RAID x SQL Server

Escolhendo a configuração certa para os arquivos certos

**Rafael Dontal Gonçalez**

[rafaeldontal@gmail.com](mailto:rafaeldontal@gmail.com)



Formado em Engenharia da Computação e com 5 anos de experiência como DBA. Após ler o SQL Server Internals (que vale como uma pós graduação), trabalha como consultor SQL Server para Performance e Tuning de ambientes de banco de dados.

|  |
| --- |
| **De que se trata o artigo** |
| O artigo aborda conceitos da montagem de arrays de discos independentes (RAID), e procura relacioná-los ao SQL Server, apontando caracteristicas de performance e de segurança e as vantagens e desvantagens de cada tipo de construção, fazendo o relacionamento dessas características com cenários comuns nas instâncias do SQL Server. |
| **Em que situação o tema é útil** |
| Na preparação de um ambiente de banco de dados, principalmente SQL Server de alta disponibilidade e alto desempenho, auxiliando e orientando DBAs desde um nível básico até os mais avançados. |

|  |
| --- |
| **RAID X SQL Server**    Sempre que está sendo planejado um ambiente de banco de dados, uma das questões mais importantes para o desempenho e segurança de todo o ambiente paira no ar: Qual configuração de RAID escolher? Para respondermos a esta pergunta, precisamos fazer algumas outras: Afinal, o que é RAID? Como ele funciona? Qual nível de RAID melhor se aplica para cada característica do SQL Server? Como o banco de dados trabalha com cada tipo? |

Na maioria dos bancos de dados OLTP que utilizam Hard Disks (Discos Rígidos) magnéticos tradicionais, um dos fatores limitantes para o operações de leitura e escrita é a latência cabeça do disco; o tempo que leva a cabeça para se mover fisicamente através do disco para encontrar as informações. Portanto, distribuindo os dados entre um grande número de discos menores agrupados, geralmente leva a um desempenho muito melhor do que usar alguns discos grandes e individuais. Por isso utilizamos o RAID.

Para entender e escolher o melhor tipo de RAID para o SQL Server, é fundamental conhecermos o que é e como é que funciona este recurso para discos. A seguir, vamos apresentál-lo e explicar alguns de seus conceitos.

RAID é a sigla para Redundant Array of Independent Disc, que traduzindo para o portugues significa Conjunto Redundante de Discos Independentes. Ou seja, esta expressão indica justamente a utilização de Hard Disks padronizados e baratos como "blocos de montagem" para a criação de sistemas que se comportam como um único disco, maior, mais rápido e/ou confiável do que suas peças individuais.

**Termos importantes em RAID**

DIVISÃO (STRIPPING): Quando falamos em STRIPPING, estamos dizendo os dados serão separados e distribuídos para mais de um disco;

ESPELHAMENTO (MIRRORING): No MIRRORING há uma cópia idêntica dos dados para mais de um disco;

PARIDADE: Os dados redundantes são armazenados usando algum algoritmo para permitir a detecção e possível reparação (Conhecido como “Tolerância a falha”).

CONTROLADORA: A placa controladora é um dispositivo que fica entre o sistema host e o sistema de armazenamento e permite que os dois sistemas se comuniquem.

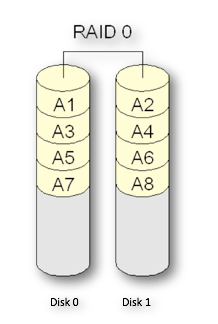
**Tipos de Implementações de RAID**

Implementação via software: O controle de gravação nos discos é feito pelo Sistema operacional, através da controladora de discos, sem a necessidade de ter uma controladora RAID, tornando-se mais barato. Porém, esse processo se torna complexo e demanda maior consumo do processador, concorrendo com o serviço de banco de dados, o que provavelmente prejudicará o desempenho da sua base.

Implementação via hardware: A mais recomendada e utilizada em um ambiente de banco de dados de alto desempenho. Requer pelo menos uma controladora especialmente dedicada para isso, o que encarece esse processo. Tem como vantagem a não utilização de recursos do processador, além de poder suportar vários sistemas operacionais. Suporta também "hot-swapping" (seria a “troca a quente”, em que não precisa parar o ambiente todo para a troca de um disco corrompido, por exemplo), evitando assim a indisponibilidade.

