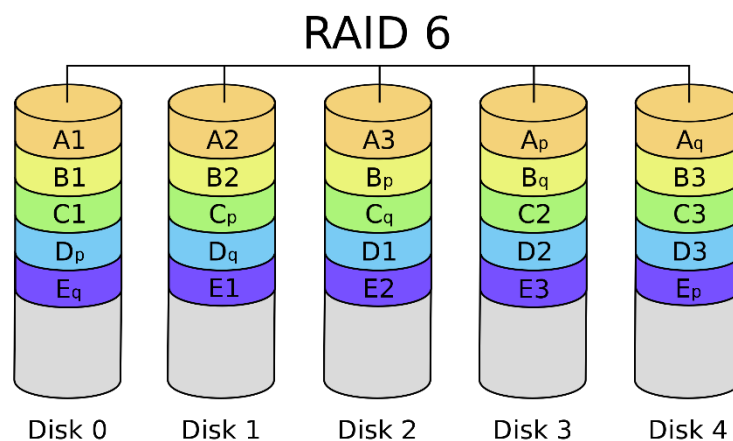
- Existe un impacto medio de los fallos del disco, y la reconstrucción lleva más tiempo debido a la necesidad de recalcular la paridad.
- El rendimiento tanto en lectura como en escritura se ve significativamente afectado si la matriz RAID 5 se degrada.
- Para discos SATA de alta capacidad (entre 500GB y 1TB), los tiempos de reconstrucción son prolongados y provocan una degradación en el rendimiento del controlador.

Importante: Si ocurre un fallo en un segundo disco mientras se está reconstruyendo el primero, **se perderán todos los datos.**



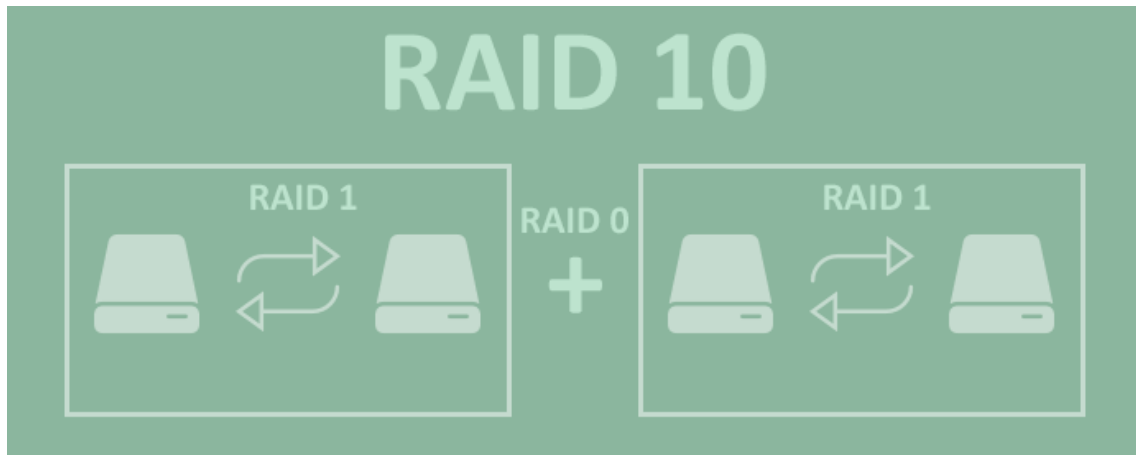
Raid6

Se necesitan al menos *cuatro* discos para implementar RAID 6, que puede tolerar el fallo de hasta dos discos duros ($N-2$). Similar a RAID 5, RAID 6 incorpora un disco de reserva que entra en acción cuando uno de los discos falla (transformándolo efectivamente en un RAID 5 hasta que se reemplace el disco averiado).



Ofrece una alta redundancia de datos y un rendimiento sólido en lectura. Sin embargo, el rendimiento en tareas de escritura es inferior al de RAID 5 debido a los dos cálculos de paridad. Requiere una inversión adicional, **ya que dos discos se reservan exclusivamente para la paridad**. Por ejemplo, si disponemos de cuatro discos de 1TB, solo tendremos 2TB de espacio utilizable, ya que los otros 2TB se destinan a la paridad.

