

# Módulo – 3

## Proteção de dados– RAID



## Módulo 3: Proteção de dados – RAID

Ao terminar este módulo você estará apto a:

- Descrever os métodos de implementação RAID
- Descrever as três técnicas de RAID
- Descrever os níveis de RAID mais utilizados
- Descrever o impacto de RAID no desempenho
- Comparar os níveis de RAID com base nos custos, desempenho e proteção

Este módulo abrange o RAID e seu uso para melhorar o desempenho e a proteção, bem como suas diversas implementações, técnicas e níveis mais utilizados . Ele também descreve o impacto do RAID no desempenho e compara os níveis do RAID utilizados.

## Módulo 3: Proteção de dados – RAID

### Aula 1: Visão geral de RAID

Durante esta aula os seguintes tópicos serão apresentados:

- métodos de implementação RAID
- componentes de array RAID
- técnicas de RAID

Esta aula apresenta os métodos de implementação do RAID e os componentes de array RAID, além de apresentar varias técnicas do RAID.

## Por quê RAID?

### RAID

É uma técnica que combina múltiplos drives de disco em uma unidade lógica (conjunto RAID) e fornece proteção e desempenho.

- Oferece desempenho limitado devido aos componentes mecânicos em um drive de disco
- Um único drive possui uma expectativa de vida determinada e é medida em MTBF:
  - ▶ por exemplo: caso o MTBF de um drive seja de 750.000 horas e têm 1.000 drives no array, então o MTBF da array é de 750 hours <sup>A2</sup> (750.000/1.000)
- O RAID foi apresentado para diminuir estes problemas

Atualmente, os data centers abrigam centenas de drives de discos em sua infraestrutura de armazenamento. Os drives de disco são suscetíveis a falhas devido ao uso ou desgaste mecânico, que podem resultar em perda de dados. Quanto maior o número de drives de discos no storage array, maior a probabilidade de falha do disco no array. Considerando um storage array de 100 drives de discos, cada um com uma expectativa média de vida de 750.000 horas, a expectativa média de vida deste conjunto em um array é de 750.000/100 ou 7.500 horas.

RAID é uma tecnologia que permite aproveitar múltiplos drives como parte de um conjunto proporcionando proteção contra as falhas no drive. No geral, as implementações do RAID também melhoram o desempenho do sistema de armazenamento enviando simultaneamente I/Os de múltiplos drives. Os arrays modernos com flash drives também se beneficiam em termos de proteção e desempenho utilizando o RAID.

Em 1987, Gibson e Katz, da Universidade da Califórnia, Berkeley, publicaram um artigo intitulado "A Case for Redundant Arrays of Inexpensive Disks (RAID)." Este artigo descreveu o uso dos drives de discos baratos e com baixa capacidade como uma alternativa para drives com grande capacidade, comuns nos mainframes dos computadores. O termo RAID tem sido redefinido para se referir a discos *independentes* a fim de refletir os avanços da tecnologia de armazenamento. A tecnologia RAID deixou de ser somente um conceito acadêmico para se tornar um padrão de mercado e é uma implementação comum nos storage arrays nos dias de hoje.



## Métodos de implementação de RAID

- Implementações do software do RAID
  - ▶ Utiliza software com base no host para fornecer a funcionalidade de RAID
  - ▶ Limitações
    - ▶▶ Utiliza os ciclos da CPU do host para executar cálculos do RAID, consequentemente há um impacto no desempenho geral do sistema
    - ▶▶ Suporta níveis limitados do RAID
    - ▶▶ O software RAID e o SO serão atualizados somente se forem compatíveis
- Implementação do hardware do RAID
  - ▶ Utiliza uma controladora de hardware especializada instalada no host ou no array

Há dois métodos de implementação do RAID, hardware e software. Ambos possuem suas vantagens e desvantagens.

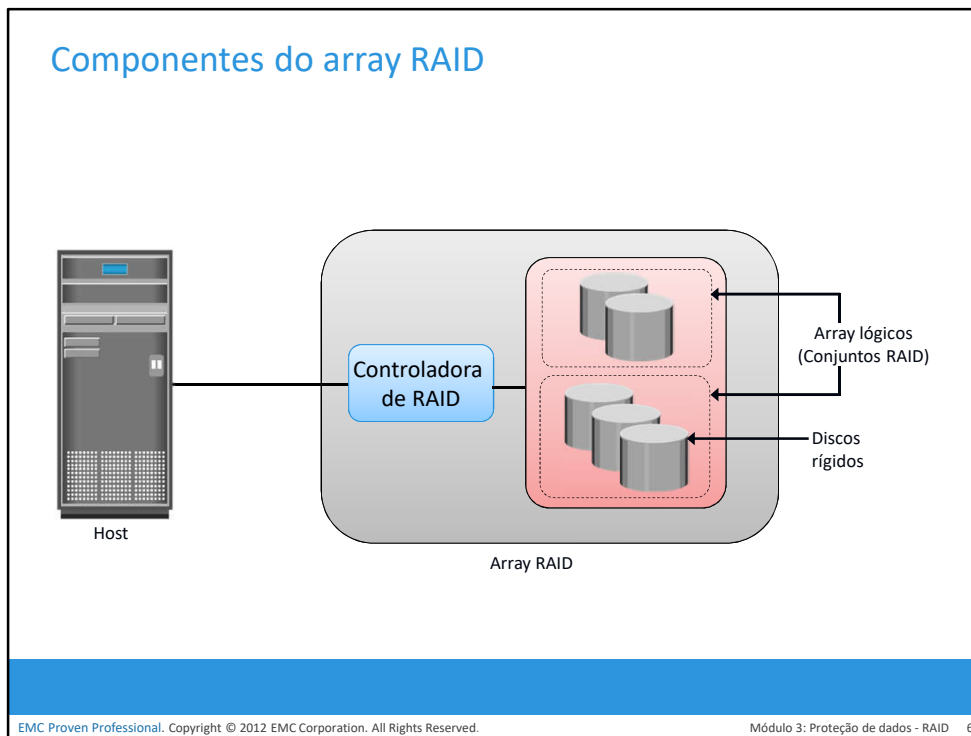
O RAID por software utiliza um software no host para oferecer funções do RAID e é implementado no nível do sistema operacional. As implementações do RAID por software oferecem benefícios de custo e simplicidade comparado com o RAID por hardware. Entretanto, eles possuem as seguintes limitações:

**Desempenho:** o software RAID afeta todo o desempenho do sistema, pois há a necessidade de ciclos adicionais da CPU para executar os cálculos do RAID.

