

**CENTRO UNIVERSITÁRIO UNINORTE**  
**PÓS-GRADUAÇÃO EM GERÊNCIA DE BANCO DE DADOS**  
**SISTEMA DE ARMAZENAMENTO MASSIVO DE DADOS**



---

## **ARMAZENANDO GRANDES VOLUMES DE DADOS**

Prof. Luiz Matos

Rio Branco, AC, setembro de 2019.

---

***continuando ...***

### ❑ ATIVIDADE 1

---

Questionário de revisão.

<https://pollev.com/luizmatos788>

### ❑ ATIVIDADE 2

---

O setor financeiro de uma média empresa está ampliando suas atribuições e infraestrutura. Novos colaboradores acessarão os arquivos (documentos, planilhas, bancos de dados etc.) do setor através da rede local da organização.

O setor de TI disponibilizou um HD com sistema de arquivos compartilhado na LAN, mas ele está atingindo sua capacidade máxima de armazenamento a cada 60 dias.

Atualmente, o espaço utilizado é de 50 gigabytes (e em crescimento contínuo), com centenas de arquivos. Os usuários se queixam do tempo de resposta e da capacidade do disco. Como gerente do setor de TI, o que você poderia recomendar para melhorar a situação?

## ❑ ATIVIDADE 2

O setor financeiro de uma média empresa está ampliando suas atribuições e infraestrutura. Novos colaboradores acessarão os arquivos (documentos, planilhas, bancos de dados etc.) do setor através da rede local da organização.

O setor de TI disponibilizou um HD com sistema de arquivos compartilhado na LAN, mas ele está atingindo sua capacidade máxima de armazenamento a cada 60 dias.

Atualmente, o espaço utilizado é de 50 gigabytes (e em crescimento contínuo), com centenas de arquivos. Os usuários se queixam do tempo de resposta e da capacidade do disco. Como gerente do setor de TI, o que você poderia recomendar para melhorar a situação? Elabore o esboço da arquitetura relacionada à sua recomendação utilizando a ferramenta Dia Diagram Editor.

## A Case for Redundant Arrays of *Independent* Inexpensive Disks (RAID)

*David A. Patterson, Garth Gibson, and Randy H. Katz*

Computer Science Division  
Department of Electrical Engineering and Computer Sciences  
571 Evans Hall  
University of California  
Berkeley, CA 94720  
(patt@cs.berkeley.edu)

<http://web2.cs.columbia.edu/~junfeng/11sp-w4118/lectures/raid.pdf>

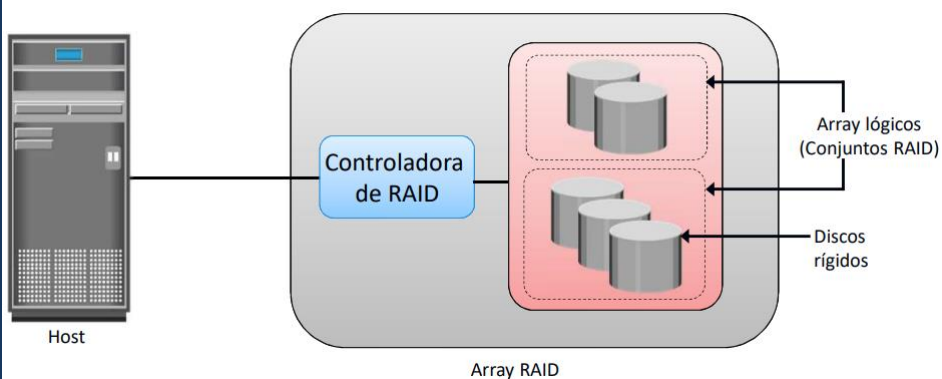
### Motivação

- ❑ Lidar com o crescimento dos volumes de dados das aplicações, sem que para isso fosse necessário adquirir novos discos rígidos (caríssimos!)
- ❑ *Single Large Expensive Disks (SLED)*

## Definição

- ❑ Uma forma de **organizar física** (ou logicamente) **discos de baixo custo** de maneira que proporcionem o **mesmo efeito (ou melhor)** do que os SLED e claro, de maneira mais barata.