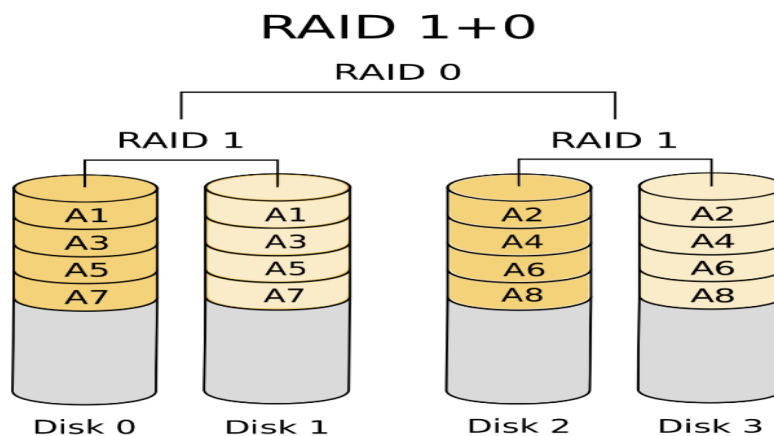
RAID 6 es una opción recomendada cuando se necesita tolerancia a fallos de varias unidades y se prioriza una redundancia más robusta para la protección de datos.



Raid10 (Raid 1+0)

Se requieren al menos cuatro discos para implementar RAID 1+0, donde se configura primero un RAID 1 y luego se aplica un RAID 0 sobre él. Esta configuración implica un costo más alto debido al número mínimo de unidades necesarias.

A cambio, se obtiene un alto rendimiento de lectura, gracias al RAID 0, y una tolerancia a fallos, gracias al RAID 1. Por ejemplo, si se utilizan cuatro discos, hasta dos pueden fallar sin perder información (N-2), siempre y cuando los discos defectuosos no pertenezcan al mismo subgrupo.



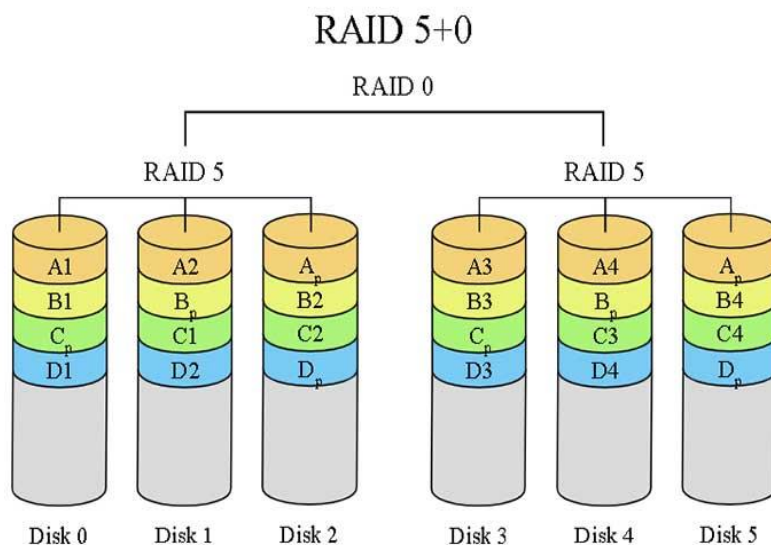
A pesar de su precio más elevado, el RAID 10 resulta ideal para aplicaciones como servidores de bases de datos. Ofrece un rendimiento elevado, una redundancia de datos completa y una rápida recuperación frente a fallos en los discos.



Raid 50 (Raid 5+0)

Se requieren al menos seis discos para implementar RAID 5+0, donde se configura primero un RAID 5 y luego se aplica un RAID 0 sobre él. Esta configuración permite la tolerancia a fallos de hasta tres discos sin perder datos, según la fórmula $s(n-1)$.*

Con RAID 50, obtenemos un volumen muy robusto, un mayor rendimiento de lectura en comparación con RAID 5 estándar y un rendimiento de escritura que va de medio a alto.



