

RAID



Conceitos de RAID

- Um simples Disco Rígido físico ou múltiplos discos em um gabinete externo sem funcionalidades RAID é geralmente chamado de
 - JBOD (Just A Bunch Of Disks)
 - SAS (Server Attached Storage) podendo estar dentro do servidor.
- O projeto e pesquisa RAID foi desenvolvido pela Universidade de Berkley (CA-USA) nos anos 80.
 - Tema Original = Redundant Array of Inexpensive Disks
 - Atualmente = Redundant Array of Independant Disks
- RAID combina um número independente de discos juntos proporcionando aumento do throughput/performance no acesso e confiabilidade dos dados
 - Níveis RAID diferentes proporcionam uma variedade de misturas de performance e tolerância a falhas

RAID

O que é RAID?

- Dois ou mais discos rígidos independentes.
- Aparece no servidor como um único Grande Disco Rígido .
- Sistema Operacional totalmente independente.
- Dispositivo com alta confiabilidade e redundância.

Aumento da Disponibilidade dos Dados

- Rápido acesso aos dados (multiple drives).
- Segurança oferecida pela redundância/paridade.

Gerenciamento Simplificado

- Partição de um único Drive para gerenciamento/partição.

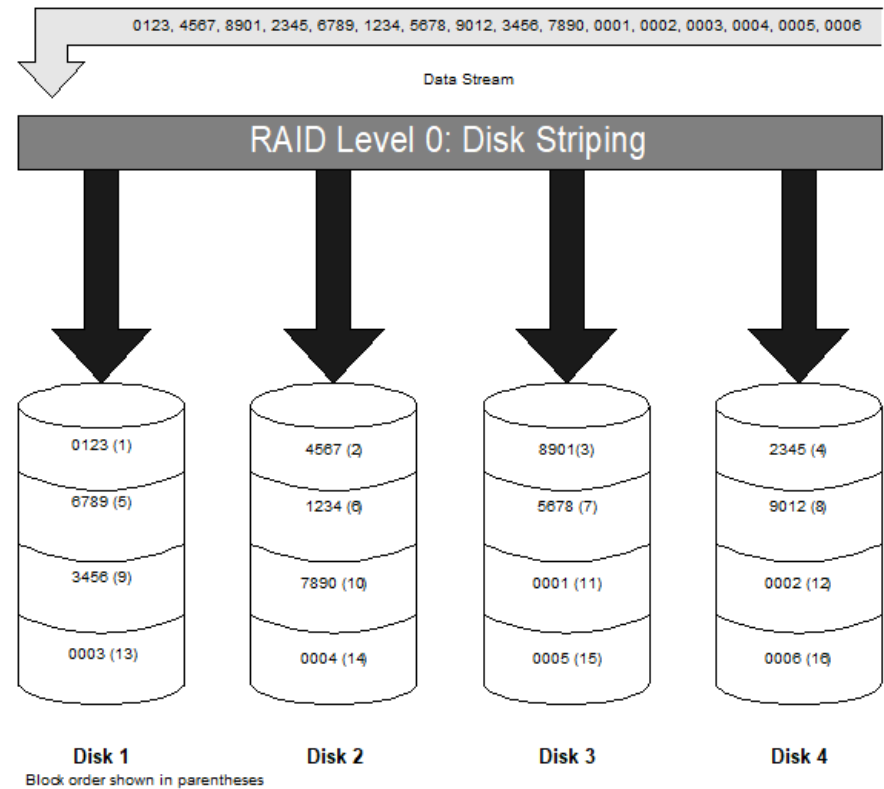
Componentes Hot-Swap

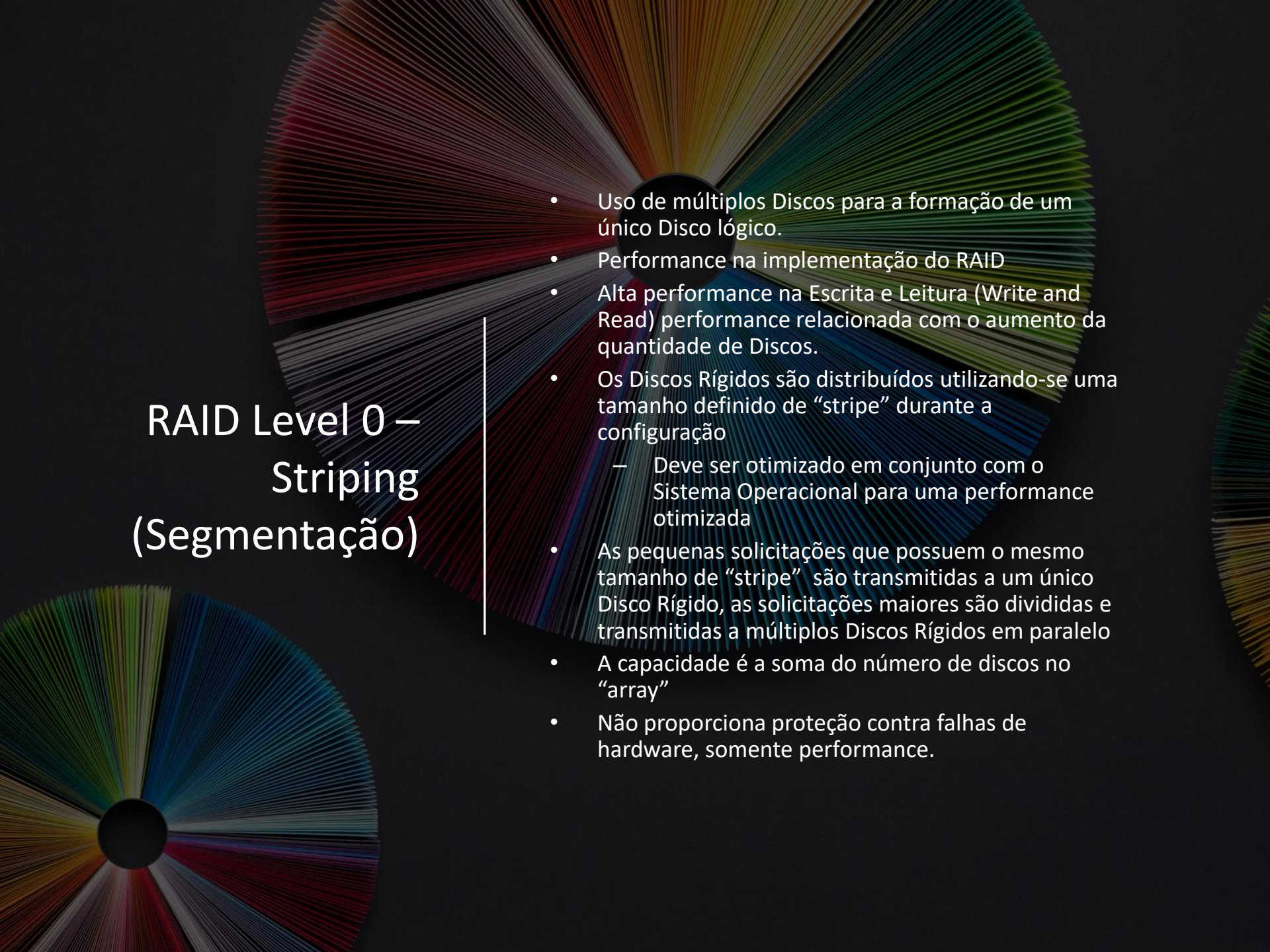
- Tanto para um ou vários drives.
- Fontes de Alimentação e Ventilação, também.