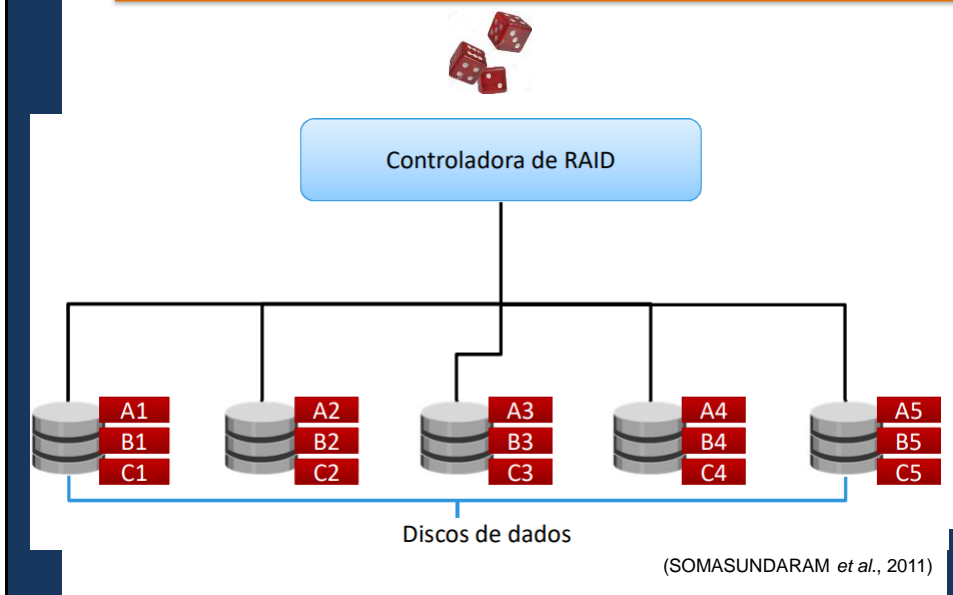
## Componentes



\*Array lógico: subconjunto de discos dentro do conjunto RAID

(SOMASUNDARAM *et al.*, 2011)

## RAID 0



## RAID 0

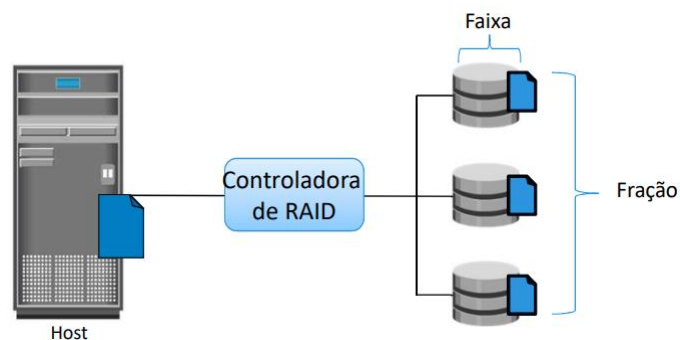
- ❑ Dados subdivididos em segmentos consecutivos que são escritos sequencialmente em cada um dos discos
- ❑ Leitura e escrita simultânea em cada disco
- ❑ Adequado para aplicações que precisam de altas taxas de transferência de I/O
- ❑ Sem tolerância a falhas (perdeu UM disco, perdeu tudo!)
- ❑ Exige 2 ou mais discos

## Striping

- ❑ **Strips:** número pré-determinado de blocos de discos endereçáveis contiguamente.
- ❑ **Striping:** técnica de distribuir os dados entre vários drives a fim de utilizá-los em paralelo.

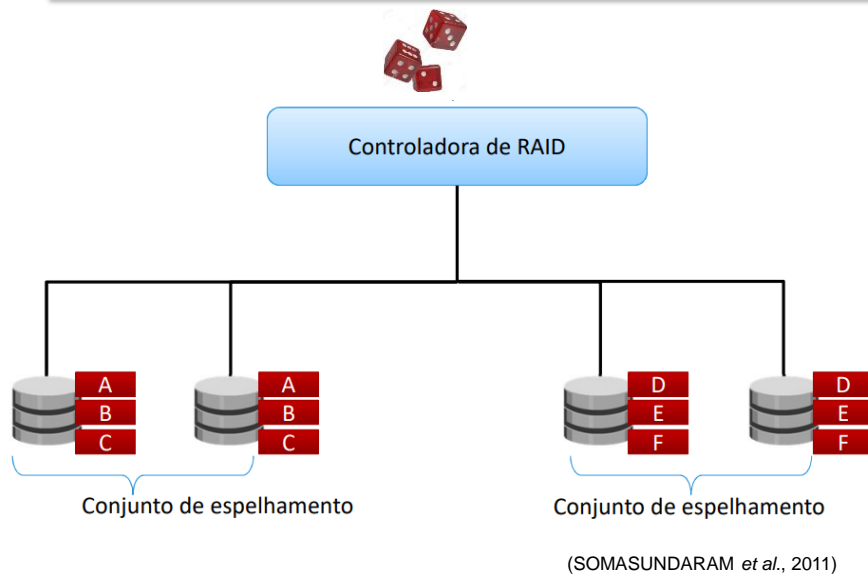
(SOMASUNDARAM *et al.*, 2011)

## Striping



(SOMASUNDARAM *et al.*, 2011)

## RAID 1



## RAID 1

- ❑ Dados replicados em 1 unidade de disco diferente
- ❑ Escrita em pares de unidades e leitura em todas as unidades simultaneamente
- ❑ Adequado para aplicações que requeiram alta disponibilidade
- ❑ Com tolerância a falhas (o + eficiente na recuperação)
- ❑ Exige pelo menos 2 discos



## Espelhamento

- ❑ Técnica de armazenamento dos dados em dois discos diferentes, produzindo duas cópias dos dados.
- ❑ Não é backup!