**Características compatíveis:** o RAID por software não é compatível com todos os níveis do RAID.

Nas implementações do RAID por hardware, uma controladora de hardware especializada é implementada tanto no host como no array. A placa controladora RAID é uma implementação do RAID por hardware no host no qual uma controladora RAID especializada está instalada no host e os drives de discos estão conectados ao host. Os fabricantes também integram as controladoras RAID nas placas-mãe. Uma controladora do RAID no host não é uma solução eficaz em ambientes de data center com muitos hosts. A controladora do RAID externa é um hardware em array e funciona como uma interface entre o host e os discos. Ela apresenta os volumes de armazenamento ao host, que, por sua vez, gerencia estes volumes como drives físicos. As principais funções da controladora do RAID são:

- Gerenciamento e controle dos conjuntos de discos
- Tradução das solicitações de I/O entre os discos lógicos e físicos
- Recuperação dos dados em casos de falhas nos discos



Um array do RAID é um compartimento com um número de drives de discos que ajudam o hardware a implementar o RAID. Um subconjunto de discos dentro do array do RAID pode ser agrupado para formar associações lógicas, chamadas de arrays lógicos também conhecidos como conjunto RAID ou grupo RAID.

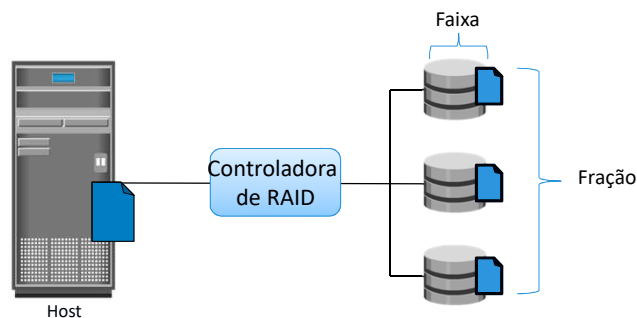


## Técnicas do RAID

- As três técnicas importantes utilizadas por RAID são:
  - ▶ Striping
  - ▶ Espelhamento
  - ▶ Paridade

As técnicas RAID, striping, espelhamento e paridade, são as bases para definir os vários níveis do RAID. Estas técnicas determinam a disponibilidade de dados e as características de desempenho de um conjunto RAID.

## Técnica de RAID – Striping



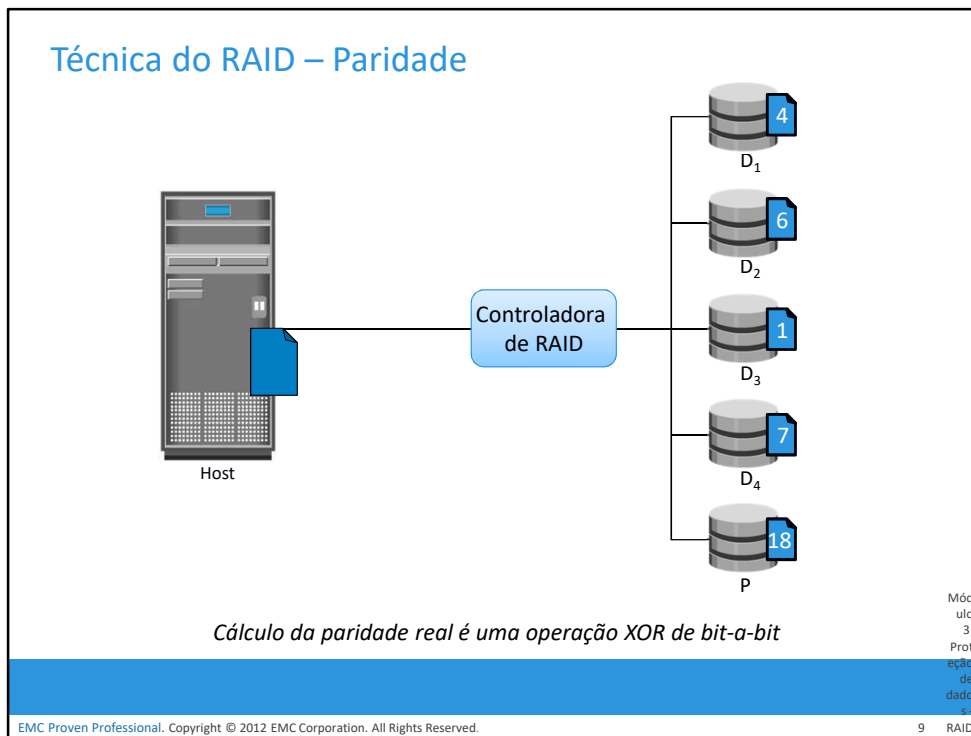
EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Módulo 3: Proteção de dados - RAID 8

Striping é a técnica de distribuir os dados entre vários drives (mais de um) a fim de utilizar os drives em paralelo. Todos os cabeçotes de leitura/gravação trabalham simultaneamente, permitindo que mais dados sejam processados num período de tempo mais curto e com maior desempenho, comparado com a leitura e gravação de um único disco.

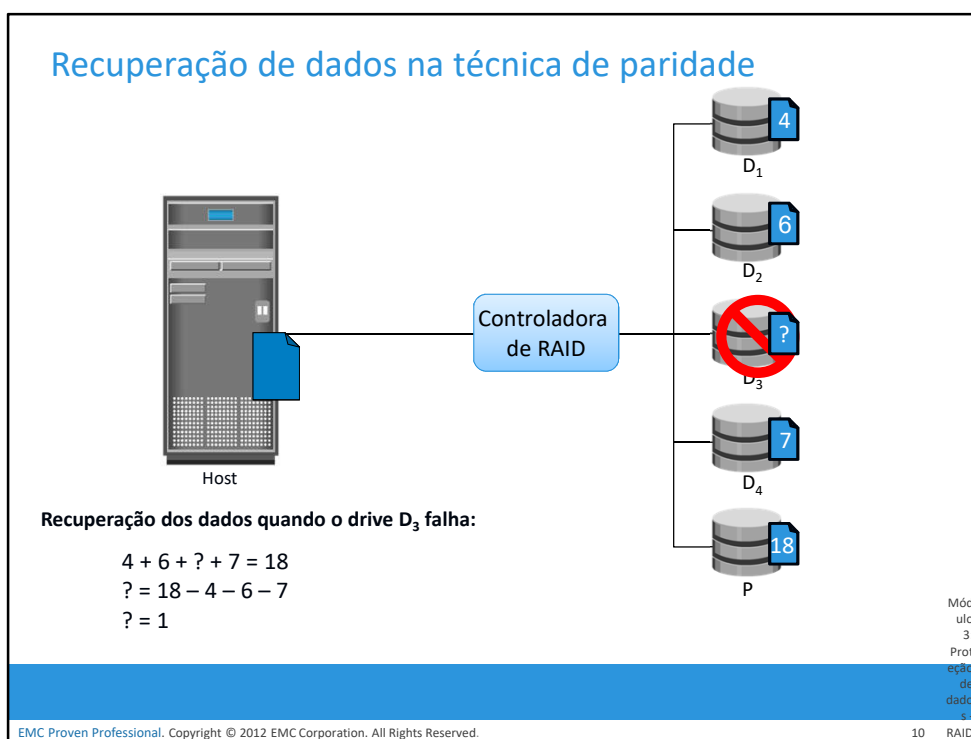
Dentro de cada disco de um conjunto RAID, um número predefinido de blocos de discos endereçáveis contiguamente são definidos como faixas. O conjunto de faixas alinhadas que se estendem pelos discos dentro do conjunto RAID é chamado de striping. A figura no slide mostra as representações físicas e lógicas de um conjunto RAID distribuído.