O que RAID não é

- Suporte a aplicações de Disaster Recovery Físicos
- (Tape Drivers).

RAID Level 0



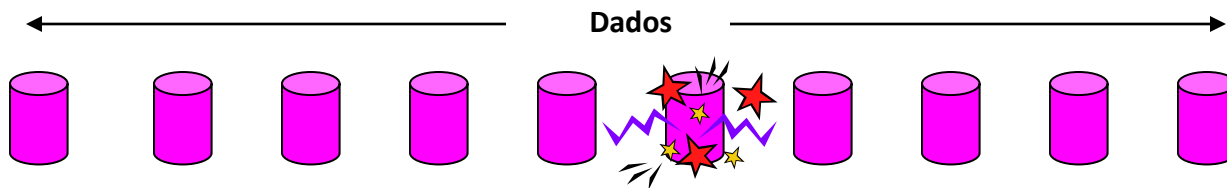


RAID Level 0 – Striping (Segmentação)

- Uso de múltiplos Discos para a formação de um único Disco lógico.
- Performance na implementação do RAID
- Alta performance na Escrita e Leitura (Write and Read) performance relacionada com o aumento da quantidade de Discos.
- Os Discos Rígidos são distribuídos utilizando-se uma tamanho definido de “stripe” durante a configuração
 - Deve ser otimizado em conjunto com o Sistema Operacional para uma performance otimizada
- As pequenas solicitações que possuem o mesmo tamanho de “stripe” são transmitidas a um único Disco Rígido, as solicitações maiores são divididas e transmitidas a múltiplos Discos Rígidos em paralelo
- A capacidade é a soma do número de discos no “array”
- Não proporciona proteção contra falhas de hardware, somente performance.

Resumo Nível: RAID 0

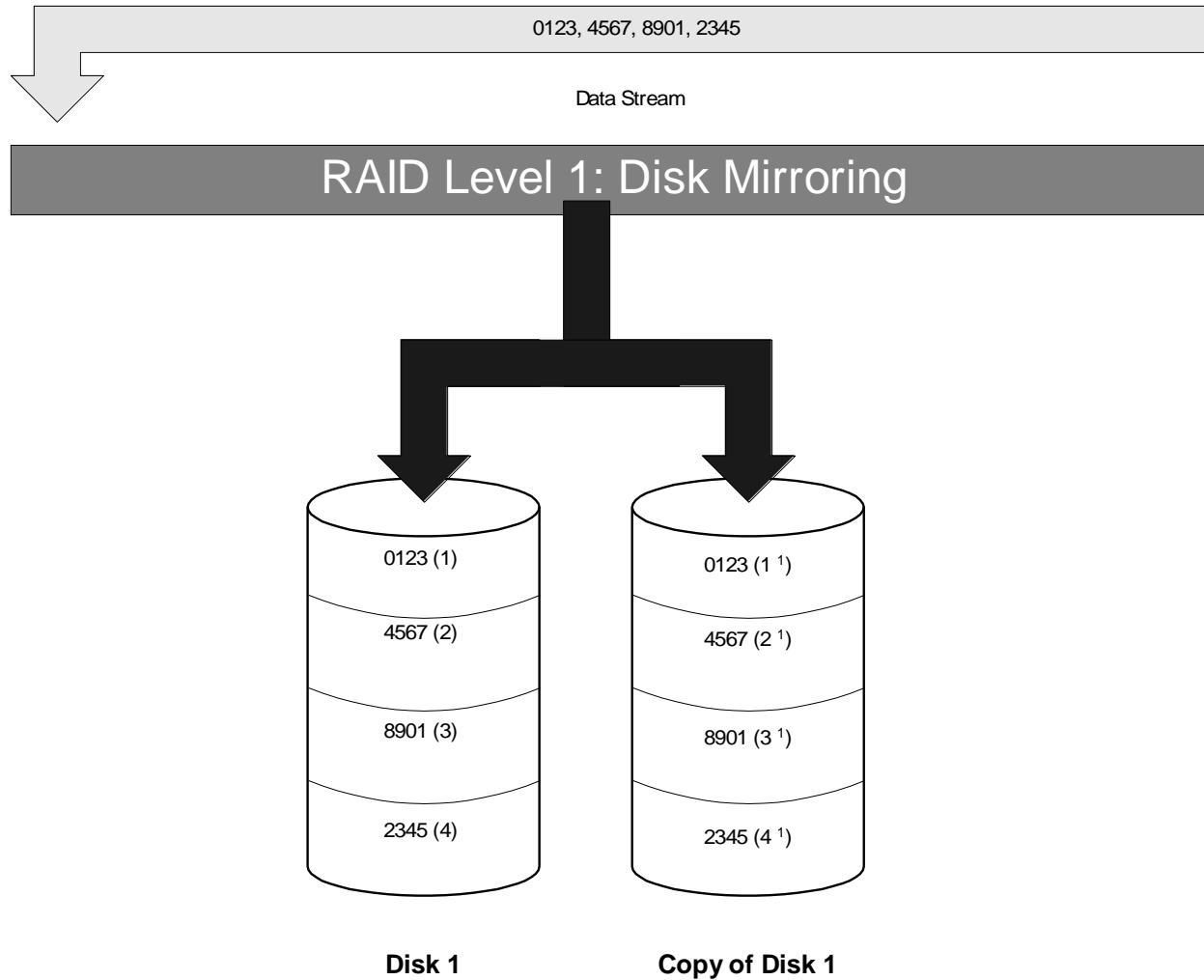
- RAID-0, “Striping”, todos os Discos estão disponíveis para dados, Inseguro.
- Utilizado para armazenamento temporário
- Qualquer Disco que falhar causa a perda dos dados



10 = Total de Discos
10 = Capacidade Usável
0 = Redundância de Discos
10 = Performance Leitura
10 = Performance Escrita