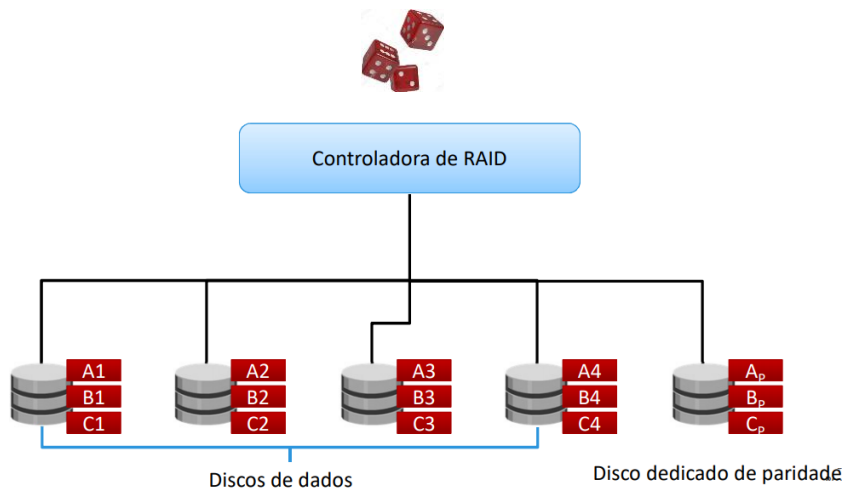
(SOMASUNDARAM *et al.*, 2011)

## RAID 2



- ❑ Escrita e leitura semelhante ao RAID 0, porém faz o cálculo do *Error Correcting Code* (ECC) e armazena o resultado em discos próprios
- ❑ Exige pelo menos 3 discos
- ❑ Tornou-se obsoleto após a inserção desse tipo de correção nos próprios discos

## RAID 3

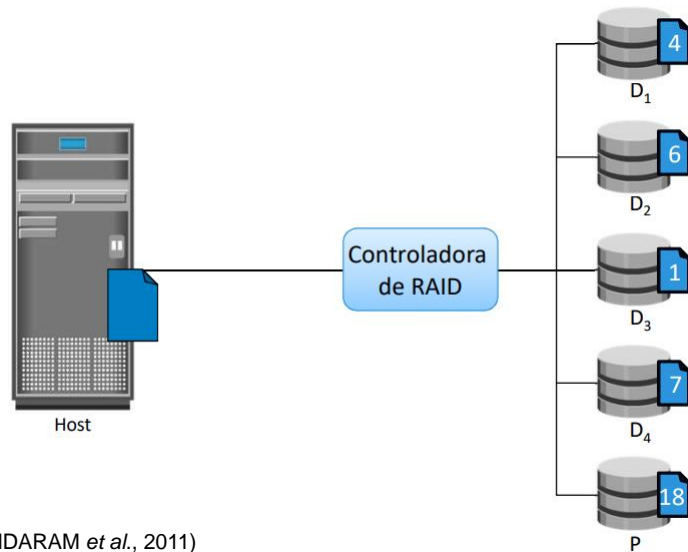


(SOMASUNDARAM *et al.*, 2011)

## RAID 3

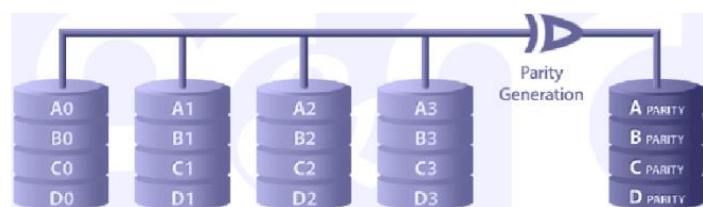
- ❑ Escrita e leitura semelhante ao RAID 0, porém faz o cálculo da paridade e armazena o resultado em disco dedicado
- ❑ Escrita em pares de unidades e leitura em todas as unidades simultaneamente
- ❑ Leitura e escrita + rápidas (exige a sincronia dos eixos)
- ❑ No caso de falhas, permite a correção de erros em nível de bit
- ❑ Exige pelo menos 3 discos