O tamanho da faixa (também chamado de profundidade do striping) descreve o número de blocos em uma faixa e é o número máximo de dados que podem ser gravados ou lidos a partir de um único disco do conjunto, supondo que os dados acessados comecem no início da faixa. Todas as faixas em um striping possuem o mesmo número de blocos. Ter um tamanho de faixa menor significa que os dados estão quebrados em pequenos pedaços nos discos. O tamanho do striping é o resultado da multiplicação do tamanho da faixa pelo número de discos de dados em um conjunto RAID. Por exemplo, num disco com 5 conjuntos RAID repartidos com o tamanho da faixa de 64KB, o tamanho do striping é de 320 KB (64KB x 5). A largura do striping refere-se ao número de faixas de dados em um striping. O RAID repartido não oferece qualquer proteção de dados a menos que a paridade ou espelhamento sejam utilizados.



A *paridade* é o método para proteger os dados repartidos de erros no drive de disco sem o custo do espelhamento. Um drive de disco é acrescentado para sustentar a paridade, uma construção matemática que permite a recriação dos dados perdidos. A paridade é um técnica redundante que assegura a proteção dos dados sem manter um conjunto completo dos dados duplicados. O cálculo da paridade é uma função da controladora RAID.

A informação da paridade pode ser armazenada em drives de discos dedicados e separados ou distribuídos por todos os drives de um conjunto RAID. Os quatros primeiros discos na figura, classificados de *D1 a D4*, contém dados. O quinto disco, classificado como *P*, armazena a informação de paridade, a qual, neste caso, é a soma dos elementos de cada linha.



Agora, caso um dos discos de dados falhe, o valor perdido pode ser calculado subtraindo a soma do restante dos elementos de seu valor de paridade. Aqui, pela simplicidade, o cálculo de paridade está representado pela soma aritmética dos dados. Entretanto o cálculo é uma operação *XOR de bit a bit*.

Comparada com o espelhamento, a implementação da paridade reduz consideravelmente os custos associados à proteção de dados. Considere uma configuração RAID com cinco discos, onde quatro deles sustentam os dados e o quinto sustenta a informação de paridade. Neste exemplo, a paridade necessita de 25% de espaço extra do disco comparado ao espelhamento, que precisa de 100% de espaço extra do disco. Porém, há algumas desvantagens no uso da paridade. A informação na paridade é gerada a partir de dados do disco de dados. Portanto, a paridade é recalculada toda vez que houver uma mudança nos dados. Este recálculo consome tempo e afeta o desempenho da controladora RAID.

Para o RAID paridade, o cálculo do tamanho do striping não inclui a faixa de paridade. Por exemplo, em um conjunto RAID paridade de cinco discos (4+1) com o tamanho da faixa de 64 KB, o tamanho do striping será de 256 KB (64KB x 4).

## Módulo 3: Proteção de dados – RAID

### Aula 2: Níveis do RAID

Durante esta aula apresentaremos os seguintes tópicos:

- Os níveis do RAID mais utilizados
- Os impactos do RAID no desempenho
- Comparação do RAID
- Hot spare

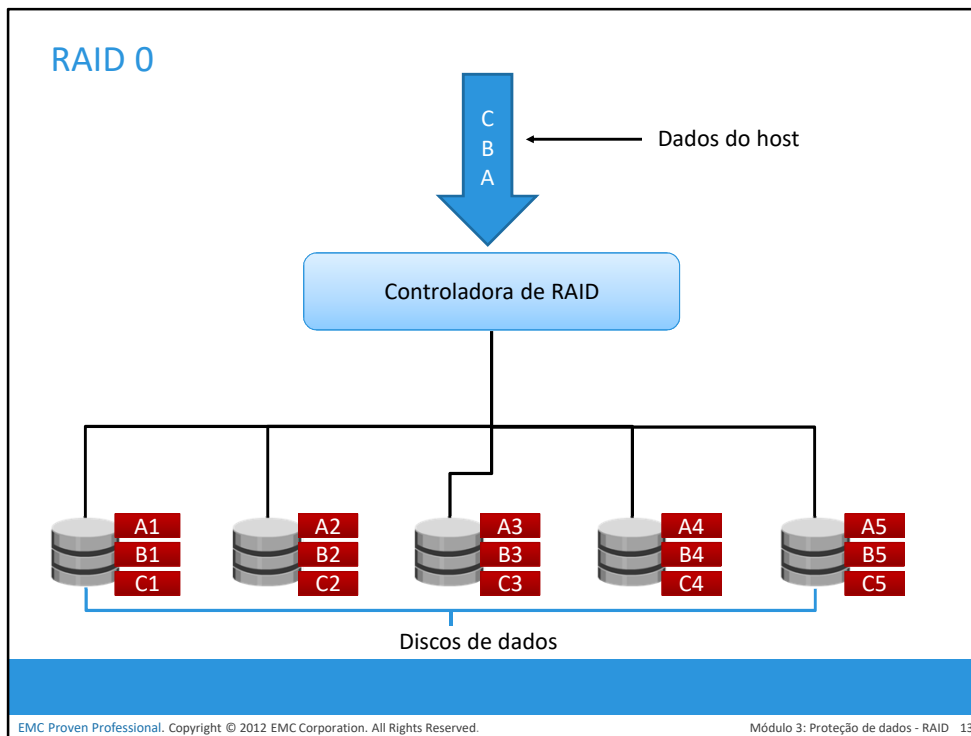
Esta aula explica o níveis do RAID mais utilizados e suas comparações e também sobre hot spare.

## Níveis do RAID

- Os níveis do RAID mais utilizados são:
  - ▶ RAID 0 – Conjunto fracionados <sup>A3</sup> sem tolerância à falhas
  - ▶ RAID 1 – Espelhamento do disco
  - ▶ RAID 1 + 0 – RAID agrupado
  - ▶ RAID 3 – Conjunto fracionado com acesso paralelo e disco de paridade dedicado
  - ▶ RAID 5 – Conjunto fracionado com acesso ao disco independente e paridade distribuída
  - ▶ RAID 6 – Conjunto fracionado com acesso ao disco independente e paridade dupla distribuída