Todos os Dados
Foram Perdidos

RAID Level 1



Block order shown in parentheses

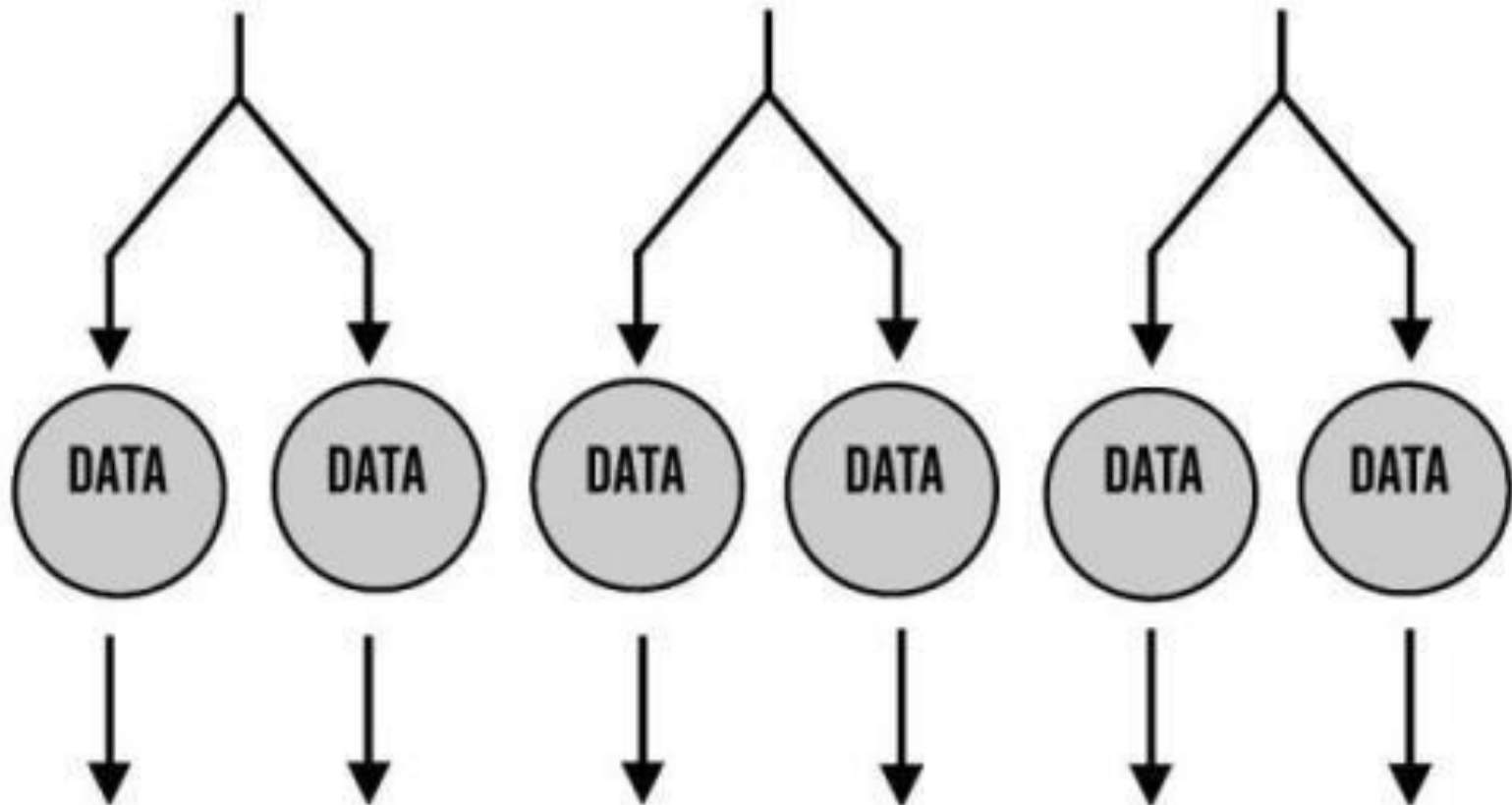
RAID Level 1 provides fully redundant disk mirroring

RAID Level 1 – Mirroring (Espelhamento)

- O RAID 1 proporciona alto nível de tolerância a falhas
 - Cada solicitação de I/O é espelhada em um segundo Disco Rígido
- O RAID 1 trabalha com múltiplos de dois Discos Rígidos- o set primário e o set espelhado – dobrando também o custo por GB da solução
 - Proporciona o mesmo throughput durante a escrita
 - Oferece mais performance durante a leitura (read) – A controladora RAID executa o mesmo comando de leitura em cada Disco – assim que um dos discos disponibiliza os dados, ele é automaticamente transferido para o servidor e o próximo comando de leitura pode ser processado.
- Oferece proteção contra falhas nos discos com taxa de 1 para 1
 - Assim que um disco falha, automaticamente o espelho assume, porém caso ocorra a falha no espelho não existe mais segurança
 - Pode proteger contra desastre naturais/físicos porém o espelho deve estar instalado remotamente (em outro local físico) acarretando em um custo adicional \$\$.
 - Não protege contra vírus digital ou acidentes/delete intencional, já que 100% do disco está sincronizado com o outro
- Implementação de Alto Custo
 - Controladora RAID +
 - Custo por GB de chega a ser o dobro se compararmos com uma unidade simples.

RAID 1

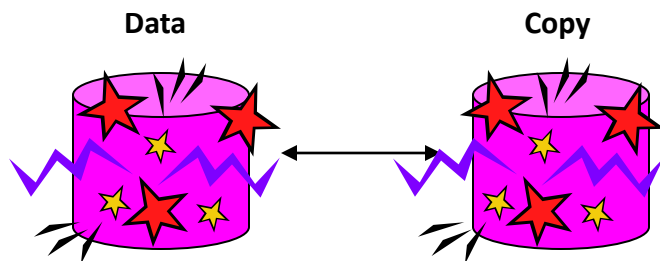
Dados duplicados são escritos em pares de unidades.



Leituras podem ocorrer simultaneamente em todas as unidades.

Resumo Nível: RAID 1

- RAID-1, “Mirroring” rápido, 50% dos Discos estão disponíveis para dados, extremamente seguro.
- Utilizados para informações importantes
- Qualquer Disco espelhado pode falhar, e os dados permanecem salvos



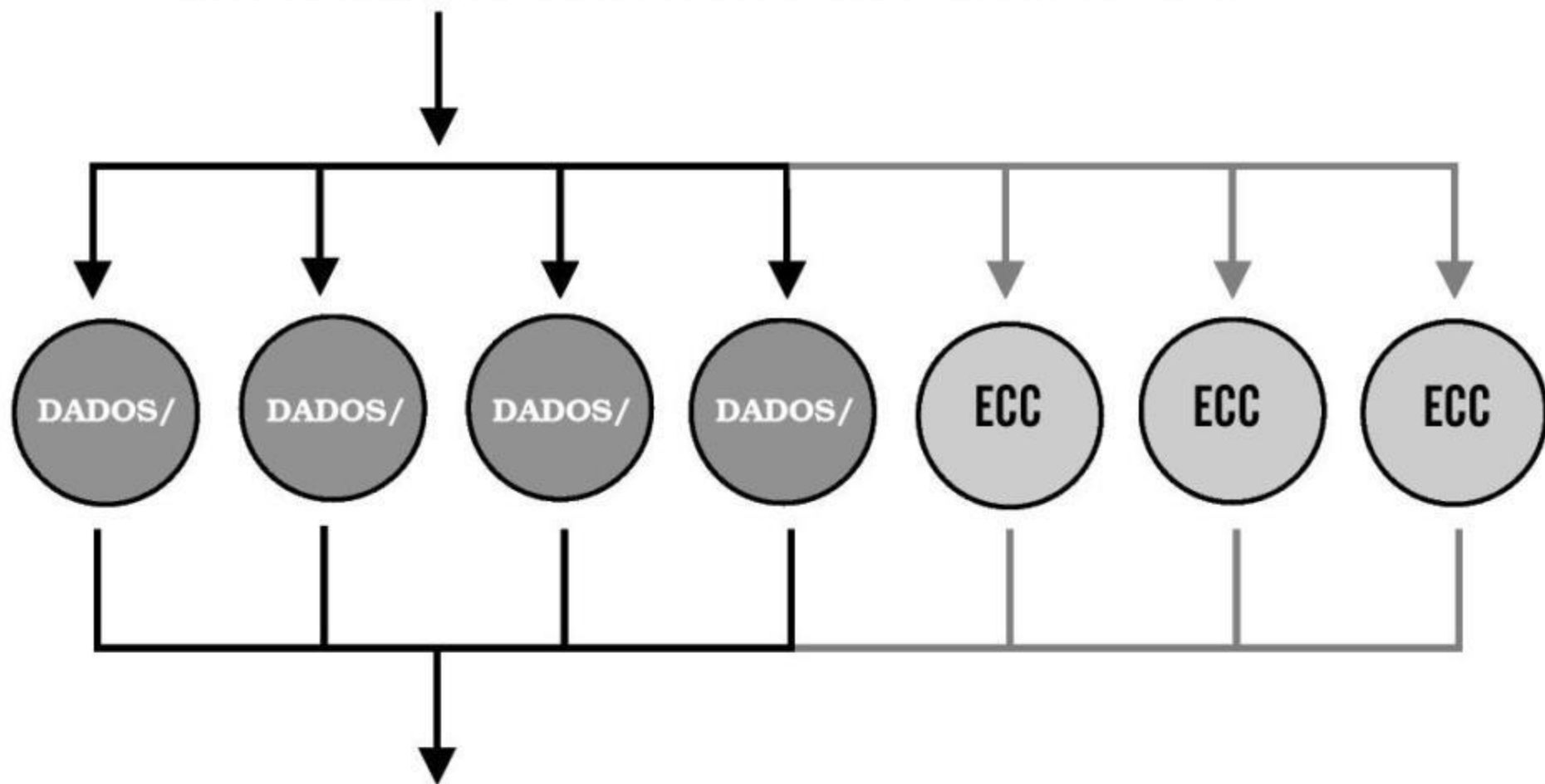
2 = Total de Discos
1 = Capacidade Usável
1 = Redundância de Discos
2 = Performance Leitura
1 = Performance Escrita