## Paridade



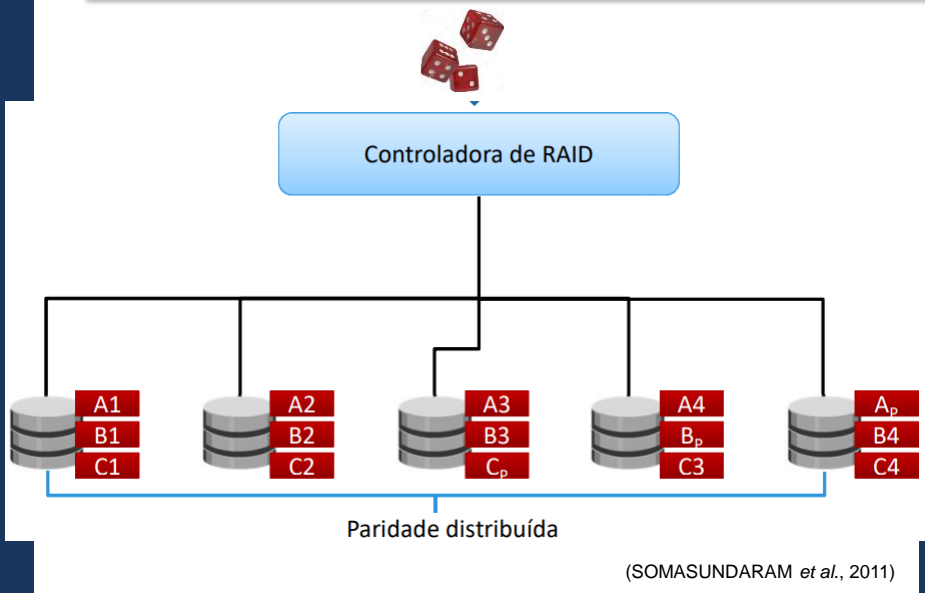
(SOMASUNDARAM *et al.*, 2011)

## RAID 4



- ☐ Paridade em nível de bloco
- ☐ Não exige leitura/gravação do striping inteiro (sem sincroniza dos eixos!)
- ☐ A cada nova escrita, atualiza a paridade, bloqueando próxima escrita
- ☐ Tornou-se obsoleto devido aos custos de recuperação

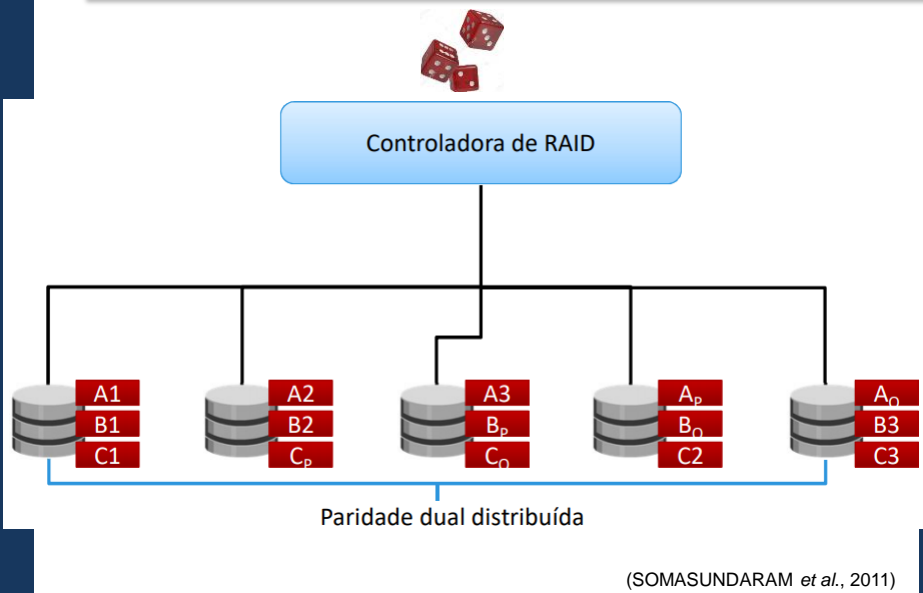
## RAID 5



## RAID 5

- ❑ Semelhante ao RAID 4, porém as informações de paridade são **distribuídas** entre os discos
- ❑ Não exige sincronização, somente partições do mesmo tamanho
- ❑ Boa escrita e ótima leitura (simultânea em todos os discos)
- ❑ No caso de falhas, permite a remontagem dos dados (mas, com overhead)