**Principais tipos de RAID**

**RAID 0**



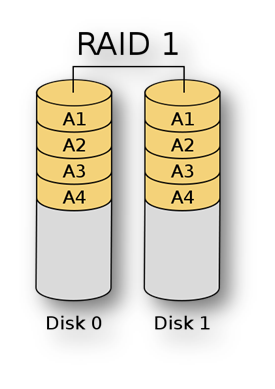
O armazenamento do nível RAID O é feita de forma distribuída (Stripping), necessitando de pelo menos 2 unidades, o RAID 0 distribui dados em cada disco. As capacidades disponíveis de cada disco são adicionadas em conjunto para que um volume lógico seja adicionado ao computador.

Vantagens: Performance de I/O é melhorada; Maior capacidade de armazenamento.

Desvantagens: Não tolera falhas;

Recomendações: Nunca deve ser usado em ambientes críticos; Em sistemas/aplicações onde ocorrem muita troca de dados e onde a possível perda de dados não será crucial para o negócio.

**RAID 1**



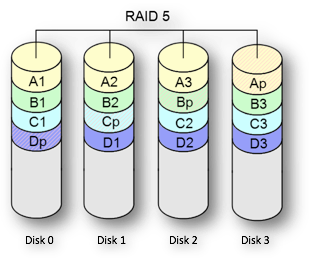
Trabalha com números pares de discos, pois realiza um mirroring (espelhamento) dos dados, ou seja, quando um arquivo é gravado no disco, o armazenamento ocorre de forma simultânea nos discos, tornando-se um sistema de armazenamento altamente confiável. Caso ocorra falha em um dos discos, a informação ainda estará disponível no disco “espelhado”, mas não podemos confundir redundância com backup, no caso de deletarmos o conteúdo do primeiro disco, automaticamente os dados do disco-espelho também serão deletados.

Vantagens: As leituras são mais rápidas; A falha de dados de um disco não causa perda de dados.

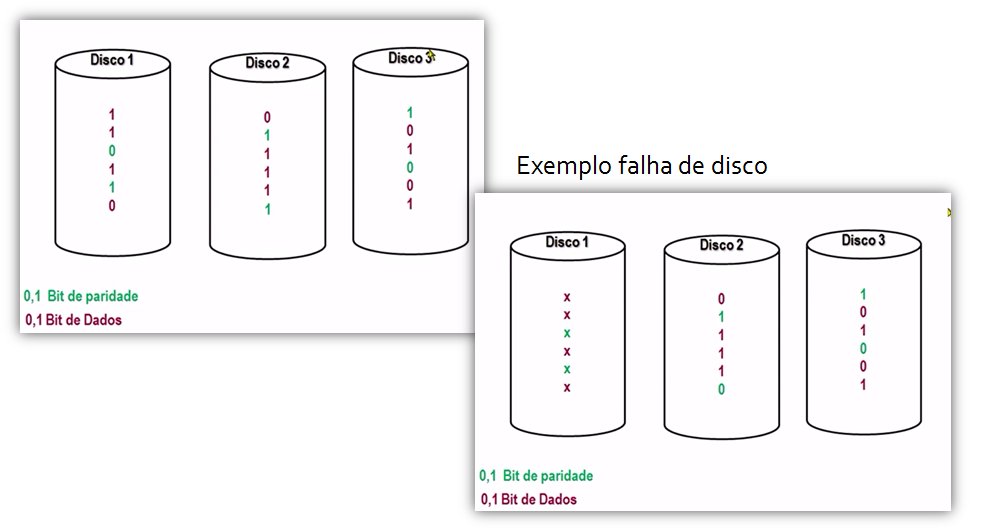
Desvantagens: O custo é mais elevado; Performance na escrita é prejudicada.

Recomendações: Para qualquer aplicação que requer alta disponibilidade; Em sistemas/aplicações mais leituras do que escrita.

**RAID 5**



Neste nível, faz-se o stripe em blocos de dados através dos discos, porém o espaço equivalente a um destes discos é dedicado apenas para fazer a paridade entre os outros discos. São utilizados no mínimo 3 discos e as informações de paridade são gravadas em cada disco (de forma distribuída no conjunto) e, se um deles falhar, será possível reconstruir a informação, como vemos na imagem abaixo:



O bit de paridade indica se a quantidade de 1s nos outros discos naquela linha de dados é par ou ímpar. Tendo essa informação, é possível reconstruir o dado perdido equivalente a um disco.