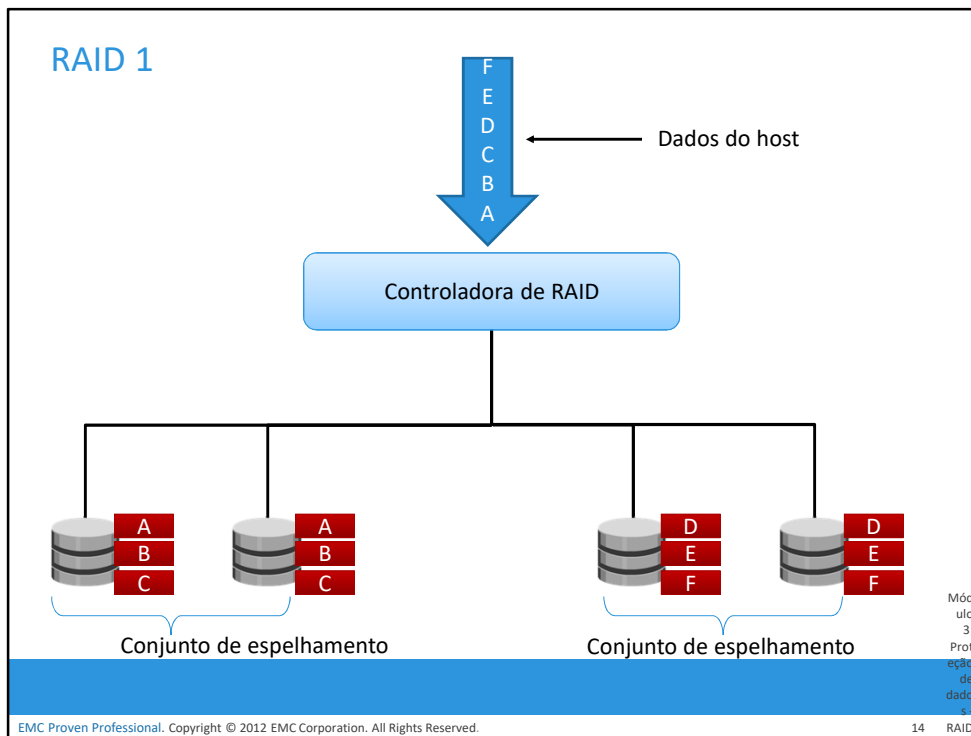
O desempenho do aplicativo, a necessidade de disponibilidade de dados e o custo determinam a escolha no nível do RAID. Estes níveis do RAID são definidos com base nas técnicas de striping, espelhamento e paridade. Alguns níveis do RAID utilizam técnicas simples, enquanto outros utilizam técnicas combinadas. Os níveis do RAID mais utilizados estão listados no slide.

**A3**      **fracionado**  
Autor; 02/05/2015

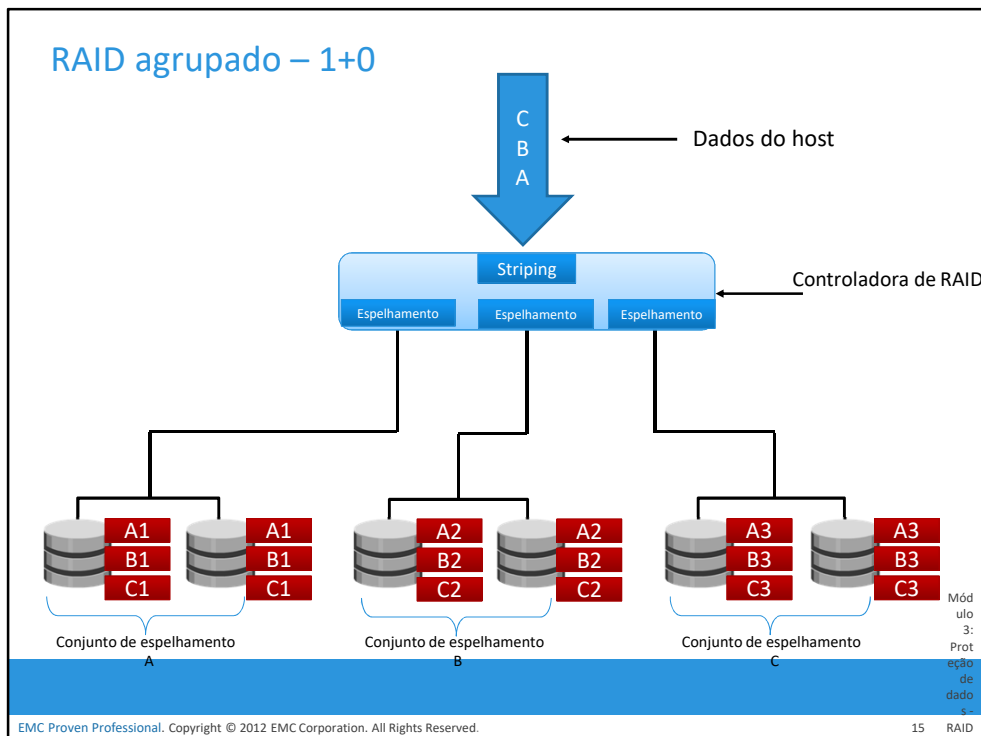


A configuração RAID 0 utiliza as técnicas de striping de dados, onde os dados são repartidos por todos os discos de um conjunto RAID. Portanto, ele utiliza a capacidade total de armazenamento do conjunto RAID. Para ler os dados, todas as faixas são colocadas juntas pela controladora. Quando o número de drives num conjunto RAID aumenta, o desempenho melhora, pois mais dados podem ser lidos ou gravados simultaneamente. O RAID 0 é uma boa opção para os aplicativos que precisam de alto rendimento de I/O. Porém, se estes aplicativos necessitarem de maior disponibilidade durante os erros dos drives, o RAID 0 não proporcionará a proteção e disponibilidade dos dados.