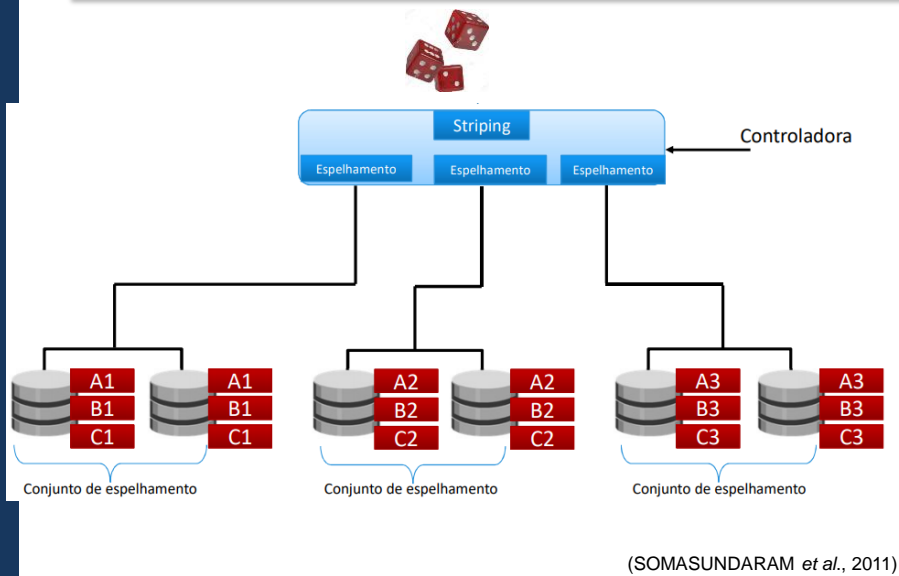
## RAID 6



## RAID 6

- ❑ Semelhante ao RAID 5, porém as informações de paridade são distribuídas e **duplicadas** entre os discos
- ❑ Suporta a falha de dois discos de um array
- ❑ Maior overhead na escrita e na remontagem dos dados
- ❑ Exige pelo menos 4 discos

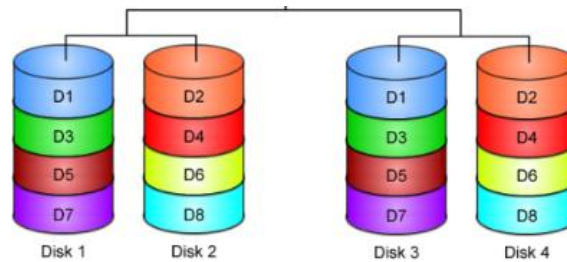
## RAID 1+0 (Mirror + Stripe)



## RAID 1+0 (Mirror + Stripe)

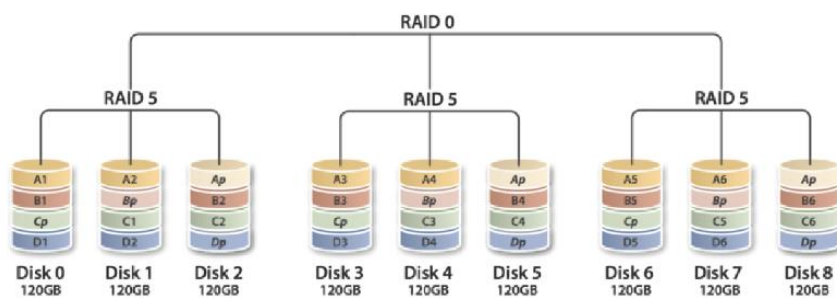
- ❑ Dados são primeiro espelhados para depois serem segmentados
- ❑ Mesma performance do RAID 1, porém, com tolerância a falhas
- ❑ Alto custo de expansão e sincronismo de velocidade para melhor performance
- ❑ Exige pelo menos 4 discos

## RAID 0+1 (Stripe+Mirror)



- ☐ Dados são primeiro segmentados e para cada segmento criado é feito um espelho
- ☐ Falhou 1 disco? Necessário espelhar todo um conjunto segmentado

## RAID 50

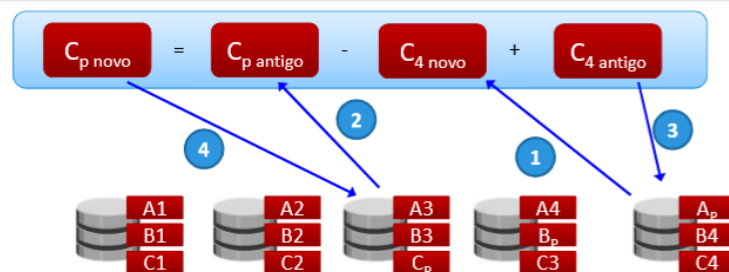


- ☐ Combina a paridade em conjunção com o striping
- ☐ Admite falhas em até 2 discos, desde que em conjuntos diferentes
- ☐ Exige pelo menos 6 discos

