

RAID pra que te quero

Sistemas RAID são matéria-prima de servidores de arquivos, estações de edição digital e outros sistemas que não podem dar pau jamais

Múltiplos HDs em conjuntos redundantes aumentam a velocidade e a segurança dos seus dados

por Márcio Nigro



Quanto valem os seus dados? Não estamos falando de valor sentimental. A pergunta diz respeito a dinheiro, *money*, grana, dim-dim, mufunfa, verdinhas, tutu. Quanto mais você achar que eles valem, melhor tem que ser o método de armazenamento para garantir a segurança desses bens tão preciosos.

A gente vive falando nas páginas desta humilde revista que é preciso fazer backup e coisa e tal para garantir a integridade das informações no seu computador. Mas, em certos casos, só backup não resolve. É o caso de quem tem aplicações de missão crítica, como servidores de arquivos, grandes bases de dados e programas *high-end* de vídeo digital, por exemplo. Para esses, a melhor resposta é um sistema RAID.

Para aqueles casos em que o backup sozinho não resolve

Mas o que é RAID?

Não sabe direito do que estamos falando? Então, expliquemo-nos.

RAID significa “*Redundant Array of Inexpensive Disks*” ou, em português claro, “fileira redundante de discos baratos” – se bem que, para usar um termo mais adequado, substituiremos “fileira” por “matriz”, porque muitas vezes os discos não formam exatamente uma fila. O conceito surgiu em 1987, como uma proposta que pregava o uso de múltiplos discos pequenos e acessíveis para substituir o uso de um disco único e grande (SLED, *Single Large Expensive Drive*) para armazenar dados com redundância e tolerância a falhas. Com um sistema RAID, embora a informação seja dividida entre vários discos

rígidos (ou outras unidades de armazenamento), tudo aparece no computador como se fosse apenas um drive.

É importante ressaltar que o termo “discos baratos” não deve ser levado muito ao pé da letra, já que a maioria dos sistemas RAID empregam as últimas versões da interface SCSI para aumentar as taxas de transferência. Discos SCSI não são exatamente baratos, mas na época em que apareceu o RAID eles eram muito mais baratos que os HDs dos computadores de grande porte, proporcionando boas taxas de transferência sustentáveis e reduzindo bastante a latência e o tempo de procura. (Posteriormente vieram os drives com interface IDE, que baixaram os preços mais ainda.) Na verdade, um sistema RAID pode ser bem caro e a aplicação precisa justificar a implementação desse tipo de tecnologia. ▶

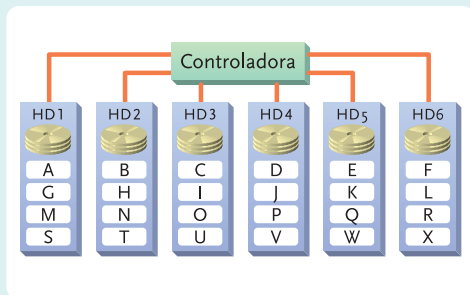
Os vários tipos de RAID

Existem várias arquiteturas RAID, cada uma oferecendo diferentes recursos de segurança, como redundância, tolerância a falhas etc. A forma de expandi-las também varia: alguns sistemas podem ser expandidos com a adição de apenas um disco, enquanto outros requerem usar cinco de uma vez. O desempenho também muda de nível para nível, de modo que é importante definir as intenções de suas aplicações e a performance desejada para saber qual tipo de RAID é adequado à sua necessidade.

Os principais fatores que diferenciam os níveis de RAID são o modo como os dados são gravados no conjunto de discos; se há alguma forma de redundância de dados; e se e como o conjunto pode ser expandido. Veja a seguir uma explicação mais detalhada de cada nível.