Vantagens: Alta velocidade de leitura; Perda do espaço de apenas 1 disco; Maior rapidez para o tratamento e controle de erros.

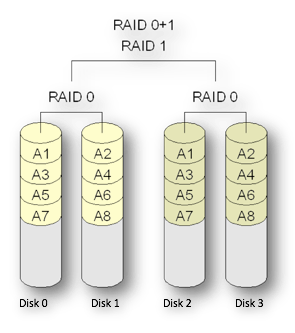
Desvantagens: Performance baixa na escrita; Falha no disco pode ter um médio impacto na performance.

Recomendações: Servidores de arquivos, aplicações, web e-mail.

**RAID 01 ou 0+1**

Os dados são segmentados em um grupo de discos, utilizando o conceito de RAID 0. Depois esse grupo é espelhado, utilizando-se o conceito de RAID 1.

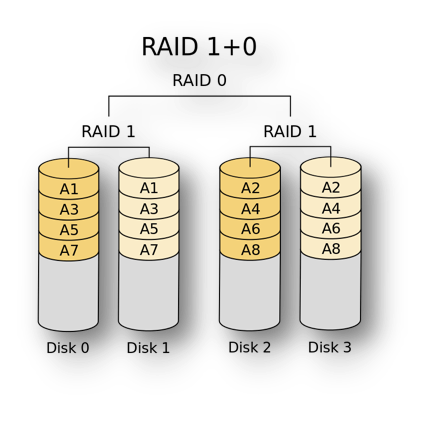
Como na imagem abaixo:



Este sistema apresenta performance de leitura e escrita excelentes.

**RAID 10 ou 1+0**

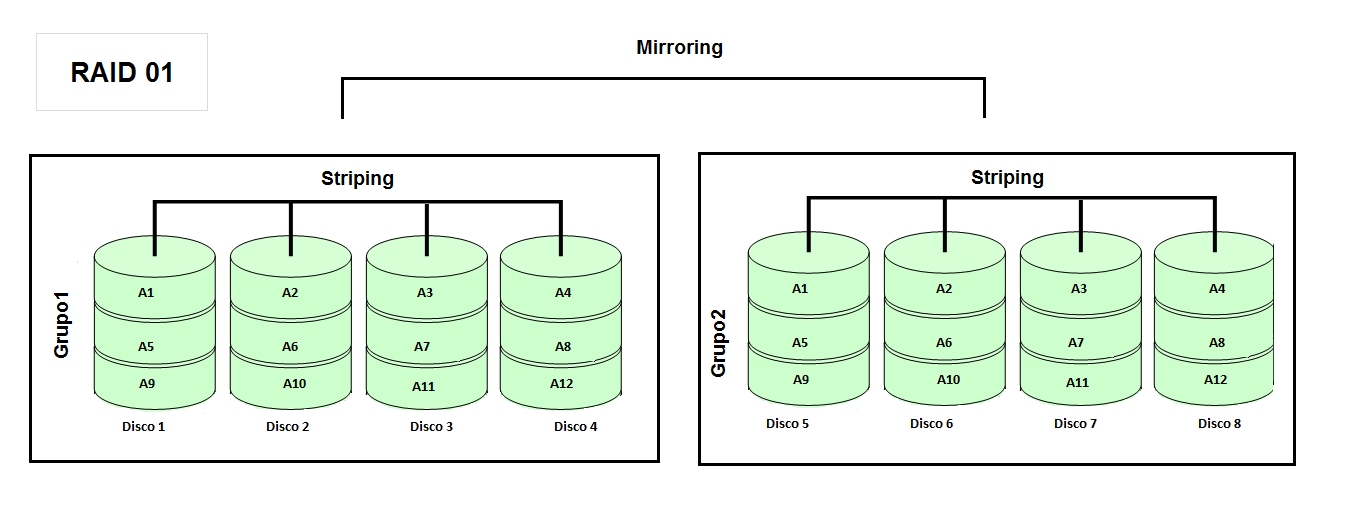
Ao contrário do RAID 01, os dados são segmentados através de grupos de discos espelhados, isto é, os dados são primeiro espelhados em duplas, para depois serem segmentados.