## Níveis do RAID

- ❑ **RAID 0** – Conjunto fracionados sem tolerância à falhas  
Striping – distribui dados em 2 ou + HD
- ❑ **RAID 1** – Espelhamento do disco (*mirror*)
- ❑ **RAID 2** – Checagem de erro (*Error Correcting Code*)
- ❑ **RAID 3** – Acesso paralelo e disco de paridade dedicado
- ❑ **RAID 4** – Blocos de dados para guardar registros
- ❑ **RAID 5** – Acesso ao disco independente e paridade distribuída
- ❑ **RAID 6** – Acesso ao disco independente e paridade dupla distribuída

## Impactos no desempenho



- ❑ **RAID 5**: toda atualização no disco exige 4 operações de I/O (2 leituras e 2 gravações no disco)
- ❑ **RAID 6**: toda atualização no disco exige 6 operações de I/O (3 leituras e 3 gravações no disco)
- ❑ **RAID 1**: toda gravação exige 2 operações de I/O (2 gravações no disco)

(SOMASUNDARAM *et al.*, 2011)



### Indicações conforme aplicação

- ☐ **RAID 1+0:** aplicativos com o perfil de I/O de gravações curtas, intensivas e aleatória (OLTP, SGBD - espaço temporário)
- ☐ **RAID 3:** leituras e gravações extensas e sequenciais (backup de dados e streaming multimídia)
- ☐ **RAID 5 e 6:** carga de trabalho pequena (e-mail, SGBD - entrada de dados)

(SOMASUNDARAM *et al.*, 2011)

### ☐ ATIVIDADE 3

- a) Explique por que o RAID 1 não substitui um backup.
- b) Explique por que o RAID 0 não é indicado para proteção e alta disponibilidade de dados.
- c) Uma organização está precisando em reconfigurar o armazenamento para seu aplicativo financeiro (contas a pagar e a receber) a fim de obter maior disponibilidade. 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 120 GB. O tamanho total dos dados do aplicativo é de 430 GB, que não serão modificados nos próximos 6 meses. Considerando que o final do ano está se aproximando e não é possível adquirir novos discos, recomende um nível de RAID que a empresa deva utilizar para reestruturar o ambiente e atender suas necessidades. Justifique sua escolha com base nos custos, desempenho e disponibilidade.

# TECNOLOGIAS

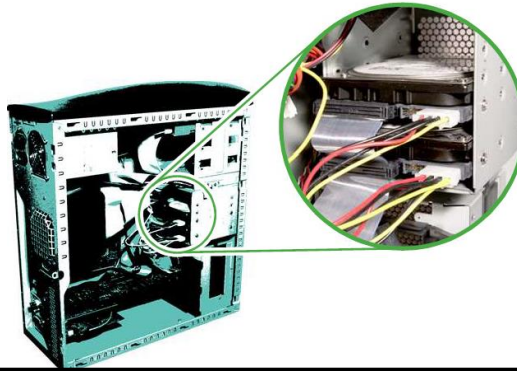
## Direct-Attached Storage (DAS)

- ❑ Dispositivos de armazenamentos estão ligados diretamente ao computador ou a um grupo de servidores em um cluster.
- ❑ Investimento inicial menor
- ❑ Configuração mais simples e mais rápida
- ❑ Limitações no armazenamento interno (número de slots, capacidade dos discos) gerando baixa escalabilidade

## Direct-Attached Storage (DAS)

### ❑ Interno

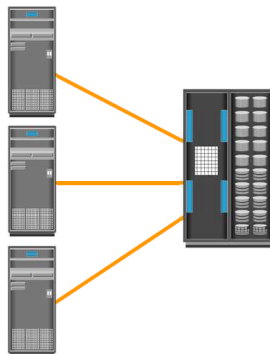
- ✓ Dispositivo de armazenamento conectado internamente ao host através de um barramento serial ou paralelo



