

Arquitectura del Computador y Sistemas Operativos

Decimoctava Clase



Manejadores de Volúmenes (1/9)

Un “Volumen” es similar a una partición, pero más flexible. A diferencia de las particiones los volúmenes pueden:

- Además de ser una parte del disco completo (como son las particiones), los volúmenes pueden abarcar más de un disco.
- Se puede manejar un conjunto no contiguo de tracks como una sola unidad, cosa que las particiones no permiten (al menos sin mover datos).
- Si bien algunos aspectos de las particiones pueden modificarse con el disco en uso, los volúmenes son mucho más flexibles, pudiendo manipularse por completo con el sistema online.

Los manejadores de volúmenes o VMs (por sus siglas en inglés Volume Managers) son programas para administrar los dispositivos de almacenamiento masivo.

Manejadores de Volúmenes (2/9)

El Volume Manager lógicamente se intercala entre los RAIDs (que proporcionan redundancia, striping, etc) y los sistemas de archivos de la siguiente forma:

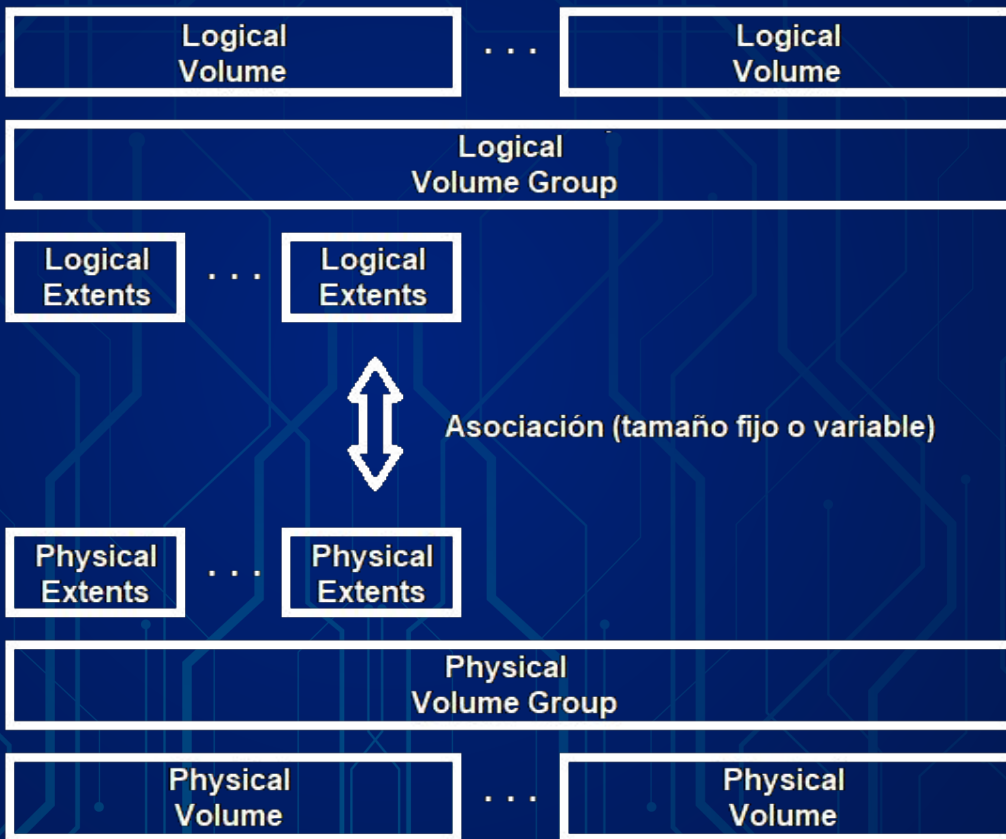
**Sistema de Archivos
(FS)**

**Manejador de volúmenes
(LVM)**

**Capa de concatenación
y redundancia (RAIDs)**

Manejadores de Volúmenes (3/9)

Diagrama de Bloques





Manejadores de Volúmenes (4/9)

Para mejorar la velocidad, algunos VMs suelen:

- Hacer mapeo de extents de tamaño variable (que se ajusten a tracks o múltiplos de tracks) de los discos que tiene debajo.
- Distribuir los extents haciendo “*striping*” entre los discos físicos que componen el volumen físico.

Como los PEs (Physical Extents) no tienen que ser contiguos (de hecho a veces hasta a propósito se separan al hacer “*striping*”), se puede hacer crecer el volumen sin tener que mover datos aún si ya hay otro volumen usando los PEs siguientes.



Manejadores de Volúmenes (5/9)

Algunos VMs además soportan:

- Snapshots
- Volúmenes híbridos
- Thin Provisioning
- RAIDs / File Systems

Manejadores de Volúmenes (6/9)

Snapshots

Los “*SnapShots*” (o “Instantáneas”) son como fotos del volumen en un determinado instante de tiempo.

A partir de recibir el comando para guardar un “*Snapshot*”, el “Volume Manager” no sobrescribe los extents cuando el Sistema Operativo le indica que guarde algo. En realidad primero mueve el valor a reemplazar a una zona libre (fuera del volumen) y recién después guarda el nuevo valor.

Si el Sistema Operativo luego quiere volver a modificar algún dato de ese extent cuya copia ya fue hecha, ahí sí se lo sobrescribe directamente.