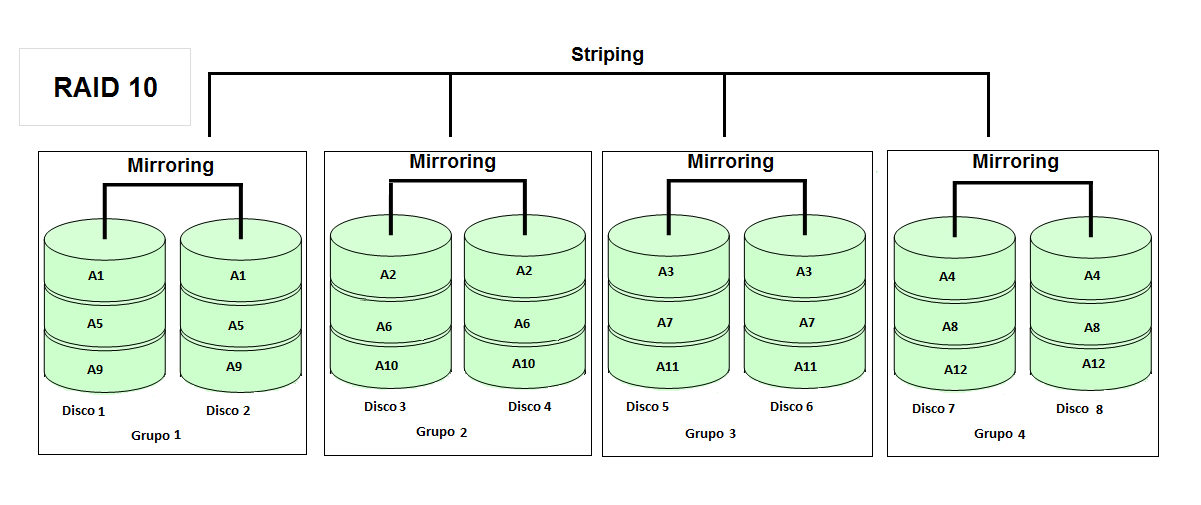


Performance de leitura e escrita bem semelhantes ao RAID 01.

**Importante**: Algumas vezes são confundidos os níveis 10 e 01, dizendo que ambos são iguais. Realmente na questão desempenho, tanto para leitura quanto para escrita, há uma equivalência entre eles. Notamos as diferenças quando o quesito é segurança.

Comparando os desenhos a seguir, imaginamos que seja corrompido apenas um disco.





Para o RAID 01, se o Disco 1 do Grupo 1 for corrompido, todo esse grupo ficará inoperante, pois foi quebrada a sequencia de stripping, jogando todo o processamento para o Grupo 2 e transformando o sistema todo em um RAID 1, ou seja, ainda terá um bom desempenho em leitura e escrita, mas não será mais tolerante a falha alguma.

Já no caso de perdermos o mesmo Disco 1 no RAID 10, note que apenas um dos grupos será afetado, e este grupo afetado está espelhado, ou seja, o ambiente de discos continua igual, o que possibilita a perda de um disco de cada grupo até que seja feita a reposição destes drives.

Outra diferença é na velocidade de recuperação, porque caso ocorra uma falha de disco, no sistema RAID 1+0 será necessário apenas reespelhar um disco, ao contrário do sistema RAID 0+1 que será necessário espelhar todo um conjunto segmentado.

**Comportamento dos arquivos do SQL Server**

Agora que já sabemos como trabalham os principais tipos de RAID, vamos relembrar o comportamento do SQL Server com seus distintos tipos de arquivos.

Sempre que iniciamos uma nova instalação de banco de dados SQL Server, como boa prática, solicitamos ao menos 5 drives diferentes: Sistema Operacional, Arquivos de dados, Arquivos de Log, Tempdb e Backup.

Os arquivos de Sistema Operacional não tem uma taxa de Entrada e Saída muito altas. Nele são feitas leitura essenciais para o funcionamento de todo o servidor, mas essas leituras não são realizadas em uma escala macro ou com concorrência, como ocorre com os arquivos de dados, por exemplo.

Os tipos de solicitações (leitura ou escrita) nos Arquivos de Dados do SQL Server variam muito, dependendo das regras de negócio da empresa. Apesar disso, na grande maioria dos casos, a leitura é sempre maior do que a escrita, ou seja, os dados são escritos uma vez e muitas outras vezes esses mesmos dados serão lidos.

