

Tarea No.2: RAID de Arquitectura del Computador

Prof. José Canache

Nombres: José Jesús Díaz y María Moreno

1) RAID 0, 1, 5, 10.

El **RAID** (*Redundant Array of Independent Disks* o Matriz de Discos Independientes Redundantes), lo definimos como la combinación o montaje de pequeños discos que forman un dispositivo con mejor rendimiento y redundancia. Gracias a este sistema, la información es distribuida sobre varios discos en función de la técnica que se use, lo que permite mejorar la velocidad, aumentar la capacidad y minimizar la pérdida de datos en caso de un fallo físico en el sistema operativo.

Los tipos de RAID más comúnmente usados son:

RAID 0: es un nivel de redundancia y rendimiento que consiste en dividir los datos en franjas y almacenarlas en varios discos. Esta configuración no proporciona redundancia ni tolerancia a fallos, ya que, la pérdida de un disco implica la pérdida de todos los datos. Sin embargo, el RAID 0 ofrece un mejor rendimiento y velocidad de lectura y escritura en comparación con un disco único.

RAID 1: es un nivel de redundancia que utiliza espejos para duplicar los datos en dos o más discos. Esto garantiza que los datos sean recuperables incluso si uno de los discos falla. El RAID 1, ofrece una alta tolerancia a fallos, pero no mejora el rendimiento de lectura y escritura en comparación con un disco único.

RAID 5: es un nivel de redundancia y rendimiento que utiliza paridad para distribuir los datos y la información de paridad en varios discos. La paridad permite calcular y recuperar los datos perdidos si uno de los discos falla. El RAID 5, ofrece una buena relación entre rendimiento y redundancia, pero puede tener un rendimiento más lento en la escritura debido al cálculo de paridad.

RAID 10 (RAID 1 + 0): es una combinación de RAID 1 y RAID 0, que utiliza espejos (RAID 1) y franjas (RAID 0) para proporcionar redundancia y rendimiento. El RAID 10, es ideal para aplicaciones de alta disponibilidad y misión crítica, como bases de datos y servidores, ya que, ofrece excelente rendimiento y alta tolerancia a fallos. Sin

embargo, el RAID 10 requiere un mayor número de discos y puede ser más costoso en comparación con otros niveles de RAID.

¿Cuál es el mejor RAID 0, 1, 5 o 10?, según la IA.

De acuerdo a la IA, no hay un RAID que sea el mejor en todos los casos, ya que, cada uno tiene sus propias ventajas y desventajas en términos de rendimiento, tolerancia a fallos y eficiencia de almacenamiento. Como se mencionó anteriormente, RAID 10 es una excelente opción para aplicaciones críticas que requieren un alto rendimiento y tolerancia a fallos; mientras que, RAID 5 es más adecuado para aplicaciones que requieren una combinación de rendimiento y protección de datos. RAID 0, es adecuado para aplicaciones que requieren un alto rendimiento, pero, no son críticas en términos de pérdida de datos; mientras que, RAID 1 es adecuado para aplicaciones que requieren protección de datos y no necesitan un alto rendimiento.

Opinión personal con respecto a, ¿Cuál es el mejor RAID?

Siempre se debe buscar ciertas condiciones, como: espacio, confiabilidad, fiabilidad y disponibilidad. Para nosotros el RAID 0, es la mejor opción, ya que, ofrece un mejor rendimiento y velocidad de lectura y escritura en comparación con un disco único, por lo que, se puede obtener un mayor rendimiento del almacenamiento, cuando el presupuesto es muy limitado y cuando una posible pérdida de los datos no supone mayor problema. Aun así, es importante recalcar que la elección del RAID adecuado va a depender de las necesidades específicas que se tenga en el momento y de los recursos con los que se cuenten.

2) RAID HW/RAID SW (RAID Hardware y RAID Software).

Un sistema RAID puede ser controlado por hardware o por software, y aquí es donde también existe diferencia tanto de funcionamiento como de rendimiento, por lo que existen dos formas de ejecutar RAID:

- 1) **RAID HW (Hardware RAID):** depende de una controladora RAID especial y potente, que se instala en la torre del ordenador o en una matriz de discos. La controladora RAID es responsable de la organización y gestión de los dispositivos de almacenamiento individuales, lo que ofrece un excelente rendimiento y una elevada velocidad de transferencia de datos. Sin embargo, el RAID hardware es costoso y menos flexible en términos de compatibilidad y traslado a un nuevo sistema.

2) RAID SW (Software RAID): es una implementación de la tecnología RAID que se apoya en el sistema operativo y la CPU del ordenador para gestionar y organizar múltiples unidades de almacenamiento en un único sistema lógico, o sea, utiliza los recursos del controlador del host como controlador, lo que significa que todo el procesamiento de datos se realiza en el motherboard del sistema.

