

# Introducción a sistemas RAID

*Cuando hablamos de velocidad y seguridad en los datos no podemos dejar de explicar el concepto de sistemas RAID, ya que estos representan una mayor eficacia y seguridad para tratar grandes volúmenes de información*

---

## **INTRODUCCION:**

Gracias a la universalización de la informática, más personas están en contacto con el mundo de los ordenadores, ya sea para uso personal o profesional. Todo ha servido para una evolución más rápida de los diferentes componentes de los ordenadores. Desde hace 20 años la capacidad de proceso de los procesadores ha aumentado un 40 % cada año. En el mismo periodo de tiempo, los discos han doblado su capacidad cada año, mientras que su costo se ha reducido a la mitad. Desgraciadamente, el aumento del rendimiento de los discos duros ha sido menos importante en comparación con el rendimiento del sistema, ya que tan solo ha mejorado un 50 % durante la última década. Por lo que teníamos y tenemos un problema con los discos duros, que son menos eficaces que el rendimiento general del sistema, provocando una descompensación entre el tratamiento de la información del sistema (muy rápido) y la lectura - grabación de datos en el disco duro (muy lenta). Para ello se inventó un sistema para guardar información en varios discos duros a la vez por lo que el acceso se hacía más rápido ya que la carga se distribuía entre los diferentes discos duros, a esto se le llama cadenas redundantes de discos de bajo costo (RAID). Pero a la hora de saber como podemos optimizar nuestro RAID nos encontramos con algunos problemas, como puede ser encontrar la perfecta relación calidad precio.

En este documento se intentará explicar que es RAID que tipos existen y que modelo es el que mejor se ajusta a nuestras necesidades.

## **¿Que es RAID?**

RAID es un conjunto de dos o más discos que funcionan de forma conjunta, para poder aumentar el rendimiento y el nivel de protección de los datos.

Una de las ideas originales de los discos RAID, era la reducción del costo. Inicialmente se pensó que si se cambiaban los discos de gran capacidad por un conjunto de discos de menor capacidad y menor precio se reduciría el costo del mantenimiento. Desgraciadamente no es así, ya que, un disco RAID sería más caro que un disco duro convencional de la misma capacidad.

No obstante, los discos RAID proporcionan una serie de ventajas respecto a los discos duros convencionales, ya que pueden ser más útiles para nuestros sistemas.

### **Ventajas de los discos RAID:**

El rendimiento general del sistema aumenta ya que pueden funcionar de forma paralela con los diferentes discos del conjunto.

Dependiendo del nivel de RAID que escojamos, si uno de los discos del conjunto falla, la unidad continúa funcionando, sin pérdida de tiempo ni de datos. La reconstrucción de los datos del disco que ha fallado se hace de forma automática sin intervención humana. En el caso de algunos sistemas operativos la regeneración de datos se hace desde software por ejemplo en el windows NT, aunque en estos sistemas se pueden usar controladoras RAID que sí regenerarían los datos automáticamente.

La capacidad global del disco aumentará, ya que se suman las capacidades de los diferentes discos que componen el conjunto.

## Los diferentes Niveles Raid.

Para empezar a acercarnos a los sistemas RAID debemos explicar qué tipos existen y en qué se caracteriza cada uno de ellos. En esta página encontraremos todo esto.

La primera definición de RAID, fue en 1987, pueden definirse inicialmente 5 niveles RAID, aunque posteriormente se están haciendo 2 más. Los niveles RAID ofrecen grandes diferencias entre rendimiento e integridad de los datos, dependiendo de las especificaciones de cada nivel. No hay un nivel RAID perfecto para todos los usuarios, ya que cada uno de ellos cumplen distintos propósitos.

A continuación, intentaremos introducirnos en el funcionamiento de cada uno de los niveles RAID, y mencionaremos algunas de las ventajas e inconvenientes que proporcionan.