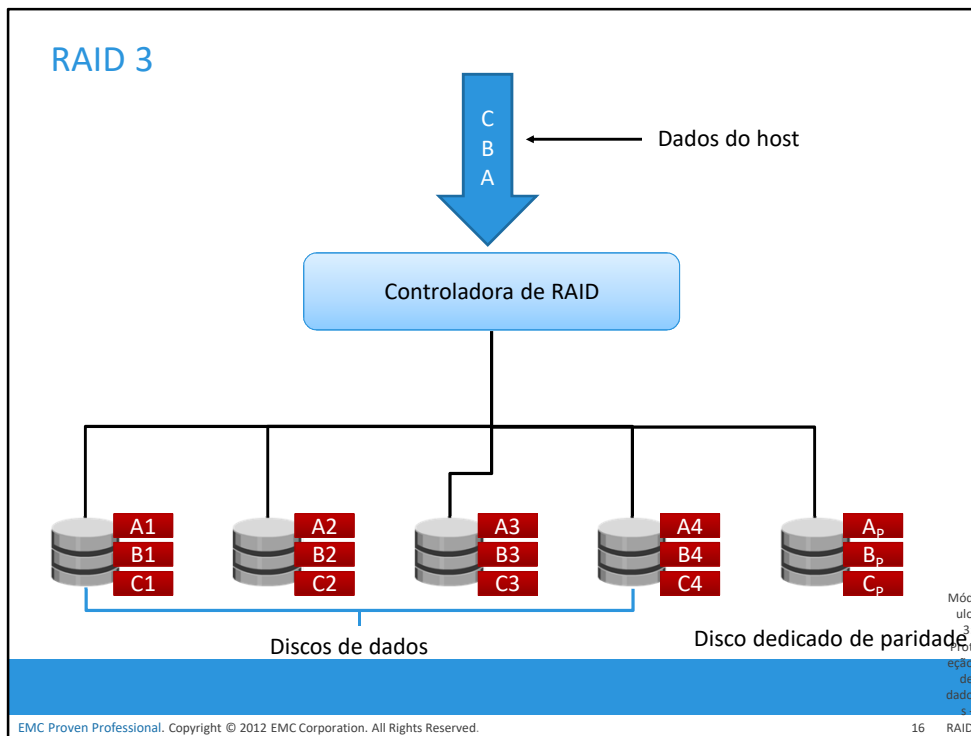


RAID 1 é baseado na técnica de espelhamento. Nesta configuração do RAID, os dados são espelhados para oferecer uma tolerância à falha. O conjunto do RAID 1 consiste de dois drives de discos e todas as gravações são feitas em ambos os discos. O espelhamento é transparente ao host. Durante uma falha do disco, o impacto na recuperação dos dados em um RAID 1 é o menor entre todas as implementações RAID. Isto por que a controladora do RAID utiliza o drive espelhado para a recuperação de dados. O RAID 1 é adequado para aplicativos que necessitam de uma alta disponibilidade sem restrições de custo.



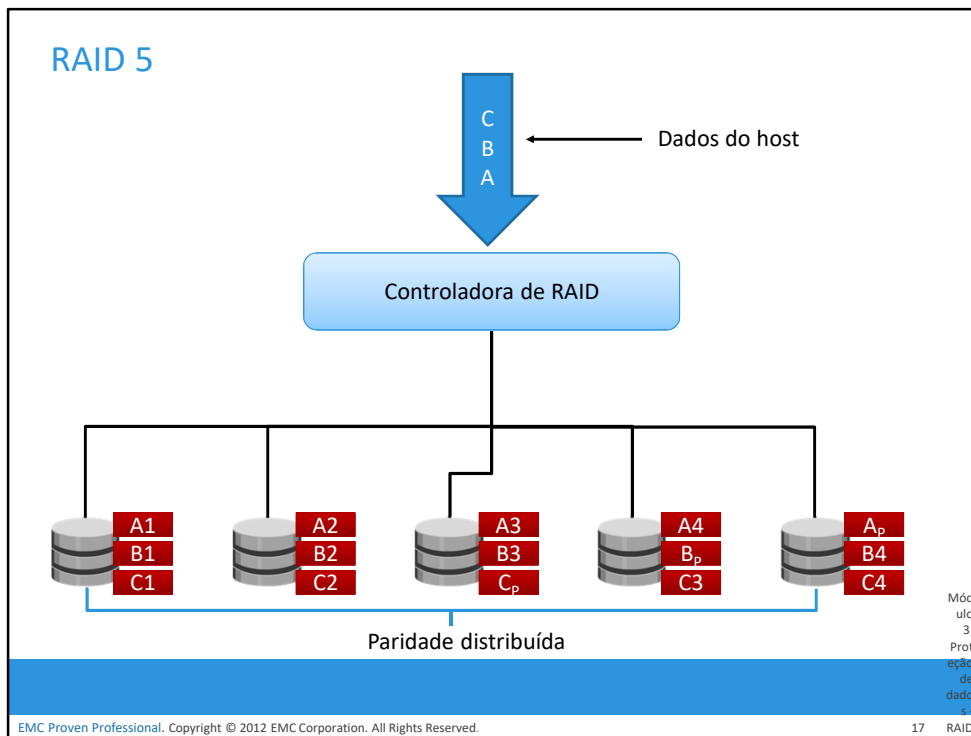
A maioria dos data centers necessitam redundância e desempenho de dados de seus arrays RAID. RAID 1+0 combina os benefícios do desempenho do RAID 0 com os benefícios de redundância do RAID 1. Utiliza-se as técnicas de espelhamento e striping e também seus benefícios. Este tipo do RAID exige um número par de discos, sendo o mínimo de quatro.

O RAID 1+0 também é conhecido por RAID 10 (dez) ou RAID 1/0. O RAID 1+0 também é chamado de striped mirror. O elemento básico do RAID 1+0 é o par espelhado, o que significa que os dados são primeiramente espelhados e depois, as duas cópias de dados são repartidas em vários pares de drives de discos de um conjunto RAID. Ao substituir um drive com erro, somente o espelho é reconstruído. Em outras palavras, a controladora array de discos utiliza o drive não afetado no par espelhado para recuperar os dados e continuar a operação. Os dados do disco não afetado são copiados para o disco substituído.

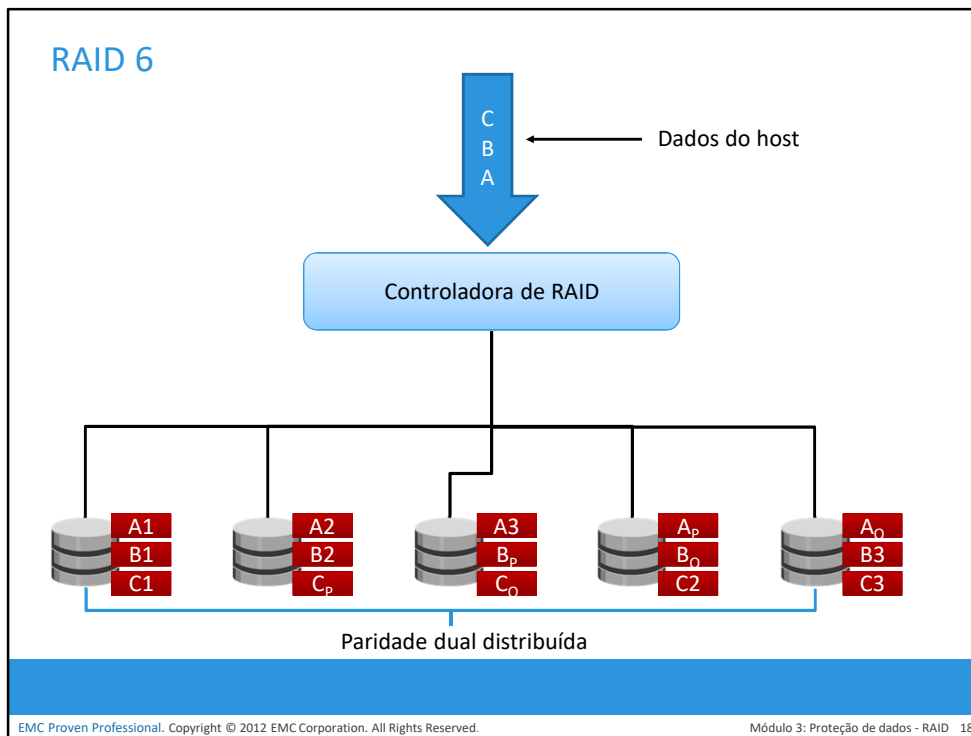


O RAID 3 fraciona os dados para obter desempenho e utiliza a paridade para gerar tolerância à erros. A informação de paridade é armazenada no drive dedicado para que os dados possam ser reconstruídos caso o drive falhe no conjunto RAID. Por exemplo, em um conjunto de cinco discos, quatro são utilizados para os dados e um para paridade. Portanto, o espaço total necessário do disco é de 1,25 vezes o tamanho dos discos de dados. O RAID 3 sempre lê e grava os stripings completos de dados por todos os discos, pois os drives operam paralelamente. Não existe gravações parciais que atualizem uma entre várias faixas em um striping.