Descubra e adote o melhor nível de RAID para a sua necessidade

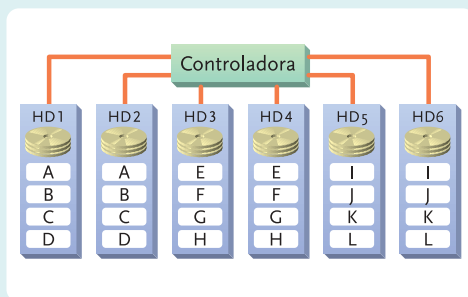
RAID-0



O RAID nível 0 também é conhecido como “fatiador de dados” (*data striping* ou simplesmente *stripe*). É um conjunto de discos que transferem os dados em paralelo. Essa categoria oferece alta performance e transparência para o usuário. Como não há redundância (para essa configuração o termo RAID não é muito adequado), a capacidade de armazena-

mento do sistema é igual à soma de todos os drives utilizados. A ausência de redundância também significa que não há paridade para recuperação de erros e tolerância a falhas. Por outro lado, é muito fácil de se implementar e expandir esse nível, que requer pelo menos dois drives, sendo possível adicionar outro a qualquer momento. De modo geral, quanto mais discos maior o desempenho do conjunto, o qual está limitado apenas pela capacidade de barramento. E mesmo assim, quando um barramento está saturado, é possível usar outros. “Fatiando” dois HDs, é possível aumentar a performance em 80%. O RAID-0 é bastante popular em aplicações de edição de vídeo digital e outras que necessitam de grande largura de banda.

RAID-1



O RAID nível 1 funciona basicamente como espelhamento de disco. O sistema requer um mínimo de dois discos para que possa haver redundância dos dados, e tende a ser mais rápido na leitura e mais lento na gravação do que apenas um HD. Cada dado é escrito ao mesmo tempo em dois drives e, se um deles falhar, o outro assume a função. Por causa da redundância, a capacidade de armazenamento é metade da soma de todos os discos, de mo-

Para saber mais

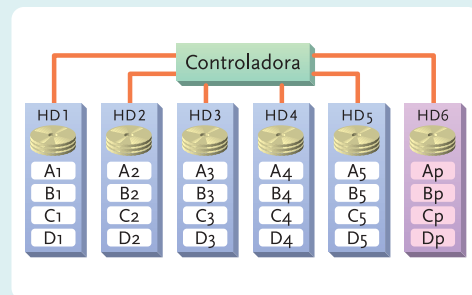
NetApp: www.netapp.com

Quantum: www.snapappliances.com

Electronix: www.raidweb.com

do que o custo por megabyte é alto. Os HDs têm que ser adicionados de dois em dois na hora de expandir o conjunto. Essa configuração é usada normalmente em aplicações de contabilidade, pagamentos e finanças.

RAID-3



O RAID nível 3 reparte os bytes entre os discos, com a paridade gravada em uma única unidade. Os dados são entrelaçados em nível de byte, e o acesso a um simples bloco requer a transferência de todos os drives que, sincronizados, funcionam como se fosse um só disco virtual. O RAID-3 oferece transferência rápida de arquivos grandes, como gráficos e vídeo digital. No entanto, pode até ser mais lento do que um único volume em aplicações normais.

Esse nível aproveita melhor o espaço em disco do que o RAID-1: 83% a 90% está disponível, dependendo da quantidade de drives do conjunto. No entanto, o disco de paridade torna o sistema 10% a 20% mais custoso do que usar o mesmo número de volumes separados. Expandir a capacidade do sistema RAID-3 é fácil. É indicado para produção de vídeo, *streaming* em tempo real e editoração.

Fique ligado: palavras-chave do mundo RAID

Os sistemas RAID também podem conter subsistemas específicos que variam com o nível e o fornecedor. Veja de alguns dos recursos e conceitos mais comuns.

•**Tolerância a falhas** – É a habilidade de manter o conjunto operando quando um ou mais dos seus componentes falham.

•**Paridade** – É um método pelo qual os dados são distribuídos em vários volumes de tal modo que, numa eventual falha de um disco, as informações são comparadas com as dos discos restantes da matriz a fim de reconstruir os dados perdidos. No processo de paridade, a soma do número total de bits ‘1’ tem que dar um número ímpar (*odd parity*) ou par (*even parity*). SCSI usa paridade ímpar. Isso significa que a soma de todos os ‘1’ de um byte mais a paridade será sempre ímpar.

•**Caching** – Algumas controladoras RAID — especialmente as de níveis 3, 4 e 5 — têm RAM para pré-alocar dados ou fazer cache. Isso ajuda a compensar as perdas de performance de gravação nos níveis mais altos de RAID. Sistemas sofisticados de *caching* incluem becape de bateria, capaz de manter os dados por horas em caso de falta de energia. Alguns até oferecem cache espelhado para redundância.

•**Hot Spare** – Alguns subsistemas RAID incluem um drive extra que só é usado quando outro drive falha.