Os dados permanecem Salvos

Todos os Dados
Foram Perdidos

RAID 2

Cada operação de escrita ocorre em todas as unidades.



Cada operação de leitura ocorre em todas as unidades.

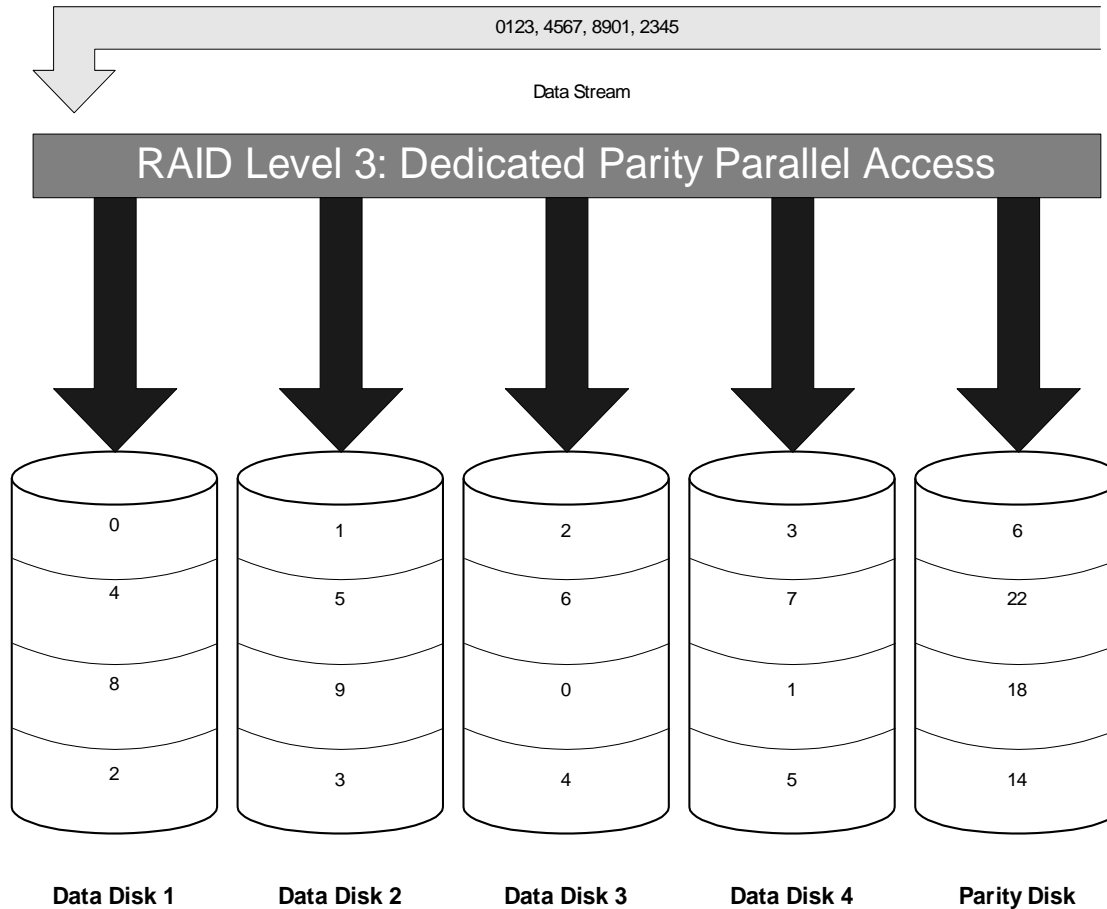
Resumo

Nível:

RAID 2

- 3 ou + discos
- Já obsoletos pelas novas tecnologias de disco, que já possuem ECC internamente
- Não é comercialmente viável
- Divisão dos dados em bits
- Exige sincronia de discos

RAID Level 3



Byte level striping shown here

RAID Level 3 uses separate parity disk

Paridade e ECC

- Utiliza metodologia de reconstrução dos dados perdidos, adicionando-se informações adicionais, comumente chamada de “overhead”
- Utiliza cálculos Matemáticos Polinomiais e operações inversas
- É utilizada uma equação de 5th Grau para encontrar a soma e o resultado.

$$(5+10+\underline{?}+2+7 = 27)$$

- O método de detecção de erros com paridade em computadores funciona apenas para a detecção do erro.

Método de Paridade no barramento SCSI

Binary Number	128	64	32	16	8	4	2	1	Parity	Parity Number
2	0	0	0	0	0	0	1	0	0	1
6	0	0	0	0	0	1	1	0	1	3
33	0	0	1	0	0	0	0	0	1	2
33	0	0	1	0	0	0	0	1	1	3

Bit Error

Parity Sum Error

RAID Level 3 – Acesso Sincronizado com disco dedicado de paridade

- O Uso de mais Discos possibilita um aumento do throughput
 - Um único de Disco de paridade pode ser um problema de gargalo na performance
- Proteção contra falha de Disco com razão de 1 para vários
 - A Performance é degradada durante o acesso, e especificamente durante a reconstrução em caso de falha (rebuild) de um disco rígido (demorando horas)
 - Não oferece proteção contra desastre físicos
 - Não protege contra vírus digital ou acidentes/delete intencional, já que os dados são protegidos por um esquema de paridade ECC (com esquema e atualização em tempo real)
- Implementação de Alto Custo
 - Custo da Controladora RAID +
 - Discos Rígidos com sincronização de spindle
 - 1 Disco Rígido adicional para Paridade
 - Implementado como um sistema RAID
 - (gabinete stand alone)