[RAID 0: Data Striping Without Parity \(DSA\)](#)

[RAID 1: Mirrored Disk Array \(MDA\)](#)

[RAID 2: Hamming Code for Error Correction](#)

[RAID 3: Parallel Disk Array \(PDA\)](#)

[RAID 4: Independent Disk Array \(IDA\)](#)

[RAID 5: Independent Disk Array \(IDA\)](#)

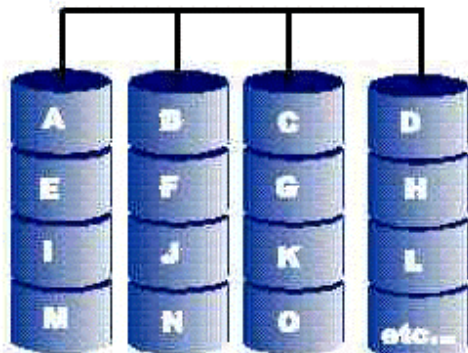
[RAID 6: Es el raid 5 pero con 2 paridades.](#)

### RAID 0: Data Striping Without Parity (DSA)

*Datos en bandas de discos sin paridad y sin corrección de errores*

El nivel RAID 0, no va ser definido inicialmente, pero es un termino comúnmente empleado.

Los datos repartidos entre los diferentes discos es información redundante. Este nivel será relacionado en muchas ocasiones como striping.



#### DISCOS DE DATOS

Ventajas RAID 0

- ?? Proporciona un alto rendimiento.
- ?? No tiene costo adicional.
- ?? Toda la capacidad del disco se emplea.

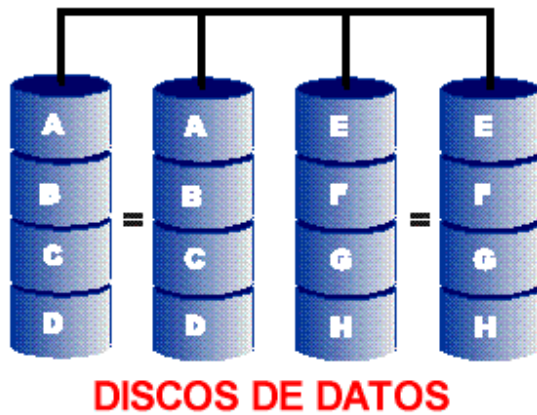
Inconvenientes RAID 0

- ?? No es verdaderamente un disco RAID ya que no tiene integridad de datos.
- ?? Un error en uno de los discos implica la pérdida total de los datos.

### RAID 1: Mirrored Disk Array (MDA)

*Conjunto de discos en espejo*

El nivel RAID 1 es más conocido como mirroring ya que los datos son escritos al mismo tiempo en dos discos diferentes. Tienen dos copias exactas del total de la información. El RAID 1 es una solución cara, ya que desaprovecha la mitad de la capacidad total del conjunto de discos.



**Ventajas RAID 1**

- ?? Mayor rendimiento en las lecturas de datos respecto a un disco convencional.
- ?? Podemos recuperar todos los datos, en caso de error en uno de los discos.

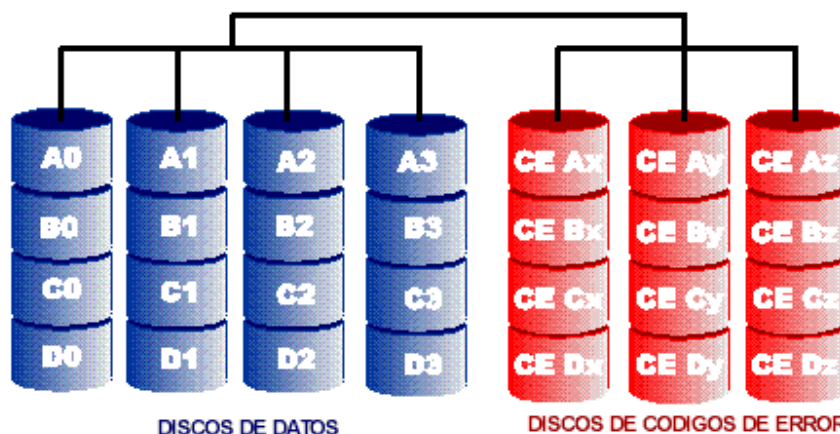
**Inconvenientes RAID 1**

- ?? Bastante caro, ya que necesitamos el doble de espacio del necesario.
- ?? Moderada lentitud en la escritura de datos, ya que los hemos de escribir en dos localizaciones distintas