En cualquier momento es posible acceder tanto al estado del disco actual como el que tenía al momento del Snapshot.

En los VMs se puede crear más de un Snapshot, por lo que puede haber varios estados del volumen a los que se puede volver.



Manejadores de Volúmenes (7/9)

Volúmenes Híbridos (1/2)

Los volúmenes híbridos utilizan discos de mayor velocidad para funcionar como cachés de los discos más grandes y lentos.

Se pueden utilizar discos de estado sólido como caché de discos magnéticos o RAM drives como caché de discos de estado sólido.

Pueden venir en dos versiones:

- Multi-drive: Se combina un conjunto de discos rápidos como cachés de discos más grandes y lentos.
- Solid state hybrid drive: Conocidos por sus siglas SSHD consisten en un solo disco con ambas tecnologías dentro.



Manejadores de Volúmenes (7/9)

Volúmenes Híbridos (2/2)

Al día de la fecha Linux contiene dentro de su kernel dos mecanismos para implementar volúmenes híbridos “Multi-Drive” por software:

- **bcache** Es un simple caché de bloques que se puede asociar a cualquier tipo de unidad. Tiene buena performance y es fácil de configurar.
- **dm-cache** Es una opción de mucho más bajo nivel incorporada al device mapper. No es muy útil por sí misma, pero es usada por LVM para implementar sus volúmenes con caché.

Por ahora Windows no incorpora opciones nativas de caché de dispositivos de almacenamiento, de ser necesarias se debe comprar una como “*PrimoCache*”. Si bien no está incluida con el SO su costo es de aproximadamente USD 50 por servidor.



Manejadores de Volúmenes (8/9)

Thin Provisioning

Esta tecnología permite crear discos de mayor tamaño que el existente. El Sistema Operativo los ve como un disco de gran tamaño, pero inicialmente ocupa pocos extents en los discos físicos reales.

A medida que el disco se va usando, se van allocating más extents de un pool y el volumen va creciendo físicamente.

Como varios volúmenes pueden sacar extents del mismo pool, los discos van creciendo a medida que los usuarios los ocupan, optimizando la cantidad de disco disponible.

En este tipo de proceso hay que tener cuidado, porque pueden terminar existiendo discos con sus extents muy fragmentados y perder performance.



Manejadores de Volúmenes (9/9)

RAIDs / File Systems

Hemos estudiado el almacenamiento en tres capas, pero hay en el mercado productos que abarcan más de una capa.

Por ejemplo ZFS es un File System que además es un manejador de volúmenes. Dice tener la ventaja de poder “conocer en detalle” cómo están ubicados los extents en los discos y usar esa información para mejorar la performance.

El Logical Volume Manager (LVM) de Linux, además de manejador de volúmenes puede administrarlos armando RAIDs.





Más allá de los PetaBytes ...

Si bien los dispositivos de almacenamientos para Data Centers (Storages) tienen lugar para colocar muchos discos, siempre hay un límite en la cantidad de almacenamiento posible.

Para sistemas que almacenan una enorme cantidad de datos, la única opción posible es trabajar con muchos sistemas más pequeños trabajando juntos (en Cluster).

Windows ofrece el S2D (Storage Spaces Direct) como parte de sus productos Windows Server®.

Bajo Linux hay varios proyectos, entre ellos GlusterFS, Ceph y ZFS.



Cambio de tamaño de volúmenes (1/2)

Hay que tener mucho cuidado al modificar los volúmenes en uso, dado que contienen datos en algún formato (FATx, EXTx, NTFS, etc.). Todos los formatos almacenaron en el volumen estructuras de datos que están ajustadas para el tamaño del mismo.

Por ejemplo, si se agranda un volumen y se verifica el tamaño del disco (por ejemplo usando “df”) lo que obtendremos es exactamente lo mismo que antes de agrandarlo. Peor aún, si lo achicáramos y luego verificáramos el tamaño del disco, también obtendríamos lo mismo que antes de achicarlo, con el componente adicional que al momento en que se trataran de usar los sectores eliminados tendríamos errores y por ende corrupción de datos.

Para evitar problemas estos problemas siempre se debe:

Agrandar el Volumen		Achicar el Volumen	
1)	Agrandar el volumen	1)	Achicar el file system
2)	Agrandar el file system	2)	Achicar el volumen



Cambio de tamaño de volúmenes (2/2)

Por ejemplo, bajo Linux, si necesitamos cambiar el tamaño de una partición con formato EXT2/EXT3/EXT4, podemos usar el comando *“resize2fs”*. El comando detecta automáticamente si el volumen está montado. Si es así comienza automáticamente a trabajo online.

El trabajo online sólo permite agrandar el formato (no achicarlo), pero tiene la ventaja que se hace con el sistema corriendo.

Bajo Windows, el formato también puede extenderse sin detener el sistema usando en *“Disk Manager”*. Lamentablemente los volúmenes deben haber sido creados con formato dinámico (caso contrario hay restricciones que impiden que los extents no sean contiguos). Tampoco pueden alterarse los volúmenes que tienen el sistema o archivos de memoria virtual.

An abstract pattern of light blue lines and dots resembling a circuit board, located on the left side of the slide.

Fin
¿Preguntas?

An abstract pattern of light blue lines and dots resembling a circuit board, located on the right side of the slide.