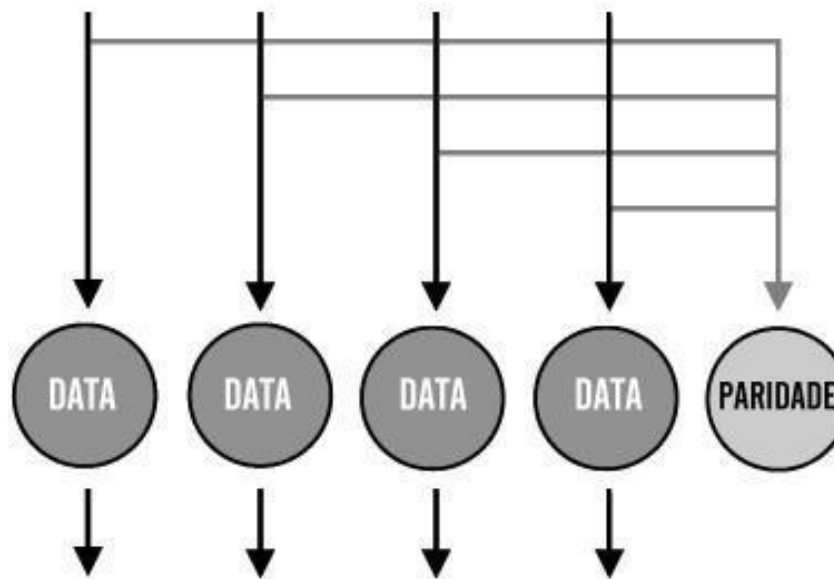
Resumo Nível: RAID 3

- 3 ou + discos
- A fim de evitar o atraso em razão da latência rotacional, o RAID 3 exige que todos os eixos das unidades de disco estejam sincronizados
- Leitura e escrita rápidas
- Disco de paridade é o gargalo

RAID Level 4

RAID 4

Toda escrita precisa atualizar a unidade de paridade dedicada.



Leituras podem ocorrer simultaneamente em todas as unidades de dados.

RAID Level 4

3 ou + discos

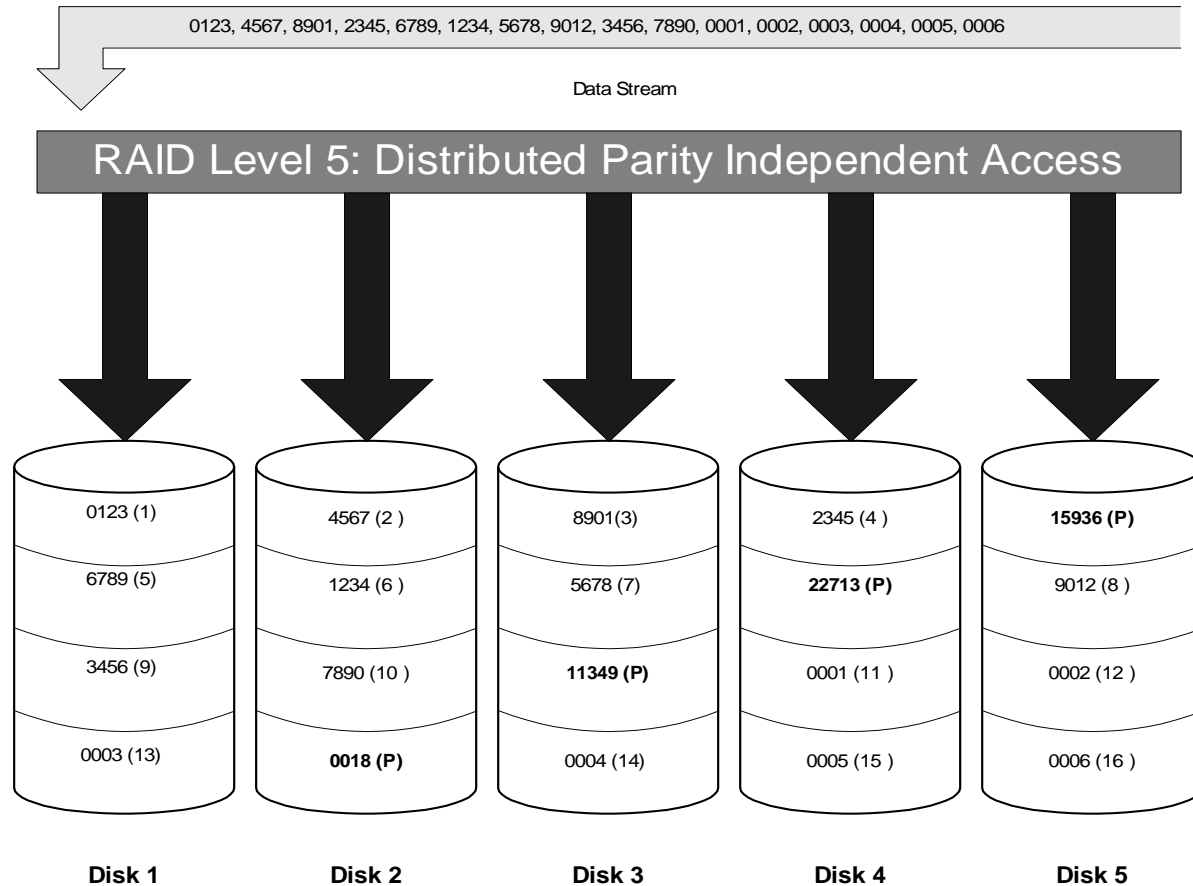
Divisão dos dados em blocos

Disco exclusivo de paridade

Leitura rápida, escrita lenta

Tecnologia não mais usada por haver melhores para o mesmo fim.

RAID Level 5



Block order shown in parentheses
(P) = Calculated parity value

RAID Level 5 distributes parity across all drives

RAID Level 5 – Acesso Independente com paridade distribuída

- O uso de mais discos possibilita o aumento da performance RAID 5, é geralmente utilizado em ambientes OLTP
- Proteção contra falha de Disco com razão de 1 para vários
 - Qualquer disco rígido (1) pode falhar e mesmo assim o sistema permanece intacto
 - Não oferece proteção contra desastre físicos
 - Não protege contra vírus digital ou acidentes/delete intencional, já que os dados são protegidos por um esquema de paridade ECC (com esquema e atualização em tempo real)
- Implementação de Alto Custo
 - Capacidade com perda pela Paridade, o crescimento dos dados aumenta o custo por GB
 - Custo da Controladora RAID +
 - Implementado como um sistema RAID
 - (gabinete stand alone)

RAID 5

- É provavelmente, o nível mais popular usado em servidores hoje.
- Tem uma junção de desempenho e uso eficiente do espaço de armazenamento, sacrificando apenas uma fração do espaço total.
- Diferente do método para armazenar cópias completas espelhadas, como no caso do RAID 1, no RAID 5 a redundância é distribuída entre todas as unidades utilizando um sistema de paridade que mantém a integridade dos dados.
- Os arquivos são divididos em fragmentos e para cada grupo destes fragmentos, é gerado um outro fragmento adicional, contendo os códigos de paridade.
- Uma operação que, ao invés de reservar um HD inteiro para a tarefa, os códigos de correção são espalhados por todos os discos. Desta forma, é possível gravar dados simultaneamente em todos os HDs, melhorando o desempenho e graças aos bits de paridade, é possível recuperar os dados de qualquer um dos HDs que eventualmente falhe, pois o sistema pode continuar funcionando normalmente, mesmo sem um dos HDs.