3) **Estado del Arte sobre los RAID.**

Según, **lifeder.com**, que es una web fundada y gestionada por amantes de la lectura. Entre ellos psicólogos, filósofos y Licenciados en Literatura. Actualmente es una de las webs de ciencia y cultura general con más visitas en Latinoamérica y España, opina que:

El estado del arte de una investigación, TFG, TGM, proyecto o tesis, se refiere al estado de las investigaciones o trabajos que ya existen sobre el tema sobre el cual se desea investigar y su objetivo es dar validez a la investigación, mostrando que generará conocimiento nuevo y no se trata de una copia de otra investigación. Es decir, se refiere a las fuentes de información existentes que abordan temas similares al propuesto en la tesis o trabajo de investigación.

Dado que todo proyecto de investigación debe aportar nueva información o resultados sobre un tema en particular, se hace necesaria la construcción de un estado del arte, por lo que, debe tener una estructura.

La estructura que debe tener el apartado de estado del arte es la siguiente:

Introducción: se presenta el tema a investigar y se mencionan los antecedentes encontrados en el área a estudiar, así como los límites de la investigación.

Investigaciones consultadas: se hace referencia a trabajos similares, que hayan tenido resultados distintos, o que fueron realizados bajo diferentes condiciones.

Conclusiones: se establecen los motivos por los cuales se llevará a cabo el estudio y cómo este aportará conocimiento nuevo al campo de investigación.

¿Cómo hacer un estado del arte?

Según, lifeder.com, para elaborar un estado del arte, se sugiere seguir los siguientes pasos:

- Determinar la pregunta de investigación: esto permitirá al investigador enfocarse en la situación actual y las categorías del tema sobre el cual pretende investigar.
- Buscar trabajos preexistentes sobre la investigación que se desea llevar a cabo y estudiarlos de manera exhaustiva.
- Escribir la introducción, detallando los antecedentes encontrados en el área a estudiar y los límites de la investigación.
- Hacer referencia a las investigaciones consultadas, citando las distintas investigaciones y estableciendo los motivos por los cuales se llevará a cabo el estudio.
- Redactar las conclusiones, mostrando cómo la investigación aportará conocimiento nuevo al campo de estudio.

Ahora bien, según **redalyc.org**, que es la Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc). Esta red de revistas, que es un proyecto académico para la difusión en acceso abierto de la actividad científica editorial de todo el mundo, bajo un modelo liderado por la academia y no lucrativo, opina sobre el Estado del Arte sobre los RAID, lo siguiente:

Es importante mencionar que, si bien se pueden identificar etapas comunes para realizar el estado del arte, también debe entenderse que la elaboración de este corresponde a un diseño investigativo cíclico y dialéctico, donde el investigador debe preguntarse constantemente hasta qué momento debe seguir haciendo una interpretación de la bibliografía que ha recuperado, si es necesario recuperar y analizar más fuentes, y finalmente debe decidir el momento en el que cerrará ese ciclo. En resumen, el estado del arte es una parte esencial en la investigación, ya que permite conocer el estado actual de las investigaciones en un tema específico y establecer cómo la nueva investigación aportará conocimiento adicional. La estructura del estado del arte incluye una introducción, las investigaciones consultadas y las conclusiones. Para elaborar un estado del arte, es necesario determinar la pregunta de investigación, buscar trabajos preexistentes, redactar la introducción, hacer referencia a las investigaciones consultadas y redactar las conclusiones.

4) **Unidades Storage para RACKS de servidores.**

En los Centros de Procesamiento de Datos (CPD), pequeñas y medianas empresas, e incluso oficinas, hay unos armarios metálicos

que tienen los principales sistemas informáticos y redes de comunicaciones. A estos armarios se les denomina armarios informáticos o simplemente rack, y son la forma de almacenar de forma segura todos los equipos en las empresas, y ordenar correctamente todo el cableado por si hay que realizar una intervención en el mismo.

Un rack es un armario con base y estructura metálica, cuyo principal objetivo es alojar todos los sistemas informáticos y de telecomunicaciones que necesiten las empresas o las organizaciones donde estén instalados. En estos rack, es donde se almacenarán los servidores, switches, routers profesionales, servidores NAS, DAS, sistemas de alimentación ininterrumpida (SAI) y, por supuesto, todo el cableado de red para proporcionar los diferentes servicios a la red local.

Tipos de rack.

Existen principalmente dos tipos de rack: los modelos de suelo y los modelos de mural.