## RAID 2: Hamming Code for Error Correction

*Es el primer nivel RAID que usa código de correcciones de error utilizando la "generación Hamming" de códigos de error.*

El nivel RAID 2 emplea múltiples discos, como en el nivel RAID 0, pero algunos de estos discos son empleados para códigos de error, los cuales los emplean para referencia de los datos en caso de que falle uno de los discos. Este nivel tiene un costo bastante elevado ya que necesitamos mucho disco para mantener los códigos de error.



Gracias a como están distribuidos los datos en los discos se consigue mejorar la velocidad de transferencia principalmente en la lectura, ya que podemos emplear todos los discos en paralelo. Estos discos, aunque proporcionen un buen rendimiento, no son muy empleados, ya que los niveles 1,3 y 5 proporcionen una mayor relación costo/rendimiento.

#### Ventajas RAID 2

- ?? Se emplea para mejorar la velocidad de demanda y también la velocidad de transferencia.
- ?? Podemos recuperar datos gracias a los discos de códigos de error.

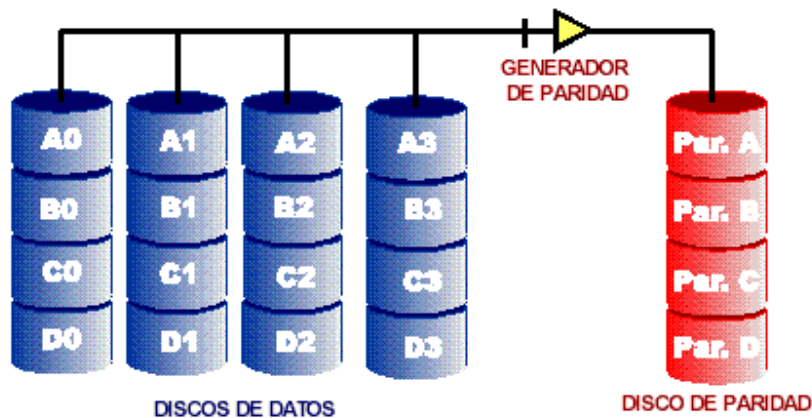
#### Inconvenientes RAID 2

- ?? Solución cara, ya que requeriremos mucho disco para guardar los códigos de error.
- ?? Tiempo de escritura de datos bastante lento, incluso aunque los datos se separen en los diferentes discos

## RAID 3: Parallel Disk Array (PDA)

*Sistema de discos en paralelo con disco de paridad para la corrección de errores*

RAID 3 emplea múltiples discos para hacer el striping, como en el nivel RAID 2, pero sólo hace falta un disco nada más para mantener la paridad por lo que reducimos el costo en discos.



Este nivel RAID es una buena alternativa para aplicaciones de velocidad de transferencia alta, ya que gracias a la distribución de datos podemos emplear todos los discos en paralelo.

#### Ventajas RAID 3

- ?? Alto rendimiento para aplicaciones de velocidad de transferencia alta.
- ?? Gracias al disco de paridad, podemos recuperar datos.