## Direct-Attached Storage (DAS)

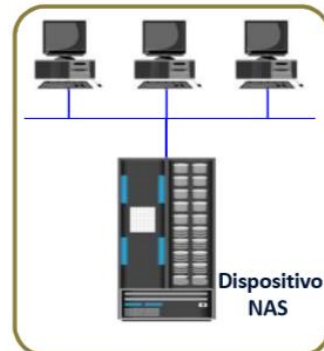
### ❑ Externo

- ✓ Servidor se conecta diretamente ao dispositivo de armazenamento externo



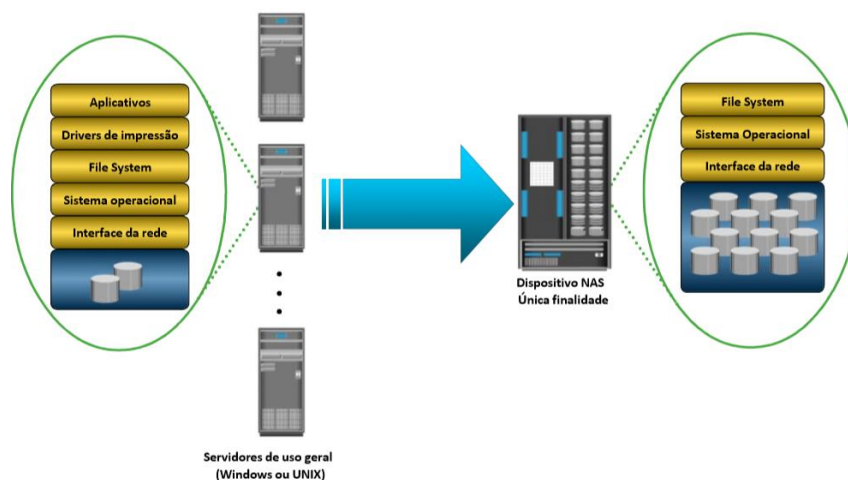
## Network-Attached Storage (NAS)

- ❑ Compartilhamento de disco que usa o TCP/IP e protocolos de sistemas de arquivos como *Network File System* (NFS) ou *Common Internet File System* (CIFS) para o acesso aos dados
- ❑ Utiliza SO especializado e otimizado para I/O de arquivos.



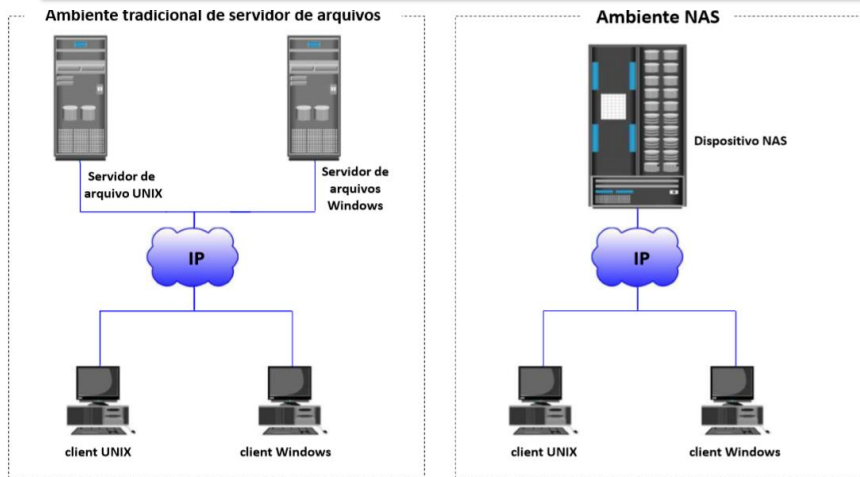
(SOMASUNDARAM *et al.*, 2011)

## DAS x NAS



(SOMASUNDARAM *et al.*, 2011)

## NAS e a consolidação de servidores

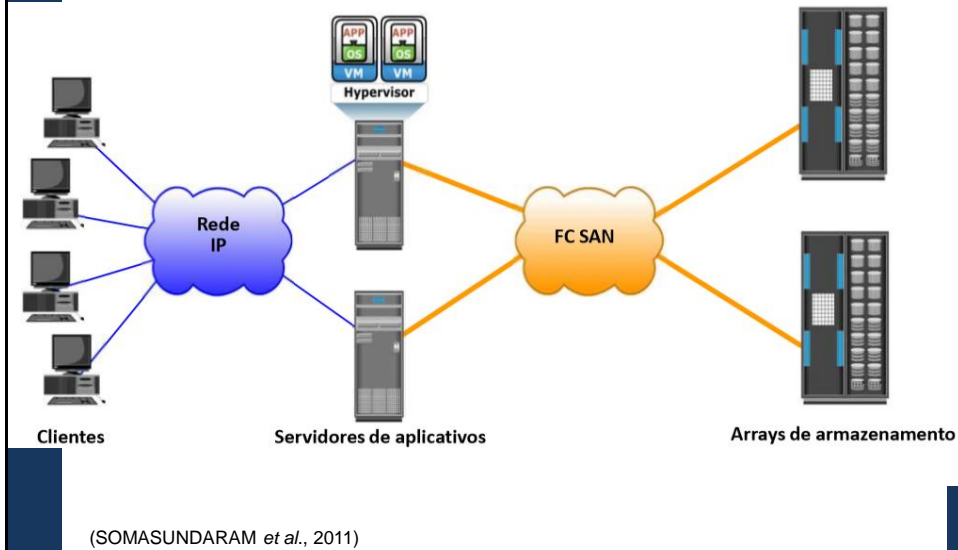


(SOMASUNDARAM *et al.*, 2011)