Vantagens e Desvantagens

Vantagens:

- O uso eficiente do espaço total combinado.
- Conta com uma leitura rápida.
Tolerância a falhas: Se você perder uma unidade seus dados ainda estarão seguros.

Desvantagens:

- Speed: RAID 5 não é tão rápido como RAID 0 ou 1
- Se você perder mais do que uma unidade ao mesmo tempo, seus dados serão perdidos.

Quando usar:

- Este nível de RAID é geralmente a melhor para armazenamento de dados, uma vez que faz um uso eficiente do espaço total combinado e fornece redundância de dados.

Resumo Nível: RAID 5

- RAID-5, rápido, todos menos um disco está disponível para dados, Segurança.
- Utilizado com Armazenamento Primário
- Caso um Disco falhe "OK", O segundo Disco ocasiona a perda total dos dados!

10 = Total de Discos
8 = Capacidade Usável
1 = Redundância de Disco
8 = Performance Leitura
8 = Performance Escrita

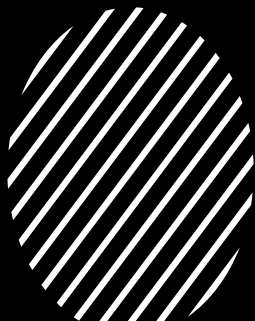


Todos os Dados
Foram Perdidos

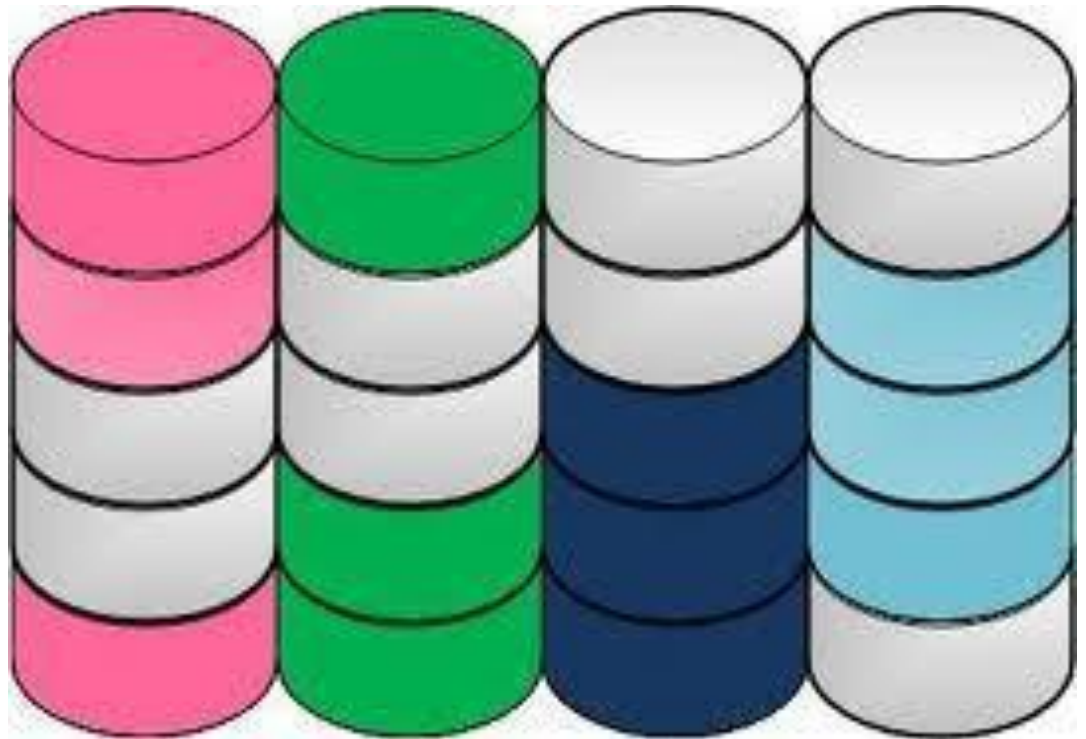


RAID Level 6

- 4 ou + discos
- Aproveitamento: $(n-2)/n$ %
- Admite quebra de até 2 discos
- Pouco implementado
- Overhead na remontagem



RAID Level 6



RAID 6

- É um padrão relativamente novo, suportado por apenas algumas controladoras. É semelhante ao RAID 5, porém usa o dobro de bits de paridade, **garantindo a integridade dos dados caso até 2 dos HDs falhem ao mesmo tempo**. Mínimo de 4 HDs para ser implementado. Ao usar 8 HDs de 20 GB cada um, em RAID 6, teremos 120 GB de dados e 40 GB de paridade.

Vantagem:

- *possibilidade falhar 2 HDs ao mesmo tempo sem perdas.*

Desvantagens:

- precisa de $N+2$ HDs para implementar por causa dos discos de paridade;
- escrita lenta;
- sistema complexo de controle dos HDs.