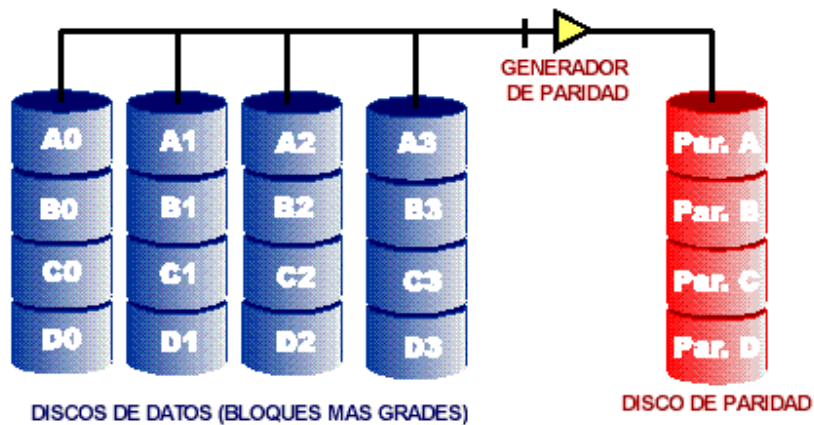
#### Inconvenientes RAID 3

- ?? Si perdemos el disco de paridad, perdemos toda la información redundante que teníamos.
- ?? Tiempo de escritura de datos bastante lento.

## RAID 4: Independent Disk Array (IDA)

*Sistema de discos independientes con disco de control de errores*

El nivel RAID 4 es más parecido al RAID 3. Los bloques de datos que distribuimos en los diferentes discos son más grandes por lo que se consigue un rendimiento superior en las escrituras.



#### Ventajas RAID 4

- ?? Buen rendimiento en los escrituras de datos.
- ?? Tiene integridad de datos.

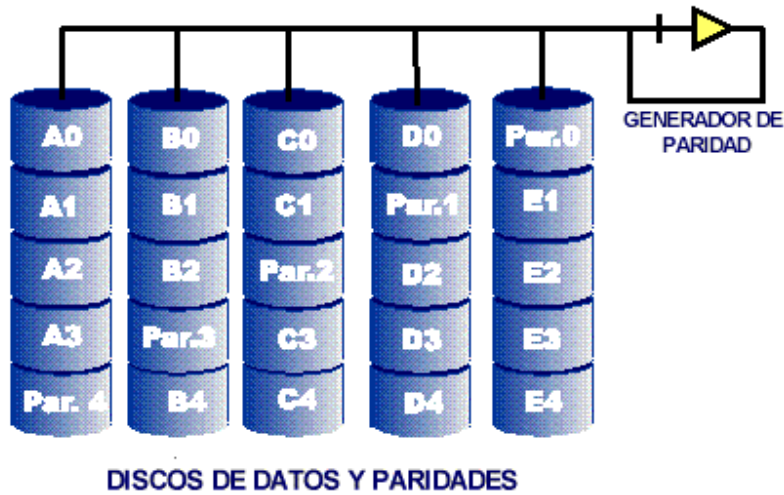
#### Inconvenientes RAID 4

- ?? Si perdemos el disco de paridad, perdemos toda la información redundante que teníamos.
- ?? Menor rendimiento en las lecturas de datos.

## RAID 5: Independent Disk Array (IDA)

*Sistema de discos independiente con integracion de codigos de error mediante una paridad.*

En RAID 5 los datos y la paridad son guardados en los mismos discos, por lo que conseguimos aumentar la velocidad de demanda, ya que cada disco puede satisfacer una demanda independientemente de los demás. Con diferencia con el RAID 3, el RAID 5 guarda la paridad del dato dentro de los discos y no hace falta un disco para guardar dichas paridades. La paridad se genera haciendo un XOR de los datos A0,B0,C0,D0 creando la zona de paridad PAR0, como se ve la paridad nunca se guarda en los disco que contienen los datos que han generado dicha paridad, ya que en el caso que uno de ello se estropeará como por ejemplo el dato A0 vastaria con regenerar las banda B0,C0,D0,PAR0 para que el dato volviera ha reestablecerse.



**Ventajas RAID 5**

- ?? Alto rendimiento en aplicaciones de velocidad de demanda interactivas.
- ?? Costo efectivo. No desaprovecha un disco exclusivamente para paridad.
- ?? Se pueden recuperar datos.