Semelhante ao RAID 3, o RAID 4 reparte os dados para obter desempenho alto e utiliza a paridade para melhorar a tolerância à erros. Os dados são fracionados por todos os discos, exceto o disco de paridade no array. A informação de paridade é armazenada no disco dedicado para que os dados possam ser reconstruídos caso o drive falhe. Diferentemente do RAID 3, os discos de dados no RAID 4 podem ser acessados independentemente para que os elementos de dados específicos possam ser lidos ou gravados em um único disco sem a leitura ou gravação do striping inteiro. O RAID 4 proporciona uma boa transferência de leitura e uma transferência razoável de gravação.

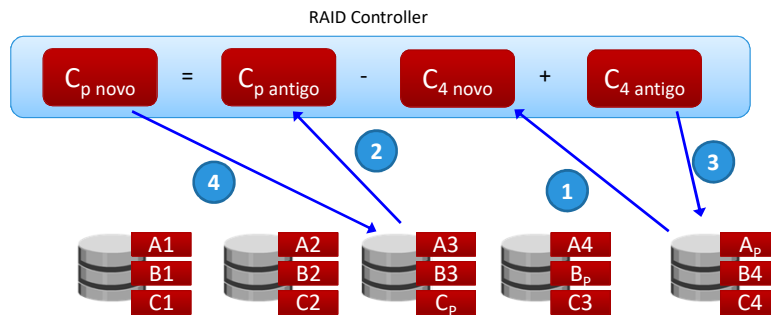


O RAID 5 é uma implementação versátil do RAID. É similar ao RAID 4 por utilizar o striping. Os drives (faixas) também são acessíveis independentemente. A diferença entre o RAID 4 e o RAID 5 é a localização da paridade. No RAID 4, a paridade é gravada no drive dedicado, criando um gargalo de gravação para o disco de paridade. No RAID 5, a paridade é distribuída por todos os discos para superar este gargalo de gravação no disco de paridade dedicado.



O RAID 6 trabalha da mesma maneira que o RAID 5, exceto que o RAID 6 inclui um segundo elemento de paridade que permite a sobrevivência caso ocorra duas falhas do disco no conjunto RAID. Portanto, a implementação do RAID 6 necessita de pelo menos quatro discos. O RAID 6 distribui a paridade por todos os discos. A degradação de gravação (explicada mais tarde neste módulo) no RAID 6 é maior do que a do RAID 5. Portanto, o desempenho de gravação do RAID 5 é melhor do que do RAID 6. A operação de reconstrução no RAID 6 pode levar mais tempo do que no RAID 5 devido a presença de dois conjuntos de paridades.

## Impacto do RAID no desempenho



- No RAID 5, toda gravação (atualização) no disco aparece como 4 operações de I/O (2 leituras do disco e 2 gravações no disco)
- No RAID 6, toda gravação (atualização) no disco aparece como 6 operações de I/O (3 leituras do disco e 3 gravações no disco)
- No RAID 1, toda gravação aparece como 2 operações de I/O (2 gravações no disco)

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Módulo 3: Proteção de dados - RAID 19

Ao escolher o tipo do RAID, é importante considerar seu impacto no desempenho do disco e no aplicativo IOPS. Em ambas as configurações RAID, espelhadas e de paridade, cada operação de gravação significa mais sobrecarga de I/O para os discos, que é chamada de *degradação de gravação*. Na implementação do RAID 1, cada operação de gravação deve ser executada em dois discos configurados como um par espelhado, enquanto que na implementação do RAID 5, a operação de gravação pode aparecer como quatro operações de I/O. Quando estiver executando I/Os para um disco configurado com RAID 5, a controladora tem que ler, recalculer e gravar um segmento de paridade para cada operação de gravação.

Este slide ilustra uma operação simples de gravação no RAID 5 que contém um grupo de cinco discos. A paridade (P) na controladora é calculada como:

$$C_p = C_1 + C_2 + C_3 + C_4 \text{ (operações XOR)}$$

Sempre que a controladora executar uma gravação de I/O, a paridade deve ser calculada pela leitura da paridade antiga ( $C_p$  antiga) e dos dados antigos ( $C_4$  antiga) do disco, o que significa duas leituras de I/O. Em seguida, uma nova paridade ( $C_p$  nova) é calculada da seguinte forma:

$$C_p \text{ antigo} = C_p \text{ antigo} - C_4 \text{ antigo} + C_4 \text{ novo (operações XOR)}$$

Depois de calcular a nova paridade, a controladora completa a gravação de I/O gravando os novos dados e uma nova paridade nos discos, significando duas gravações de I/O. Portanto, a controladora executa duas leituras de discos e duas gravações de discos para cada operação de gravação, e a degradação de gravação é 4.

No RAID 6, que mantém a paridade dual, a gravação do disco necessita de três operações de leitura: duas de paridade e uma dos dados. Após calcular as duas novas paridades, a controladora executa três operações de gravação: duas de paridade e uma de I/O. Portanto, em uma implementação RAID 6, a controladora executa seis operações de I/O e a degradação de gravação é 6.

### Exemplo de cálculo de degradação do RAID

- O total de IOPS no pico de carga de trabalho é 1.200
- Relação entre leitura/gravação 2:1
- Calcule a carga do disco no pico da atividade para:
  - ▶ RAID 1/0
  - ▶ RAID 5

Considere um aplicativo que gere 1.200 IOPS no pico de carga de trabalho, com uma taxa entre leitura/gravação de 2:1. Calcule a carga do disco no pico da atividade para as configurações RAID 1/0 e RAID 5.



## Solução: degradação do RAID

- Para RAID 1/0, a carga do disco (leitura + gravação)  
$$= (1.200 \times 2/3) + (1.200 \times (1/3) \times 2)$$
$$= 800 + 800$$
$$= 1.600 \text{ IOPS}$$
- Para RAID 5, a carga do disco (leitura + gravação)  
$$= (1.200 \times 2/3) + (1.200 \times (1/3) \times 4)$$
$$= 800 + 1.600$$
$$= 2.400 \text{ IOPS}$$

## Comparação do RAID

Nível do RAID	Qtde. min. discos	Capacidade disponível de armazenamento(%)	Desempenho de leitura	Desempenho de gravação	Degradação de gravação	Proteção
1	2	50	Melhor do que um único disco	Mais lento do que um único disco pois cada gravação deve ser enviada à todos os discos	Moderada	Espelhamento
1+0	4	50	Bom	Bom	Moderada	Espelhamento
3	3	$[(n-1)/n]*100$	Razoável para leituras aleatórias e bom para leituras sequenciais	Para gravações curtas e aleatórias: ruim à razoável Razoável para extensas gravações sequenciais	Alta	Paridade (Suporta falha de um único disco)
5	3	$[(n-1)/n]*100$	Bom para leituras sequenciais aleatórias	Razoável para gravações sequenciais aleatórias	Alta	Paridade (Suporta falha de um único disco)
6	4	$[(n-2)/n]*100$	Bom para leituras sequenciais aleatórias	Para gravações sequenciais aleatórias: ruim à razoável	Muito alta	Paridade (Suporta falha de dois discos)

n = número de discos

A tabela no slide compara os diferentes níveis do RAID.