RAID Level

10 / 1+0

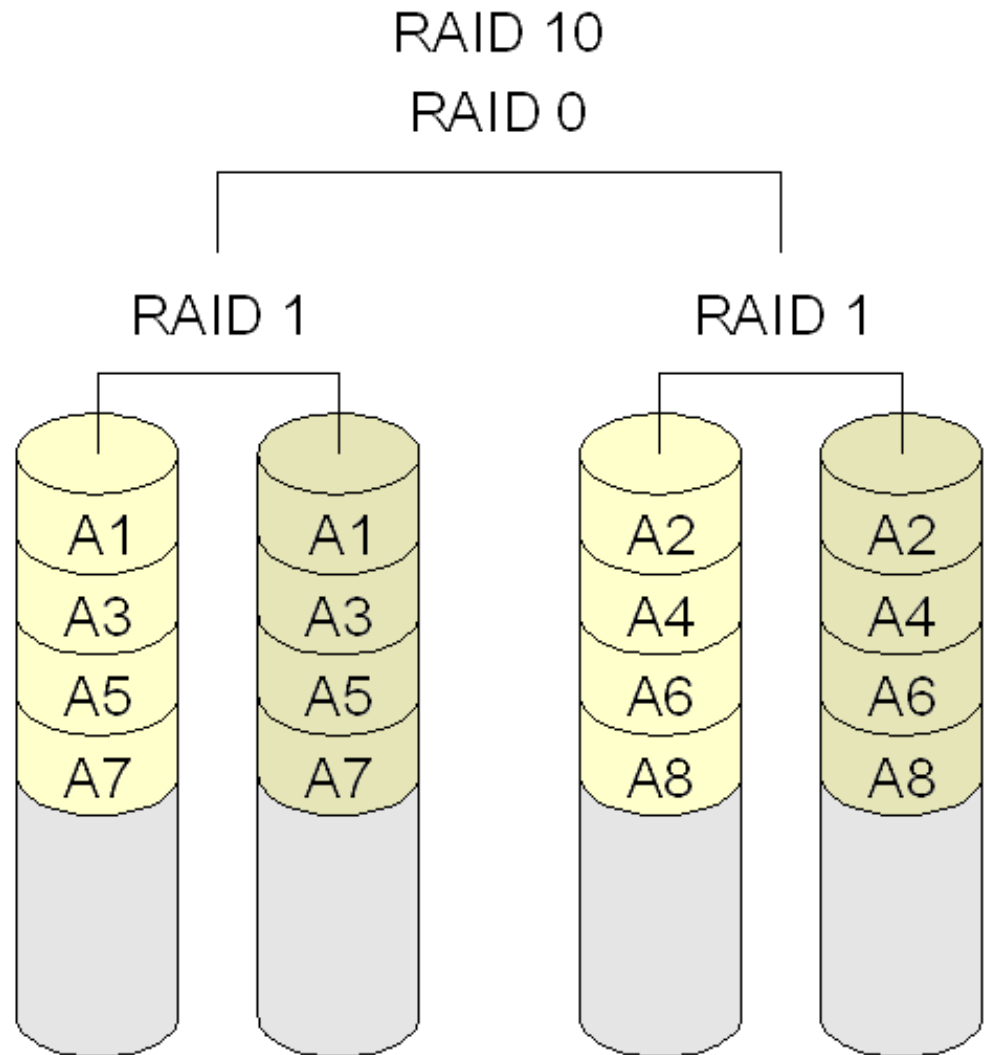
4 ou + discos

Falha decrementa para RAID 1

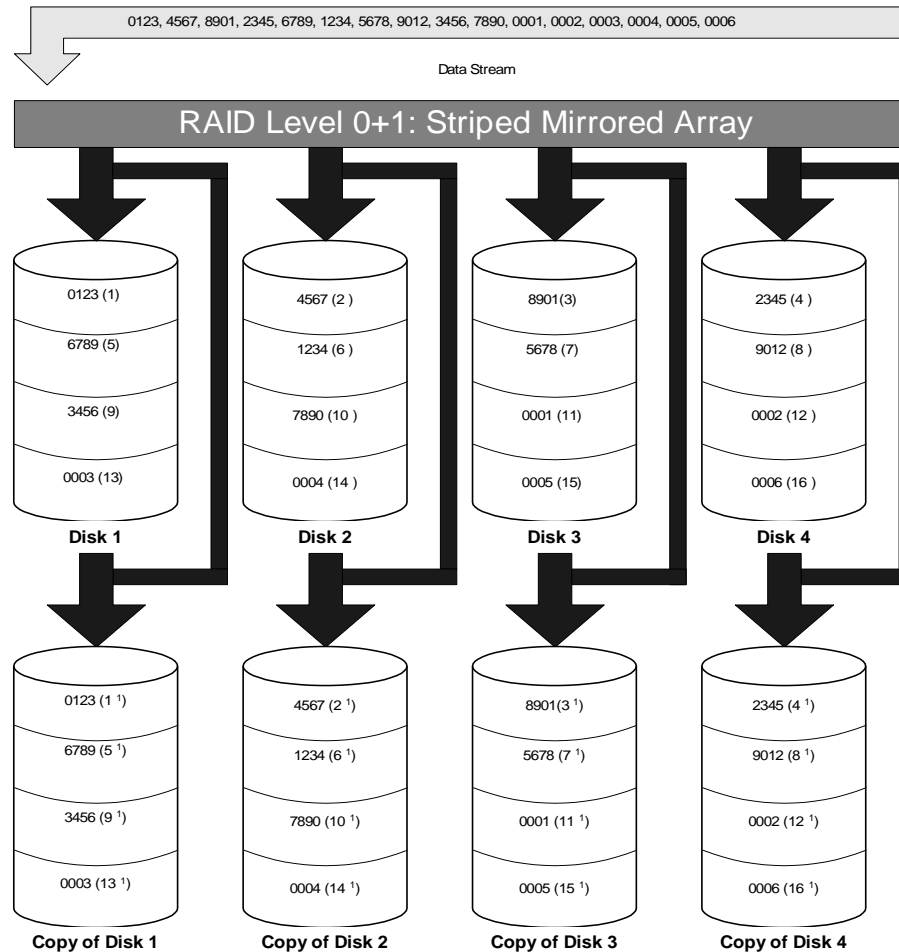
Até metade dos discos pode falhar simultaneamente, desde que não falhem os dois discos de um espelho qualquer

É o nível recomendado para bases de dados, por ser o mais seguro e dos mais velozes

RAID Level
10 / 1+0



RAID Level 0+1



Block order shown in parentheses

RAID Level 0+1 combines data striping and disk mirroring

RAID Level 0+1

4 ou + discos

Falha decrementa para RAID 0

Pode falhar 1 dos HDs, ou os dois HDs do mesmo DiskGroup

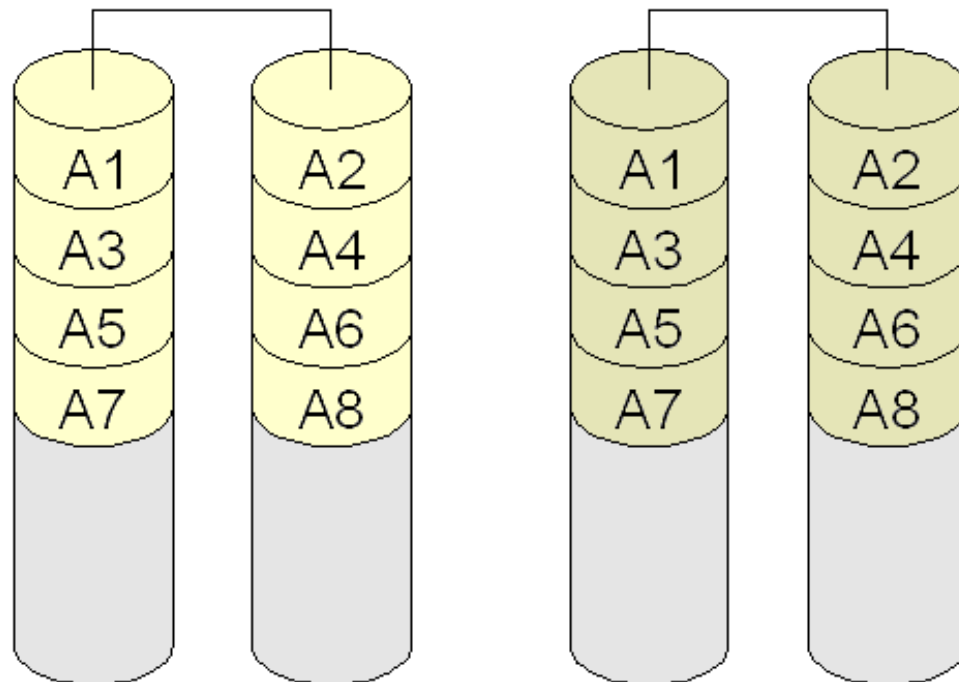
É o mais rápido e seguro, porém o mais caro de ser implantado de todos os RAID

RAID 0+1

RAID 1

RAID 0

RAID 0





RAID Level 50 / 5+0

- 6 ou + discos
- Admite falhas em até 2 discos, desde que em segmentos diferentes
- É um arranjo híbrido que usa as técnicas de RAID com paridade em conjunção com a segmentação de dados