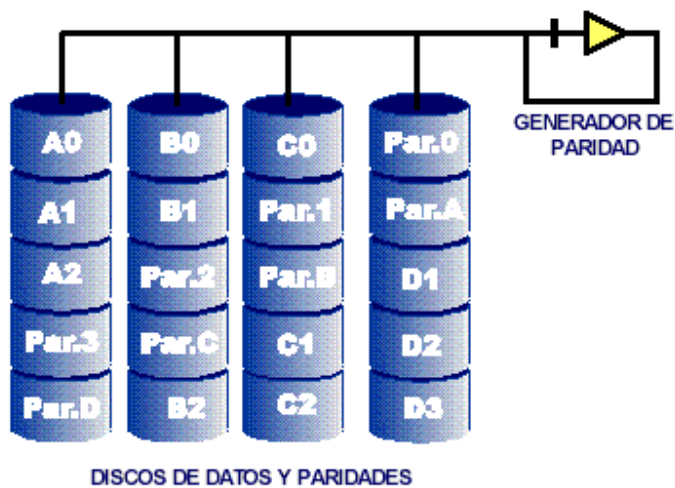
**Inconvenientes RAID 5**

- ?? El rendimiento en las escrituras de datos es bajo.
- ?? No aumenta el rendimiento en las aplicaciones, aunque la velocidad de transferencia de datos es alta.

## RAID 6: Independent Disk Array (IDA)

Sistema independiente de disco con integración de códigos de error mediante una doble paridad.

RAID 6 es esencialmente una extensión del nivel RAID 5, para ello guarda una segunda paridad. Este nivel proporciona muy buena integridad de los datos y repara diversos errores en los discos.



**Ventajas RAID 6**

- ?? Podemos recuperar diversos errores simultáneamente.
- ?? Nivel de integridad muy elevado. Solución perfecta para aplicaciones críticas.  
Inconvenientes RAID 6
- ?? El rendimiento en escrituras de datos es bastante lento.
- ?? No se dispone de muchas implementaciones comerciales del nivel RAID 6.

## Como seleccionar un disco RAID.

En esta página veremos lo que hay que tener en cuenta para seleccionar el nivel de RAID que nos interesa, así como los diferentes sistemas que existen en el mercado para hacer un sistema mas seguro y rápido en el caso de que se estropeara algún disco duro.

El seleccionar un disco RAID puede ser una dura tarea, ya que no existen criterios establecidos para saber cual es la mejor solución. Una equivocación bastante común es juntar todos los discos RAID en el mismo nivel, por lo que debemos elegir una solución correcta para el grupo completo de disco, no metiendo todos en el mismo nivel, sino estructurando los discos dependiendo de las necesidades en seguridad e integridad de los datos.

Para seleccionar un disco RAID en primer lugar tenemos que conocer nuestras necesidades o saber los tipos de aplicaciones que se emplean. Podemos separar las aplicaciones en dos tipos:

?? Aplicaciones de demanda:

Podríamos seleccionar el RAID 5 ya que ofrece una alta velocidad de demandas, tanto en escrituras como en lecturas.

?? Aplicaciones de transferencia:

?? La mejor opción pasaría por RAID 3 ya que ofrece una alta velocidad de transferencia para gráficos, imágenes y aplicaciones en general en las que necesitamos gran transferencia de datos.

Además del tipo de aplicación, habremos de tener en cuenta los factores que nos guiarán a optar por la elección de un RAID u otro como son:

?? Integridad y costo:

Normalmente nos decantaremos por una solución RAID que asegure la integridad de los datos del disco, ya que de esta manera tendremos preservados nuestros datos sobre posibles errores del disco. El RAID 0 es el único nivel que no proporciona integridad.

?? Costo:

?? El costo, vendrá marcado por nuestro poder adquisitivo, será uno de los factores más importantes a la hora de seleccionar nuestro disco. No solo tenemos que pensar en la integridad de los datos sino en cuanto que cuando se estropee un disco tenemos que apagar el sistema. Para ello existen varios dispositivos como son:

[Hot swap](#)

[Hot spare](#)

[Duplex controler](#)

[Redundat Power Supplies.](#)

Son más caros pero tenemos que tenerlos en cuenta para que el sistema nunca este apagado.