1.- **Rack mural:** Este tipo de armarios nos permite ubicarlos encima de una mesa o directamente en la pared, a modo de estantería. Si necesitamos un armario para instalar un switch y un router únicamente, es una muy buena opción porque su costo es muy bajo, y es idóneo para centros de formación, colegios donde no necesiten demasiada conectividad, hogar doméstico etc.

2.- **Racks de suelo:** Este tipo de armarios están específicamente diseñados para tenerlos directamente en el suelo, tienen ruedas bloqueables para moverlo de forma fácil y rápida, no necesitaremos anclarlo a la pared, se pone directamente en el suelo y se bloquean las ruedas para evitar su movimiento. Si necesitamos un armario para instalar un router, varios switches y servidores, esta es la opción porque proporciona el mayor espacio de altura. Este tipo de armarios están orientados a pequeñas y medianas empresas donde hayan más equipos informáticos que almacenar.

Al elegir unidades de almacenamiento para racks de servidores, es importante tener en cuenta la capacidad de almacenamiento, el rendimiento y la confiabilidad necesarios para las aplicaciones y los usuarios finales. También, es fundamental considerar la gestión del calor y la ventilación adecuada dentro del armario de servidores. Los racks de servidores suelen estar diseñados con ventiladores y puertas perforadas para permitir la circulación del aire y mantener los equipos frescos.

En resumen, al elegir unidades de almacenamiento para racks de servidores, es importante considerar factores como capacidad, rendimiento, confiabilidad y gestión del calor.

5) NAS (Almacenamiento conectado a la red).

La NAS (Network Attached Storage), se refiere a un servidor dedicado al almacenamiento compartido. El propósito principal de un NAS es compartir datos entre varios computadores, ya sean Macs o PCs, en una red local. También se conoce como caja NAS, unidad NAS, servidor NAS o cabezal NAS. Funciona como un banco de datos centralizado al cual se puede acceder de forma remota dentro de la red a la que está asociado. En otras palabras, los NAS son dispositivos de almacenamiento que se conectan a una red TCP/IP y permiten a los usuarios almacenar y compartir archivos de manera centralizada, facilitando el acceso y la administración de los datos.

Los usuarios de una red pueden acceder al almacenamiento compartido de la NAS a través de una conexión Ethernet estándar. Cada NAS se configura y administra con una utilidad basada en navegador y es un nodo de red independiente definido por su propia dirección IP. Los NAS se utilizan para manejar datos no estructurados, como: audio, video, sitios web, archivos de texto y documentos de Microsoft Office. Los NAS se utilizan en el hogar y en pequeñas empresas. Los usuarios pueden agregar almacenamiento o configurar copias de seguridad dobles. Además, algunos centros de datos utilizan el almacenamiento de archivos en la nube en lugar de o junto con la NAS física.

La NAS se conecta a la LAN (Red de Área Local) como un nodo de red independiente y utiliza una microkernel especializada para manejar solo solicitudes de entrada/salida. Los NAS son muy sencillos de agregar o eliminar de la red, pero están sujetos al comportamiento variable y la sobrecarga de la red.

La seguridad es un aspecto importante a tener en cuenta al utilizar un NAS, ya que, estos dispositivos son vulnerables a ataques externos y espionaje.

6) Análisis del 1er. Caso crítico sobre RAID 0, 1, 5, 10.

RAID 0:

El primer caso crítico en RAID 0, es la falla de una sola unidad de disco en el conjunto RAID. RAID 0, distribuye los datos en bloques a través de múltiples unidades de disco, lo que aumenta el rendimiento y la velocidad de lectura/escritura. Sin embargo, no ofrece redundancia de datos. Cuando una de las unidades falla, todos los datos almacenados en el conjunto RAID 0 se pierden, ya que, la información se distribuye uniformemente entre los discos y no hay copias adicionales de los datos.

Aquí, hay un ejemplo para ilustrar el primer caso crítico en RAID 0:

Supongamos que tenemos 2 unidades de disco (D1 y D2) en un conjunto RAID 0. Los datos se distribuyen en bloques entre D1 y D2 de la siguiente manera:

D1: | A1 | A2 | A3 | A4 |

D2: | B1 | B2 | B3 | B4 |

Cada letra representa un bloque de datos, y los números representan la secuencia en la que se distribuyen los datos. En este ejemplo, A1 y B1 son partes del mismo archivo, A2 y B2 son partes de otro archivo, etc.

Si la unidad D1 falla, todos los bloques A1, A2, A3 y A4 se perderán. Debido a que los datos están distribuidos, no es posible recuperar los archivos completos solo con los bloques restantes en D2 (B1, B2, B3 y B4). En este caso, todos los datos almacenados en el conjunto RAID 0, se consideran irrecuperables.