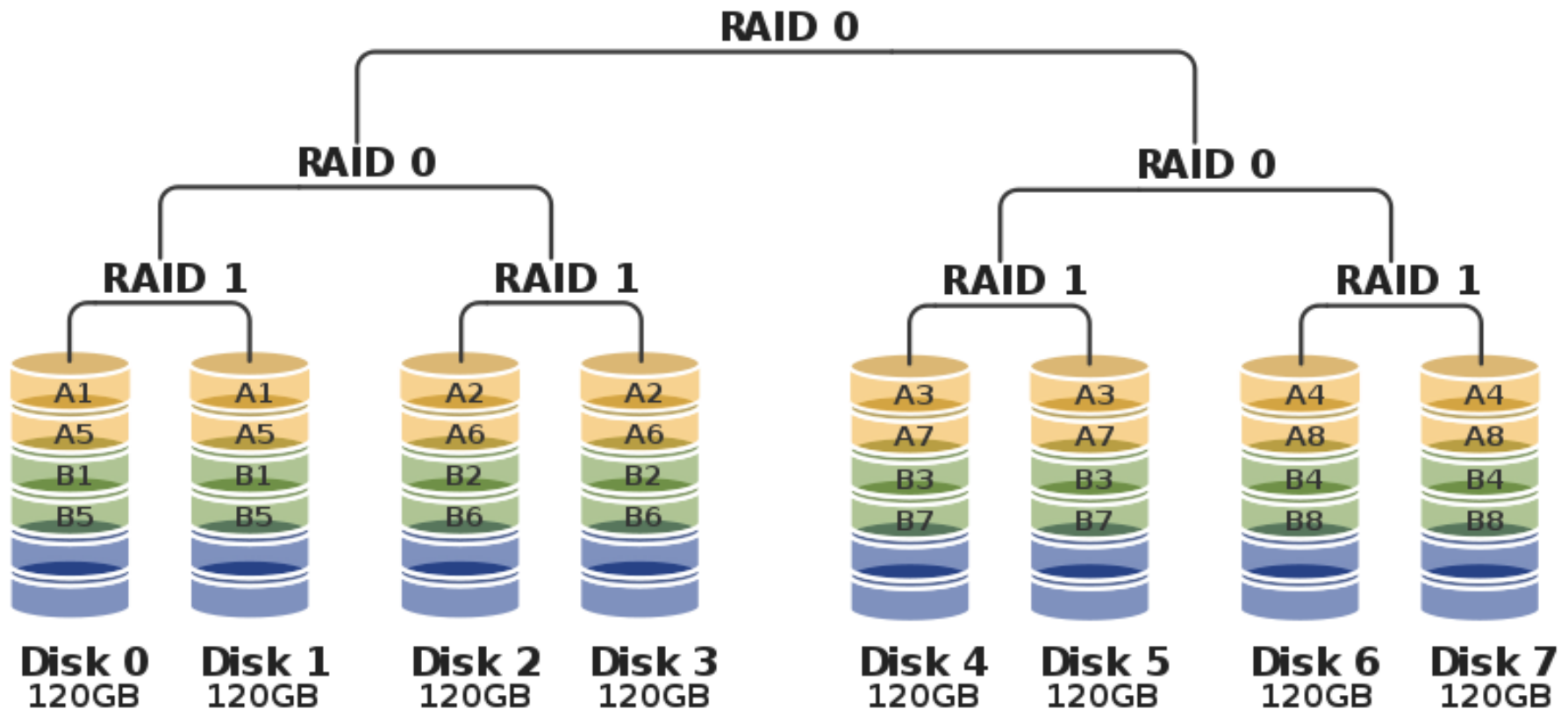


RAID Level 100 10+0

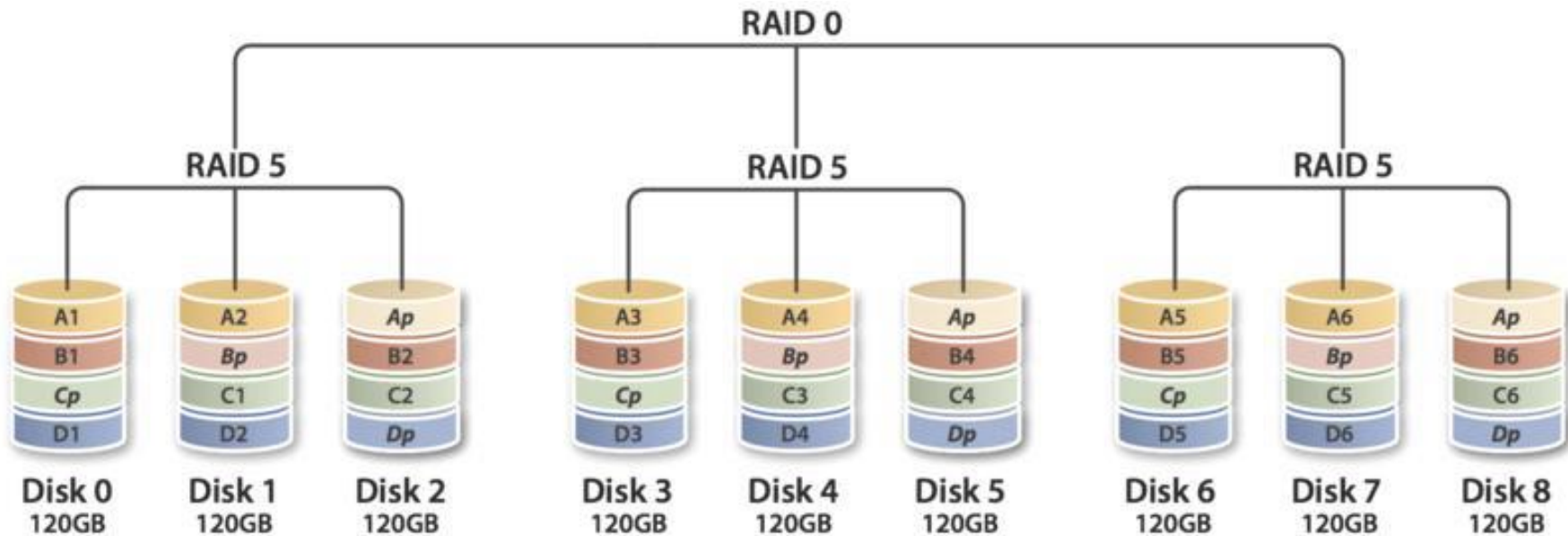
- 8 ou + discos
- É composto do RAID 10+0
- Implementa-se o RAID 0 via software sobre o RAID 10 via Hardware



RAID Level 100 10+0



RAID Level 50 / 5+0



Aplicações RAID

- Aplicações com Alta Taxa de Transferência (Ambientes tipicamente RAID 0)
- RAID striping é o ideal para aplicações com intensa transferência de dados
 - Aplicações que necessitam de uma grande quantidade de dados a ser processada em um intervalo de tempo pré-definido
 - Streaming Media
 - Video on demand, Digital Cable, Transmissões ao vivo
 - A taxa de dados fixa é crítica, Transmissões múltiplas não devem degradar a qualidade
 - “QoS” = Quality of Service
 - Stream contínuo de dados – sem espaço para o reenvio
 - Processamento de Imagem, manipulação e Renderização

Aplicações RAID

- Aplicações com alto índice de solicitações de informação (ambiente típico para o RAID 5)
- RAID é utilizado para aplicações de alto uso de multitasking, alta taxa de retorno de informação
- OLTP = **O**n **L**ine **T**ransaction **P**rocessing (aplicações típicas)
 - Validação de cartão de crédito, Reservas, ECommerce, Gerenciamento de estoque on-line
 - Banco dados e aplicações Web based
 - Alto número de transações randômicas