

Tipos de RAID's e tipos de clusters

Apresentar os principais tipos de RAID's responsáveis pelo aumento da disponibilidade dos servidores e Storages e outros equipamentos que permitem utilizar essa tecnologia, bem como detalhar os principais tipos de cluster que também são utilizados para aumentar a disponibilidade de um ambiente de servidores.

Como foi visto no módulo anterior, a necessidade de a organização ter que possuir a alta disponibilidade nos principais ativos que suportam o negócio da empresa fez com que os responsáveis pela infraestrutura de servidores implementassem soluções que aumentassem o grau de proteção dos seus servidores.

Essas medidas partiram desde a duplicação de componentes dos servidores, bem como a duplicação de dispositivos que ajudavam esse servidor a ficar ligado.

Dentre os componentes que ajudam a manter a alta disponibilidade dos servidores existe um que é responsável pela segurança dos discos (HD) desses servidores.

A ideia seria colocar uma proteção para esses discos, de forma que se houvesse uma falha, os equipamentos continuariam funcionando mesmo que parcialmente, até o momento que pudesse ser feita a troca do dispositivo.

Antes de abordar esse tema, faz-se necessário o entendimento de alguns itens relacionados a seguir:

- **Disco SCSI (Small Computer System Interface)** – trata-se de uma tecnologia criada para acelerar a taxa de transferência de dados e utilizadas em discos que suportam HDC. O HDC possibilita conectar e desconectar discos rígidos sem a necessidade do desligamento da máquina.
- **HDC** – (Hot Disconnect and Connect ou Conectar e desconectar HDs a quente).
- **Hot Spare** – está relacionado ao fato de um dispositivo estar pronto para ser usado para que seja feito o processo de reconstrução de um disco quando ele apresenta uma falha.
- **Rebuild** – esse termo está associado à palavra reconstrução. Geralmente quando um disco falha e o processo de reconstrução é iniciado, o termo Rebuild dos discos é utilizado.
- **Array** – agrupamento ou conjunto de disco.
- **O RAID** (Redundant Array of Independent Disks) – matriz ou conjunto redundante de discos independentes.

Podemos definir o RAID como sendo um sistema formado por um conjunto de discos (HD) independentes que trabalharão juntos (formando uma única unidade lógica) para deixar um equipamento com melhores critérios de tolerância a falhas "Redundantes". Essa tecnologia foi desenvolvida para aumentar o desempenho de acesso e a confiabilidade dos dados em um servidor.

O sistema de RAID possibilita que o responsável pela segurança dos servidores possa escolher qual tipo implementar, de acordo com a necessidade de capacidade, desempenho, confiabilidade que o seu cenário de proteção necessitar.

Tipos de RAID

A tecnologia RAID funciona de várias "maneiras", ou seja, de acordo com o agrupamento dos discos que fazem parte da sua estrutura. Tais "maneiras" como os discos são agrupados são conhecidas como "níveis de RAID".

Dessa forma, existem vários níveis de RAID que podem ser utilizados como forma de aumentar a disponibilidade do sistema, porém no nosso módulo abordaremos os mais utilizados no mercado. (RAID 0, RAID 1, RAID 1+0, RAID 0+1, 5 e RAID 1 + 5).

RAID nível 0

Esse tipo de RAID recebe o nome de "striping". É feito com a utilização de no mínimo dois HD's, em que o armazenamento dos dados serão divididos entre os discos que fazem parte do sistema. Possui como vantagem a alta performance na transferência de dados. Em contrapartida, não oferece segurança aos dados que estiverem armazenados nesse tipo de RAID, uma vez que no caso de haver alguma pane em um disco rígido, todo o conteúdo que estiver gravado nos discos serão perdidos. Dessa forma, esse tipo de RAID não é indicado para sistemas que necessitem segurança dos dados. Quanto mais discos houver, mais velocidade poderá ser obtida, pois a gravação dos dados não será mais sequencial.

RAID Nível 1 – Disk mirroring (Espelhamento de discos rígidos)



Esse tipo de RAID é conhecido como do tipo **mirroring** "espelhamento". Os dados a serem armazenados são divididos e gravados simultaneamente em dois ou mais discos, de forma que se tenha uma redundância dos dados que estão sendo gravados, uma vez que cada um dos discos

rígidos terá uma cópia exata do dado que foi gravado no disco.

Em razão da necessidade da replicação dos dados no segundo disco, esse tipo de RAID tem com desvantagem a perda de desempenho durante o processo de gravação dos dados, porém terá como vantagem a redundância, pois caso ocorra alguma falha física em um dos dois discos rígidos, outro terá todos os dados intactos e continuará funcionando sem que haja indisponibilidade do servidor. Esse tipo de RAID também tem um problema relacionado ao desperdício de 50% dos seus recursos, uma vez que um disco rígido "o da cópia" dos dados é desperdiçado enquanto ele não está sendo utilizado.