Sin embargo, hereda las mismas desventajas que RAID 5, como el impacto medio ante fallos de disco y tiempos de reconstrucción más largos debido a la necesidad de recalcular la paridad. Además, tiene un costo más elevado.

RAID 60

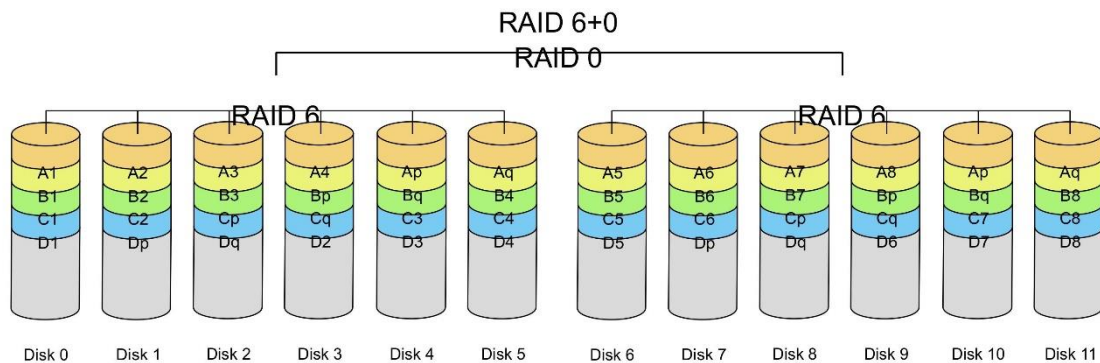


Raid 60 (Raid 6+0)

Para implementar RAID 6+0 *se requieren al menos ocho discos*, lo que permite la tolerancia a fallos de hasta cuatro discos sin perder datos.

Este tipo de configuración combina RAID 6 con RAID 0. Proporciona un alto rendimiento, especialmente en operaciones de lectura.

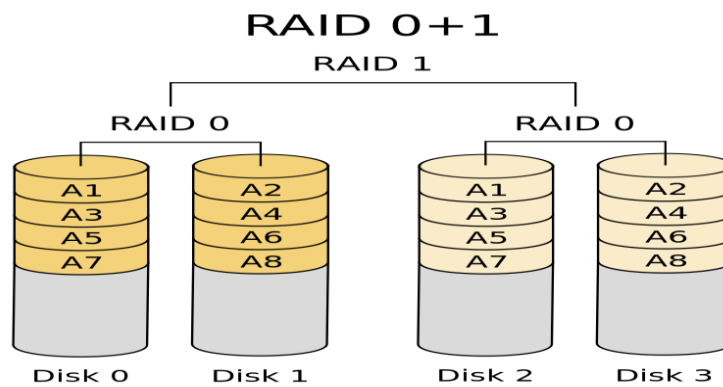
Sin embargo, hereda las mismas desventajas que RAID 6, como un rendimiento más bajo en operaciones de escritura debido a los dos cálculos de paridad, y requiere una inversión adicional en hardware.



Raid 0+1

Se precisarán 4 discos duros para esta configuración.

Los discos se emparejan para formar dos conjuntos RAID 0, y luego estos dos bloques se combinan en un RAID 1. A diferencia de la RAID 10, esta **configuración no tolera dos fallos simultáneos**, por lo que es menos segura.



Tipos de RAID más usados en servidores.

Entre los tipos de RAID más comunes en servidores se encuentran RAID1, RAID5, RAID6 y RAID10.

Algunos términos importantes que conviene conocer son éstos:

- **Datos de paridad:** Distribuidos entre todos los discos físicos en el sistema, permiten la reconstrucción de un disco físico fallido a partir de la paridad y los datos de los discos restantes. Este concepto se aplica en RAID 5, 6, 50 y 60.
- **Modo Degradado:** Se activa cuando uno de los discos en el RAID se vuelve ilegible. En este caso, el disco afectado se considera corrupto y se elimina del RAID. Los datos y la paridad del disco se redistribuyen entre los discos restantes. Sin embargo, este proceso degrada significativamente el rendimiento del RAID, y se conoce como "Modo Degradado".