## Storage Area Network (SAN)

- ❑ Rede dedicada e de alto desempenho para facilitar a comunicação em nível de blocos entre os servidores e o armazenamento.
- ❑ Usa tecnologia de rede de alta velocidade executada em cabos de fibra óptica (rede *Fibre Channel* – FC)

## Storage Area Network (SAN)



## Storage Area Network (SAN)

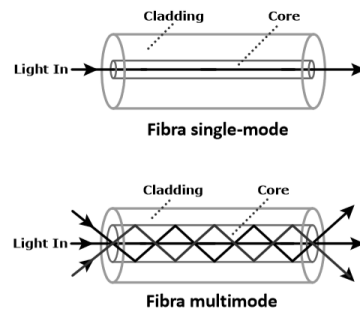
### Componentes:

- ☐ Portas dos nós (servidor e arrays de armazenamento)
- ☐ Cabos
- ☐ Conectores
- ☐ Switches e hubs
- ☐ Software de gerenciamento

(SOMASUNDARAM *et al.*, 2011)

## Storage Area Network (SAN)

Single-mode	Multimode
Leva um único feixe de luz	Pode carregar múltiplos feixes de luz simultaneamente
Distâncias até 10km	Utilizado para distâncias curtas (dispersão modal enfraquece a força do sinal após certa distância)



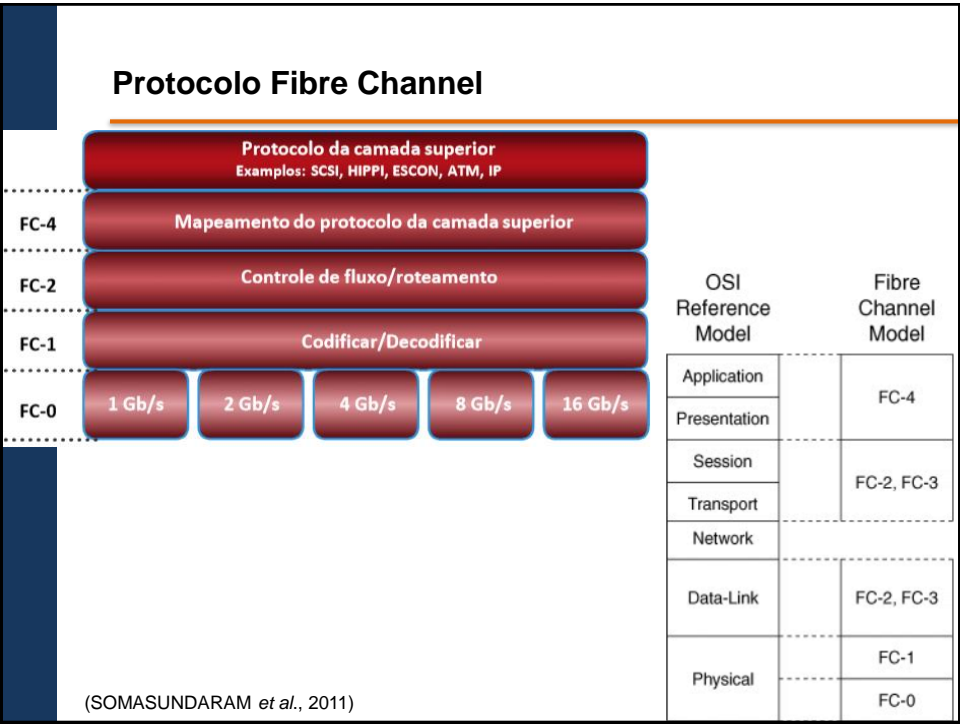
(SOMASUNDARAM *et al.*, 2011)

## Arquitetura Fibre Channel

### Componentes:

- ❑ Alto desempenho
- ❑ Alta escalabilidade
- ❑ Dispositivos de armazenamento ligados à SAN aparecem como se fossem locais aos hosts

(SOMASUNDARAM *et al.*, 2011)



❑ ATIVIDADE 4

Considerando ainda o cenário da Atividade 2, sendo que, o problema relatado pelo setor financeiro passou a ocorrer com frequência nos demais setores. Com base nisso, o Diretor de TI está analisando a viabilidade de implantação de uma infraestrutura de armazenamento escalável e de alta disponibilidade.

Que tecnologia e arquitetura de armazenamento você recomendaria (SAN, NAS, IP SAN) e por quê? Adote o projeto feito na Atividade 2 e expanda-o para este cenário.

Feito isso, sugira uma implantação de backup apropriada que minimize o tráfego na rede, evitando congestionamento e ao mesmo tempo não cause impacto nas operações de produção. Justifique sua resposta.