La siguiente Tabla nos muestra gráficamente las opciones más acertadas a la hora de elegir el Nivel RAID que mejor se adapte a nuestras necesidades teniendo en cuenta todos estos factores:



Costo	Rendimiento	Integridad	Nivel RAID
-	-	X	RAID 1
-	X	-	RAID 0
-	X	X	RAID 1
X	-	-	RAID 0
X	-	X	RAID 3/5
X	X	-	RAID 0
X	X	X	RAID 3/5

## Diccionario del RAID.

En este diccionario se intentará aclarar algunos conceptos referentes a los sistemas RAID y a la integridad del sistema de datos.

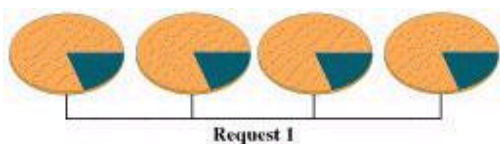
- ?? **Duplex controler:** Controladora doble. Son controladoras de disco que pueden trabajar de modo simple o de forma doble para que la carga de la controladora sea distribuida y más eficaz.
- ?? **Hot swap:** Son discos que se reemplazan en caliente. Esto quiere decir que si alguno de los discos RAID se estropeara físicamente bastaría con quitarlo y poner otro sin tener que apagar el sistema.
- ?? **Hot spare:** Disco sobrante. Es un disco que permanece siempre en el sistema esperando a que uno se estropee y él entre directamente en funcionamiento.
- ?? **Integridad de los datos:** es la capacidad que tiene un disco de aguantar un error de grabación, de corrupción o pérdida de datos. Para tal efecto se tiene que seleccionar un disco RAID o una alternativa. El nivel de integridad es uno de los primeros criterios que se han de investigar.
- ?? **Mirroring:** Haciendo espejo. Es un nivel de RAID (el nivel 1) que pasa por hacer una copia íntegra de un disco en otro.
- ?? **Paridad:** es una información redundante que es guardada para regenerar datos perdidos por un error en el disco. La paridad se genera haciendo un XOR sobre los datos de los discos y guardándolo en otro disco o en un disco dedicado a este efecto, esto dependerá del nivel RAID que usemos.
- ?? **RAID:** Redundat Array of Inexpensive Disk Conjunto de discos con información redundante de bajo costo.
- ?? **Reconstrucción o regeneración:** Cuando un disco falla la información redundante en los discos y los datos en los discos buenos son usados para regenerar la información del disco averiado.
- ?? **Redundat Power Supplies:** Fuente de alimentación redundante. El sistema consta de dos fuentes de alimentación. Si una se estropea, se pone en marcha la otra, pudiéndose cambiar la estropeada en caliente.
- ?? **Sistemas basados en velocidad de demanda:** son aquellos en los que se requiere un gran número de demandas de Entrada/Salida en el menor tiempo posible. Una alta velocidad de demanda es típica en sistemas multiusuario o en aplicaciones de bases de datos en las que se generan un gran número de pequeñas demandas de Entrada/Salida.

En este caso cada disco del conjunto puede procesar una demanda (request) simultáneamente.



- ?? **Sistemas basados en velocidad de transferencia:** son aquellos que intentan satisfacer una sola demanda con la máxima velocidad. Un claro ejemplo son las aplicaciones de ingeniería, científicas, imágenes, etc.

En estos sistemas el tamaño de la demanda es mucho más grande que en los sistemas basados en velocidad de demanda, para el que emplea todos los discos del conjunto en paralelo para satisfacer una mayor demanda en el mínimo tiempo posible.



**Striping:** es el acto de unir dos o más discos físicos en un solo disco lógico con el fin de dividir los datos entre los diferentes discos para ofrecer una significativa mejora en el rendimiento del conjunto de los discos.