Com excessão de alguns tipos de transações não logadas (que não são tão comuns, mas que deixaremos para falar em outro artigo), o predomínio das transações escritas nos Arquivos de Dados é escrita antes disso nos Arquivos de Log. Já as leituras no Log são feitas poucas vezes, apenas por processos internos do SQL Server, isto é, nenhuma aplicação que precisa ler dados, lê os dados dos Arquivos de Log.

No caso dos discos para Tempdb, a quantidade de leituras e escritas é praticamente a mesma, e, devido a quantidade e frequência de acessos, deve ter uma performance excelente para os dois processos.

No Backup, é claro, faremos muitas escritas e raramente (assim esperamos) haverá leitura.

**A Escolha do RAID ideal**

RAID 1 é excelente para prover redundancia para qualquer tipo de arquivo. Mas ele não melhora o desempenho de nenhum deles. Pelo contrário, utilizando o RAID 1 temos uma pequena piora em escrita, pois ao escrever os dados estes devem ser espelhados. Portando não seria a escolha ideal principalmente para Logs e Tempdb. Talvez pudesse ser usado para Data Files históricos de poucas leituras e pouca concorrência, algo que não necessitasse um alto desempenho. Mas vemos que a aplicação ideal para o RAID 1 são os arquivos de Sistema Operacional, em que a tolerância a falhas é necessária, mas não precisamos de algo rendimento elevado.

Muitos DBAs e empresas optam por utilizar o RAID 0 para os arquivos de Tempdb, devido ao seu alto desempenho de IO, contudo devemos lembrar que RAID 0 para tempdb é um risco: se alguma unidade falhar, então o TempDB falha e o serviço do SQL Server é desligado. Sabemos com certeza que não há dados importantes lá, mas também sabemos que o Tempdb é essencial para a operação SQL e se ele estiver em RAID 0 e uma unidade falhar, SQL Server não poderá ser iniciado até que esta unidade não seja substituída (que pode levar de minutos a semanas) ou até que alguém consiga iniciar o SQL Server sem TempDB e realoque tempdb para alguma outra unidade. O custo e tempo perdido deste processo podem ser muito grandes, então o emprego deste tipo de RAID deve ser avaliado pela equipe técnica com muita cautela. Particularmente, não recomendo RAID 0 em nenhum momento para meus clientes.

Consideramos o RAID 5 como a melhor relação custo X benefínio. Pensando em um servidor de alta necessidade de desempenho, RAID 5 é terrível para os logs e disco de backup, pois, como vimos, para toda escrita deve se fazer o calculo do bit de paridade, causando uma sobrecarga no processo. Supor o uso de RAID 5 para o drive de Sistema Operacional soa como um desperdício de dinheiro, teremos muita redundancia e desempenho para um disco que requer muito pouco disso. A utilização mais comum deste nível de RAID é para arquivos de dados. Teremos uma performance elevada de leitura e o custo será relativamente baixo em relação ao RAID 10.

Embora seja o mais caro, nenhum outro nível melhor desempenho e segurança do que o RAID 10. Ele tem as vantagens de ambos RAID 0 e RAID 1 e usa todas as unidades do array para obter taxas de I / O mais altas. Quanto mais unidades maior será o rendimento. É recomendado para a maior parte das aplicações.

**Conclusão**

Sempre que me perguntam qual tipo de RAID devem escolher para o banco de dados SQL Server, minha resposta é rápida: RAID 10. Por quê? Melhor desempenho e segurança do que qualquer outro.

Porém, sempre que você for planejar o ambiente da sua empresa ou cliente, deve avaliar com cautela as necessidades do sistema e o budget disponível. É possível montar algo que atenda às necessidades, mas que os custos não sejam tão altos.

A tabela abaixo faz um resumo geral e pode ajudar no seu planejamento:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Nível RAID** | **Tolerância a Falhas** | **Performance de leitura** | **Performance de escrita** | | **Custo da escrita ( 1 a 4)** | **Custo por perda de espaço em disco** |
| 0 | Nenhuma | Bom | Excelente | 1 | | Nenhum |
| 1 | Bom | Bom | Bom | 2 | | Médio |
| 5 | Regular | Bom | Ruim | 4 | | Médio/Baixo |
| 1+0 | Excelente | Excelente | Excelente | 2 | | Alto |

E então? O que acharam? É assim que planejam o seu ambiente de banco de dados? Ficarei feliz em saber a opinião de vocês.