•**Warm Spare** – Requer que, na hora de reposição de discos ou componentes, o sistema fique *offline*, mas não desligado.

•**Hosts múltiplos** – Recurso que possibilita que múltiplos adaptadores ou computadores *host* (hospedeiro) sejam conectados ao mesmo tempo.

•**Controladoras redundantes** – Os subsistemas

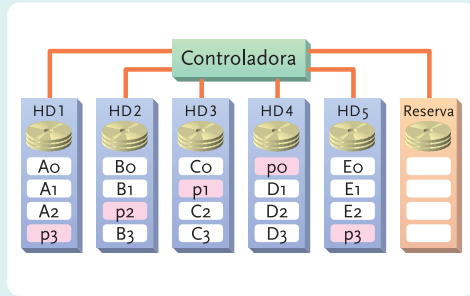
RAID que contêm duas controladoras garantem que, se uma falhar, a outra assume a função imediatamente. Em alguns casos, as duas funcionam simultaneamente e, em outros, uma fica em *standby*.

•**Hot Swappable** (“troçável a quente”) – Vários sistemas RAID permitem que os discos e fontes de alimentação sejam trocados enquanto o sistema está ligado e funcionando.

•**Software de configuração** – Os subsistemas RAID podem ser configurados de três maneiras: software com interface gráfica, painéis LCD e terminal baseado em conexões seriais.

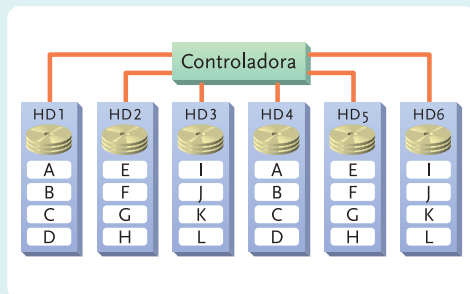
•**Software de monitoração** – Sistemas RAID mais avançados incluem programa para checar o *status* dos componentes na rede. Alguns até incluem recursos de notificação automática.

RAID-5



Oferecendo alta segurança, o RAID nível 5 é similar ao RAID-3, só que distribui a informação de paridade entre os drives. Mais de um disco pode ser gravado ao mesmo tempo e há melhor desempenho na transferência de blocos pequenos de dados (mas é mais lento para grandes arquivos). A gravação, no entanto, não é das mais rápidas. Um sistema com oito unidades pode gravar quatro discos simultaneamente, e não é necessário nenhum volume dedicado de checagem, já que a paridade está espalhada pelos volumes. 10% a 20% do espaço disponível é usado para criar paridade. O RAID-5 requer pelo menos três discos e é simples de expandir. Os dados perdidos num disco defeituoso podem ser reconstruídos num disco extra que funciona como “estepe”. Esse nível é ótimo para servidores de arquivos, bancos de dados e servidores da Internet, mas não é aconselhável para vídeo e gráficos.

RAID-0+1



Esse sistema é implementado na forma de um espelho de subconjuntos (como no RAID-1) de drives segmentados em *stripe* (como no RAID-0). Utiliza quatro ou mais HDs, em números pares. Oferece altas taxas de transferência e a mesma tolerância a falhas do nível 5. Por outro lado, a falha de apenas um disco faz com que toda a matriz vire, em essência, um sistema RAID-0. Há parentesco entre o RAID-0+1 (espelhos de drives em *stripe*) e um outro esquema, chamado RAID-10 (*stripe* de drives espelhados), que é mais tolerante a falhas.

Variações tecnológicas

Os sistemas RAID podem ser baseados em vários tipos de tecnologia:

• **Software** – Um programa calcula a paridade e os comandos necessários para criar o RAID.

• **Hardware** – Realiza os cálculos e envia os comandos, liberando o computador para outras tarefas.

• **Hardware em placa** – A própria placa controladora já possui o hardware necessário para realizar a tarefa.

• **Hardware externo** – Equipamento que possui um conector para ser ligado ao computador e um ou mais para se ligar aos discos rígidos.

Para o computador, o uso do esquema RAID é transparente, enxergando a controladora externa como se fosse um único disco. Muitas dessas controladoras possuem dimensões similares às de um HD tradicional, permitindo a sua fácil montagem interna em gabinetes para servidores. Porém, você continua precisando de uma placa PCI no computador para se conectar a essa controladora externa.