RAID Nível 0+1 e RAID nível 1 + 0

Existe a possibilidade de fazer uma combinação do RAID nível 0 e do RAID nível 1, formando combinações como RAID 0 + 1 e RAID 1 + 0

Esse nível de RAID pode ser implementado utilizando para isso quatro discos rígidos, e esses quatro discos serão separados em grupo de dois discos e os que fizerem parte de cada um dos grupos serão implementados como RAID 0 (Strip). Depois em cada um dos grupos será implementado o RAID 1, ou seja, serão espelhados. A ideia é que esse tipo de RAID suporte a perda de um dos grupos de disco, pois estará espelhado no outro grupo.

Caso a perda de um disco ocorra nos grupos, todos os dados serão perdidos, pois não haverá mais redundância.

Esse tipo de RAID é feito com a utilização de pelo menos quatro discos rígidos que trabalharão de forma parecida com o sistema anterior, apenas alterando o fato de que cada par de discos será espelhado entre si (RAID 1), garantindo assim, a redundancia dos dados deles. Após esses espelhamentos, os discos serão unidos (Strip) para melhorar o desempenho. Dessa forma, até a metade dos discos pode falhar simultaneamente, sem colocar o conjunto de discos a perder no caso dessa falha, desde que ela não ocorra em dois discos de um espelho qualquer.

Com a utilização desses tipos de RAID's, é possível utilizar o bom rendimento que tem o RAID do nível 0 e a redundância que possibilita RAID do nível 1.

RAID nível 5 – (discos com setores de paridade)

O RAID nível 5 utiliza no mínimo três discos rígidos que permitem aos dados serem gravados entre todos eles, e um pedaço de cada disco é reservado para seja guardada a paridade desses dados. Nesse tipo de RAID, se uma parte de um disco físico que compõe o agrupamento falhar, é possível recriar os dados que estavam nele a partir do dado e da paridade que está distribuída entre os outros discos.

Pelo fato da paridade não ficar destinada a um único disco, e sim a todo o ARRAY de disco, faz com que a gravação de dados seja mais rápida, pois não é necessário acessar um disco específico que guarda a paridade em cada gravação.



A utilização do RAID-5 possibilita ter uma melhor performance na leitura dos dados dos discos do que o RAID 1, porém quando o processo de falha ocorrer, a performance da leitura será prejudicada por causa da necessidade de recuperação dos dados que estão distribuídos entre as paridades dos outros discos.

A utilização de discos Spare possibilita uma imediata reconstrução do sistema como um todo, depois que o processo de Rebuild é realizado.

RAID 1 e RAID 5

Essa é uma solução implementada na maioria dos servidores como uma forma de aumentar o processo de tolerância à falha e a alta disponibilidade, pois são implementados os RAID's nível 1 (Espelhamento) e nível 5 (Paridade).

Para implementar esse tipo de RAID, são utilizados no mínimo cinco discos que são separados em um conjunto de dois discos e outro conjunto de três discos. No primeiro conjunto é implementado o RAID 1, aumentando a segurança em razão do espelhamento, e no segundo é implementado o RAID 5, aumentando a segurança pela paridade dos dados. No primeiro conjunto será criado um disco lógico para ser utilizado no servidor para instalar o seu sistema operacional, enquanto que no segundo conjunto será criado um disco lógico para que sejam guardados os dados. Nesse agrupamento poderiam ser alocados outros discos para melhorar o processo da contingência com a

presença de Spare e discos reservas.

RAID via hardware x software

O processo de RAID pode ser feito de duas maneiras, sendo uma baseada em hardware e a outra baseada em software. Cada uma dessas maneiras possui suas vantagens e desvantagens.

O RAID feito por meio de Hardware é a maneira mais utilizada, pois funciona independentemente do sistema operacional e possui melhor performance, porém apresenta como desvantagem o custo pela necessidade da existência de uma placa controladora, que é responsável pelo gerenciamento dos discos.



O RAID feito por meio de software é criado pelo sistema operacional da máquina em que está instalado e, apesar de ser menos custoso, tem uma performance menor do que o baseado em Hardware, uma vez que o sistema operacional tem de se preocupar em controlar os discos, além das suas atividades básicas. Esse tipo de RAID também vai ter uma dependência do poder de processamento do computador no qual for implementado.

Depois de configurar um sistema de RAID, é importante que seja feito um teste para simular o processo de falha de um dos discos, de forma que se tenha certeza de que o sistema fará o que está sendo esperado dele. Dependendo do tipo de sistema de RAID escolhido, ele deverá funcionar normalmente depois do processo de reinicialização, a não ser no caso do RAID 0, em que o sistema não possui tolerância a falhas.

Observações finais

- É importante que quando ocorrer uma falha no disco, seja feito um backup dos dados antes da troca, como forma de se prevenir de outros incidentes de segurança.
- Também é importante que o equipamento tenha mecanismos de alerta sonoros, luminosos ou de sistema para que o administrador da rede saiba do problema quando ele ocorrer.