La falta de redundancia y la posibilidad de perder todos los datos en caso de falla de una sola unidad de disco, hacen que RAID 0 sea un caso crítico para la protección de datos. Es importante tener en cuenta este riesgo al utilizar RAID 0 y asegurarse de realizar copias de seguridad regulares de los datos importantes almacenados en el conjunto RAID 0.

RAID 1:

El primer caso crítico en RAID 1, ocurre cuando uno de los discos falla, y se puede visualizar a través del siguiente ejemplo:

En RAID 1, se crea un espejo de los datos almacenados en dos discos diferentes (disco A y disco B). Si representamos los datos almacenados en cada disco como un conjunto, podemos ilustrar RAID 1 de la siguiente manera:

Conjunto A = {a1, a2, a3, ..., an} Conjunto B = {b1, b2, b3, ..., bn}

Donde "n" es el número total de bloques de datos en cada disco, y cada elemento "ai" en el Conjunto A es igual a su correspondiente elemento "bi" en el Conjunto B ($a_i = b_i$). En otras palabras, los dos conjuntos son iguales:

Conjunto A = Conjunto B

Supongamos que el disco A falla, entonces el Conjunto A, ya no está disponible. El sistema aún puede funcionar utilizando solo el Conjunto B, pero, la redundancia se pierde temporalmente. Para mantener la redundancia, es necesario reemplazar el disco A y reconstruir el Conjunto A, a partir del Conjunto B:

Conjunto A' = Conjunto B

Donde "A'" representa el nuevo disco reemplazado y reconstruido. Al finalizar la reconstrucción, los dos conjuntos vuelven a ser iguales, y la redundancia se restaura:

Conjunto A' = Conjunto B

RAID 5:

El primer caso crítico sobre RAID 5, estaría relacionado con la falla de un disco en la configuración y se inicia el proceso de reconstrucción. Los datos se distribuyen en paridad a través de varios discos, lo que proporciona una tolerancia a fallos de un solo disco. Sin embargo, si un disco falla, la reconstrucción de los datos se vuelve crítica, ya que, una segunda falla de disco durante la reconstrucción resultaría en la pérdida de todos los datos en la matriz RAID 5.

Por ejemplo, un análisis del primer caso crítico de RAID 5, relacionadas con la paridad y la reconstrucción de datos, sería:

Consideremos un RAID 5 de 3 discos con un tamaño de bloque de 64 KB.

Cada bloque de paridad se calcula utilizando la operación XOR en los bloques de datos correspondientes. Por ejemplo, en la banda #0, el valor del byte #0 en el bloque de paridad (en el Disco 2) se calcula de la siguiente manera:

Parity byte #0 = (Disk 0 byte #0) XOR (Disk 1 byte #0)

Cuando un disco falla, por ejemplo, el Disco 2, los datos perdidos se pueden reconstruir aplicando la operación XOR en los discos sobrevivientes. En este caso, para recuperar el byte #0 del Disco 1, se aplica XOR entre el byte #0 del Disco 0 y el byte #0 de paridad:

Disk 1 byte #0 = (Parity byte #0) XOR (Disk 0 byte #0)

Este proceso de reconstrucción se realiza para cada byte en los bloques afectados por la falla del disco.

RAID 10:

El primer caso crítico en RAID 10, sería la pérdida simultánea de ambos discos en un par de espejos, lo que resultaría en la pérdida de todos los datos en ese par. RAID 10, puede soportar la pérdida de un solo disco en cada par de espejos sin pérdida de datos. Sin embargo, si se pierden dos discos en el mismo par de espejos, todos los datos en ese par se perderán.

Supongamos que tenemos un conjunto de n discos, donde cada par de discos forma un espejo en RAID 10. La probabilidad de que un disco falle es p , y la probabilidad de que un disco no falle es $(1-p)$.

Para que ocurra el primer caso crítico (la pérdida de ambos discos en un par de espejos), ambos discos en el mismo par deben fallar simultáneamente. La probabilidad de que esto ocurra en un par específico de discos es:

$p * p$, es decir, p^2 .

Ahora bien, dado que hay $n/2$ pares de espejos en un sistema RAID 10, la probabilidad de que ocurra el primer caso crítico en cualquier par de espejos es:

$P(\text{Primer caso crítico}) = 1 - P(\text{Todos los pares de espejos no tienen fallas críticas})$
 $P(\text{Primer caso crítico}) = 1 - (1 - p^2)^{(n/2)}$

Esta fórmula nos permite, calcular la probabilidad del primer caso crítico en función del número de discos (n) y la probabilidad de falla de un disco individual (p). Se pueden ajustar estos parámetros según el hardware y las condiciones específicas del sistema RAID 10, para estimar la probabilidad de que ocurra el primer caso crítico.