## Níveis do RAID adequados para diferentes aplicativos

- RAID 1+0
  - ▶ Adequado para aplicativos com o perfil de I/O de gravações curtas, intensivas e aleatória (gravações normalmente maiores do que 30%)
  - ▶ Exemplo: OLTP, RDBMS – espaço temporário
- RAID 3
  - ▶ Leituras e gravações extensas e sequenciais
  - ▶ Exemplo: backup de dados e fluxo contínuo de multimídia
- RAID 5 e 6
  - ▶ Carga de trabalho pequena (gravações normalmente menores que 30%)
  - ▶ Exemplo: e-mail, RDBMS – entrada de dados

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Módulo 3: Proteção de dados - RAID 23

Entre os aplicativos comuns que se beneficiam dos diferentes níveis do RAID estão:

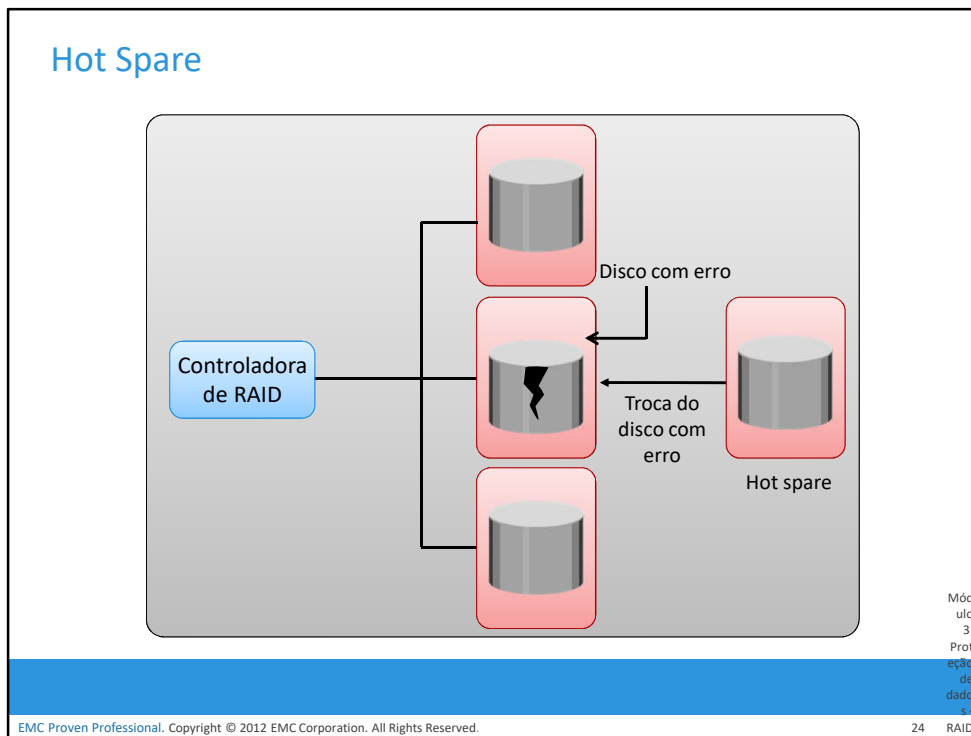
O RAID 1+0 tem um bom desempenho para cargas de trabalho que utilizam I/Os com gravações curtas intensivas e aleatórias.

Alguns dos aplicativos que se beneficiam do RAID 1+0 são os OLTPs (Online Transaction Processing, processamentos de transação on-line) com taxas altas, espaço temporário de RDBMS e etc.

O RAID 3 oferece um bom desempenho para aplicativos que utilizam extensos acessos de dados sequenciais tais como backup de dados ou fluxo contínuo de vídeo.

O RAID 5 é bom para aplicativos com leitura intensiva e aleatória de I/Os e de preferência, para mensagens, serviços de mídia com desempenho médio e implementações de RDBMS (Relational Database Management System), nas quais os DBAs (Database Administrators) otimizam o acesso aos dados.

**A4**      **aleatórias**  
Autor; 02/05/2015



Um *hot spare* é um drive sobressalente num array RAID que substitui, temporariamente, um drive de disco com erro ao tomar sua identidade. Com o hot spare, dependendo da implementação RAID, é realizado um dos métodos de recuperação de dados a seguir:

Se a paridade RAID for utilizada, os dados são reconstruídos no hot spare a partir da paridade e os dados bons dos drives de disco, no conjunto RAID.

Se o espelhamento for utilizado, os dados do espelhamento bom são copiados no hot spare.

Quando um novo drive de disco é colocado no sistema, os dados do hot spare são copiados nele. O hot spare volta ao seu estado inicial, pronto para ser utilizado quando houver uma nova falha no drive. De um outro modo, o hot spare substitui permanentemente o drive de disco com erro. Isto significa que não será mais um hot spare, e um novo hot spare deve ser configurado no array.

Um hot spare deve ser grande o suficiente para acomodar os dados do drive com falha. Alguns sistemas utilizam múltiplos hot spares para melhorar a disponibilidade dos dados.

Um hot spare pode ser configurado como automático ou inicializado pelo usuário, que especifica como ele será utilizado no caso de erro no disco. Na configuração automática, quando a taxa de recuperação de erros para um disco exceder o limite pré-determinado, o

subsistema do disco tenta copiar automaticamente os dados do disco com erro para o hot spare. Se esta tarefa for concluída antes que o disco danificado falhe, o subsistema muda para o hot spare e marca o disco com erro como não utilizável. Caso contrário, ele utiliza o disco de paridade ou espelhado para recuperar os dados. No caso da configuração ser iniciada pelo usuário, o administrador tem o controle sob o processo de recuperação. Por exemplo, a recuperação pode ocorrer durante a noite para evitar qualquer problema no desempenho do sistema. Porém, o sistema corre o risco de perder dados caso um novo erro no disco ocorra.

## Módulo 3: resumo

Principais pontos apresentados neste módulo:

- Métodos e técnicas de implementação do RAID
- Níveis comuns do RAID
- Degradação de gravação do RAID
- Comparação dos níveis do RAID com base nos custos e desempenho

Este módulo abordou os dois métodos de implementação RAID por hardware e software .

As três técnicas nas quais os níveis do RAID são criados são striping, espelhamento e paridade.