Cluster

Nesse tipo de sistema são usados dois servidores completos, e a única função do segundo é assumir a posição do outro em caso de falhas. O sistema opera em um modo conhecido como ativo/passivo, em que o segundo equipamento somente entrará em operação quando houver a necessidade, ou seja, quando o servidor principal tiver saído de operação. Outra forma de operar é construir um sistema em que os servidores trabalhariam de modo a dividir as requisições de serviços constantemente entre si, funcionando em um formato ativo/ativo. Os dois servidores dividiriam a tarefa de disponibilizar um serviço para a rede.

Cada um dos servidores é uma máquina completa e possui pelo menos duas placas de rede para que possam ficar ligados simultaneamente à rede, bem como ficarem ligados entre si. É criada uma rede privada entre esses dispositivos para que o software que controlará o sistema de cluster possa verificar a disponibilidade dos equipamentos, fazer as funções de monitoramento e sincronismo de acesso. Esse processo é necessário para que o segundo servidor, que faz parte do cluster, possa assumir imediatamente a função do primeiro servidor quando for necessário.

Em alguns sistemas que exigem maior tolerância, pode ser colocada outra placa de rede em cada servidor, que estará ligado a outro switch. Isso é feito no caso de ocorrer alguma falha no switch principal e o sistema poder se comunicar com a rede por esse outro caminho.

Também dentro de cada um desses servidores toda a alta disponibilidade deve estar contemplada, por exemplo, a existência de fonte redundante, processadores redundantes, sistema de RAID para os discos etc.

Cada servidor que faz parte de um sistema de cluster é denominado como sendo um nó ou nodo, e cada um deles deve estar interconectado de maneira a formar uma única rede, que também deverá permitir o acréscimo ou a retirada de um equipamento que faz parte de cluster, sem interromper o funcionamento desse sistema.

Os usuários que acessam os recursos que estão dentro de sistema de cluster têm a falsa impressão de que estão acessando dados de um único e não de vários computadores, como pode ser o caso em uma determinada situação. Dessa forma, se consegue uma elevada confiança, distribuição de carga e desempenho no ambiente em que os servidores estão alocados.

Independentemente de qual seja o sistema operacional utilizado, é necessário que seja usado um software para fazer o sistema de gerenciamento e montagem do sistema de cluster. Ele será responsável, entre outras coisas, pelo processo da distribuição do processamento e identificação dos possíveis erros e defeitos que possam ocorrer durante o seu funcionamento.

Tipos de clusters

- **Cluster de alta disponibilidade (ha "high-availability")** – nesse tipo de cluster são utilizados vários equipamentos "servidores" para manter um determinado recurso ou serviço sempre disponível. Quando um equipamento "ativo", que faz parte desse sistema, apresenta algum tipo de falha, o outro equipamento "passivo" assume imediatamente o seu lugar de forma a deixar o acesso ao recurso novamente disponível.



- **Cluster para balanceamento de cargas (load balancing)** – esse tipo de cluster é constituído por equipamentos "servidores" que rodam, "executam" um mesmo serviço de forma que a resposta de solicitação a alguns desses serviços possa ser distribuída entre esses equipamentos "servidores". Desse modo, ocorre um balanceamento da carga que é distribuída de maneira equilibrada aos nós do cluster. Nesse tipo de cluster, se um servidor vier a falhar (failover*), o acesso às aplicações ou serviços estarão disponíveis por meio do outro servidor, que faz parte do cluster. Todos os equipamentos são responsáveis por controlar as solicitações de acessos aos serviços. *(O sistema age automaticamente sem que seja necessária a intervenção humana).
- **Combinação HA & load balancing** – nesse tipo de cluster ocorre a combinação dos dois tipos de cluster descritos anteriormente, aumentando dessa forma a disponibilidade, distribuição dos acessos e a escalabilidade de serviços utilizados pelos clientes.

Agora que você já estudou esta aula, resolva os exercícios e verifique seu conhecimento. Caso fique alguma dúvida, leve a questão ao Fórum e divida com seus colegas e professor.



objeto disponível apenas na versão web



objeto disponível apenas na versão web

Estimule seu raciocínio com o jogo da forca, clique no botão a seguir.



objeto disponível apenas na versão web

Referências

FERREIRA, Fernando Nicolau Freitas; ARAÚJO, Márcio Tadeu de. *Política de segurança da informação: guia prático para elaboração e implementação*. Rio de Janeiro: Ciência Moderna, 2006.

MAGALHÃES, Ivan Luizio; PINHEIRO, Walfrido Brito. *Gerenciamento de serviços de TI na prática: uma abordagem com base na ITIL*. São Paulo: Novatec, 2007.

WADLOW, Thomas A. *Segurança de redes: projeto e gerenciamento de redes seguras*. Rio de Janeiro: Campus, 2001.