• **Com uma interface** – Inclui uma interface SCSI e todos os discos são ligados nela.

• **Com duas ou mais interfaces** – Permite que os discos sejam distribuídos entre as várias interfaces, aumentando a performance, já que cada interface (ou canal SCSI) fica mais livre para os discos.

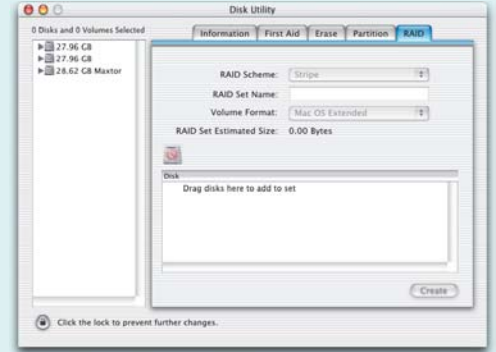
• **IDE x IDE** – É uma controladora RAID (hardware externo) que se conecta na controladora IDE, possibilitando assim utilizar discos IDE.

• **SCSI x SCSI** – Conecta-se na controladora SCSI e usa discos SCSI.

• **SCSI x IDE** – Conecta-se na controladora SCSI e usa discos IDE.

• **Fiber Channel x Fiber Channel** – Conecta-se na controladora Fiber Channel e usa discos Fiber Channel.

• **Fiber Channel x SCSI** – Conecta-se na con-



O Disk Utility do Mac OS X 10.1 pode configurar seus HDs em RAID

troladora Fiber Channel e usa discos SCSI.
• **Fiber channel x IDE** – Conecta-se na controladora Fiber Channel e usa discos IDE.

Qual escolher?

Para os usuários profissionais de Mac, os sistemas RAID que devem ser mais interessantes são os de grande porte e alta velocidade, que se podem montar a preços razoavelmente baixos – para uso em aplicações de vídeo, por exemplo – a partir de discos IDE e conexão SCSI. É possível montar um RAID de 100 GB e alta performance por menos de R\$ 5 mil e usá-lo com o Media 100 ou Final Cut Pro. Para ter uma base de comparação, há quatro anos um sistema de 50 GB para Flame (em SGI) saía por uns R\$ 31 mil. Enfim, tudo está mais fácil agora. **M**

MÁRCIO NIGRO

Colaboraram (e muito) Oswaldo Bueno e Carlos Freitas

NAS x SAN

Uma questão adjacente ao RAID são a NAS e a SAN. Uma SAN (*Storage Area Network*) é uma rede de armazenamento. Por exemplo, é possível configurar alguns servidores ligados na mesma cadeia SCSI, acessando os mesmos drives. O padrão mais atual é utilizar a interface Fiber Channel, na qual vários servidores são conectados entre si e aos HDs ou controladoras RAID. Na maioria dos casos, cada servidor acessa partições diferentes. A vantagem é a centralização do armazenamento num único local ou num local distante e seguro. Diferente do que o nome sugere, a tecnologia Fiber Channel não funciona apenas com fibra óptica, mas também com cabos de cobre. Alguns servidores conseguem compartilhar até a mesma partição, mas para isso utilizam software que faz a comunicação entre um servidor e outro avisando quais arquivos estão abertos e não podem ser acessados pelo outro.

As NAS (*Network Attached Storage*) são *network appliances*, ou seja, sistemas completos que incluem computador, sistema operacional, controla-

dora SCSI, interface de rede e, em alguns casos, os próprios HDs. Esses equipamentos requerem um mínimo de configuração inicial e são ligados na rede Ethernet, comportando-se como servidores para os outros computadores e suportando vários protocolos compatíveis com Windows, Unix e Mac (AppleShare). No entanto, uma NAS não é um RAID. Um dos fornecedores de NAS mais conhecidos é a NetApp. Seus equipamentos possuem sistema operacional e tecnologia de RAID próprios, que não são comparáveis aos níveis RAID.

O sistema da NetApp faz acesso aos discos para leitura mas, se recebe um comando de gravação, pega o bloco mais próximo disponível, grava e continua a leitura. A maior vantagem do NAS é que, na maioria dos modelos, você pode “espeter” mais um disco sem que o equipamento precise reformatar e reparticioná-los, ao contrário do que acontece com os sistemas RAID.

Outro fornecedor é a Quantum, que tem sistemas NAS de apenas um disco.