Os níveis do RAID mais utilizados são o 0,1 1+0, 3, 5 e 6.

Ao escolher um tipo do RAID, é importante considerar o impacto no desempenho do disco e no aplicativo IOPS. Tanto na configuração espelhada como na paridade, cada operação de gravação significa mais sobrecarga de I/O para os discos, que é chamada de *degradação de gravação*. Para finalizar, este módulo comparou os diferentes níveis do RAID baseado nos seus custos, desempenho e degradação de gravação.

## Verifique seu conhecimento – 1

- Qual afirmação abaixo é verdadeira sobre a implementação do software RAID?
  - A. Os upgrades do sistema operacional não exigem validação de compatibilidade com o software RAID
  - B. É mais cara do que a implementação RAID
  - C. Suporta todos os níveis de RAID
  - D. Utiliza os ciclos de CPU para executar os cálculos de RAID
- Um aplicativo gera 400 IOPS pequenos aleatórios com a proporção de 3:1 de leitura/gravação. Qual é o IOPS corrigido de RAID em um disco para RAID 5?
  - A. 400
  - B. 500
  - C. 700
  - D. 900



## Verifique seu conhecimento – 2

- Qual é a degradação de gravação em uma configuração RAID 6 para pequenos I/Os aleatórios?
  - A. 2
  - B. 3
  - C. 4
  - D. 6
- Qual aplicativo é mais beneficiado ao utilizar o RAID 3?
  - A. Backup
  - B. OLTP
  - C. E-commerce
  - D. E-mail

## Verifique seu conhecimento – 3

- Qual é o tamanho da faixa num disco com 5 conjuntos RAID repartidos que possuem o tamanho da faixa de 64 KB?
  - A. 64 KB
  - B. 128 KB
  - C. 256 KB
  - D. 320 KB

## Exercício 1: RAID

- Uma empresa está pensando em reconfigurar o armazenamento para seu aplicativo de contabilidade a fim de obter maior disponibilidade
  - ▶ Configuração atual e desafios
    - ▶▶ O aplicativo executa 15% de gravações aleatórias e 85% de leituras aleatórias
    - ▶▶ Atualmente implementado com cinco discos com configuração RAID 0
    - ▶▶ Cada disco possui capacidade formatada anunciada de 200 GB
    - ▶▶ O tamanho total dos dados do aplicativo é de 730 GB, que não serão modificados nos próximos 6 meses
    - ▶▶ O final do ano fiscal está se aproximando, não é possível adquirir novos discos
- Tarefa
  - ▶ Recomende um nível do RAID que a empresa deva utilizar para reestruturar o ambiente e atender suas necessidades
  - ▶ Justifique sua escolha com base nos custos, desempenho e disponibilidade

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Módulo 3: Proteção de dados - RAID 29

**Perfil da Empresa:** empresa de serviços de telefonia sem fio, espalhada por todo o país, com 5.000 empregados no mundo. Esta empresa possui 7 escritórios regionais no país. Embora a empresa esteja bem financeiramente, ela ainda sente a pressão competitiva. Assim, a empresa precisa garantir que a infraestrutura de TI supere os erros tolerantes.

**Configuração e desafios atuais:** A empresa utiliza aplicativos diferentes para comunicação, contabilidade e gerenciamento. Todos os aplicativos estão hospedados em servidores individuais com os discos configurados no RAID 0. Toda atividade financeira é gerenciada e monitorada por um único aplicativo de contabilidade. É muito importante que os dados contábeis estejam disponíveis. O aplicativo executa por volta de 15% de operações de gravação aleatória e os outros 85% restantes para leitura aleatória. Os dados da contabilidade estão armazenados em um conjunto RAID 0 de 5 discos. Cada disco tem a capacidade formatada anunciada de 200 GB e o volume total de seus arquivos é de 730 GB. A empresa faz a backups todas as noites e remove as informações antigas, portanto, é improvável que haja mudanças no volume de dados nos próximos 6 meses. A empresa está próxima de encerrar seu ano fiscal e o orçamento para TI já está no fim. Não será possível comprar nem mesmo um drive de disco novo.

**Tarefa:** recomende um nível RAID que a empresa possa utilizar para reestruturar seu ambiente dentro de suas necessidades. Justifique sua escolha baseado em custo, desempenho e disponibilidade da nova solução.

## Exercício 2: RAID

- Agora a empresa (como discutida no exercício 1) está planejando reconfigurar o armazenamento para seu aplicativo de data base para alta disponibilidade
  - ▶ Configuração atual e os desafios:
    - ▶ O aplicativo executa 40% de gravações e 60% de leituras
    - ▶ Configuração atualmente implementada com seis discos em RAID 0 com capacidade anunciada de 200 GB em cada disco
    - ▶ Tamanho do data base é de 900 GB e aproximadamente 30% dos dados sofrerão modificações nos próximos 6 meses
    - ▶ É início do ano fiscal e a empresa possui um aumento no orçamento
  - Tarefa
    - ▶ Recomende um nível do RAID adequado para atender as necessidades da empresa
    - ▶ Faça uma estimativa dos custos de novas soluções (discos de 200GB custam \$1.000)
    - ▶ Justifique sua escolha com base nos custos, desempenho e disponibilidade

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Módulo 3: Proteção de dados - RAID 30

**Perfil da Empresa:** uma empresa de serviços de telefonia sem fio , no país todo, com 5000 empregados no mundo. Esta empresa possui 7 escritórios regionais no país. Embora a empresa esteja bem financeiramente, ela ainda sente a pressão competitiva. Como resultado, a empresa precisa garantir que a infraestrutura de TI supere os erros tolerantes.

**Configuração e desafios atuais:** A empresa utiliza um aplicativo para contabilidade que fica hospedado em um servidor individual com discos configurados em RAID 0. É o início de um novo ano fiscal e o departamento de TI teve um aumento no orçamento. Você foi chamado para recomendar as mudanças no ambiente do data center. Você investiga o ambiente do data center e percebe que os dados estão armazenados em um conjunto RAID 0 de 6 discos. Cada disco tem uma capacidade formatada anunciada de 200 GB e o tamanho total dos arquivos é de 900 GB. O volume de dados provavelmente será mudado em 30% nos próximos 6 meses e sua solução deve acomodar este crescimento. O aplicativo executa 40% de operações de gravação e os outros 60% são de leitura.

**Tarefa:** recomende um nível RAID que a empresa possa utilizar para reestruturar seu ambiente dentro de suas necessidades. Qual é o custo desta nova solução? Justifique sua escolha baseado em custo, desempenho e disponibilidade da nova solução.

**Note:** Um drive de disco novo de 200 GB custa \$1000. A controladora suporta todos os níveis RAID mais utilizados, portanto não será necessário ser trocada.



# Thank You!

EMC Proven Professional. Copyright © 2012 EMC Corporation. All Rights Reserved.

Módulo 15: Gerenciamento da infraestrutura de armazenamento 31

Isto conclui o treinamento. Obrigado pela sua participação.