Dispositivos de Armazenamento em massa

José Roberto B. Gimenez

Estrutura da apresentação

- O meio físico de armazenamento
- Interfaces de conexão ATA, SCSI, FC
- RAID array
- Sistemas de Armazenamento DAS, NAS, SAN

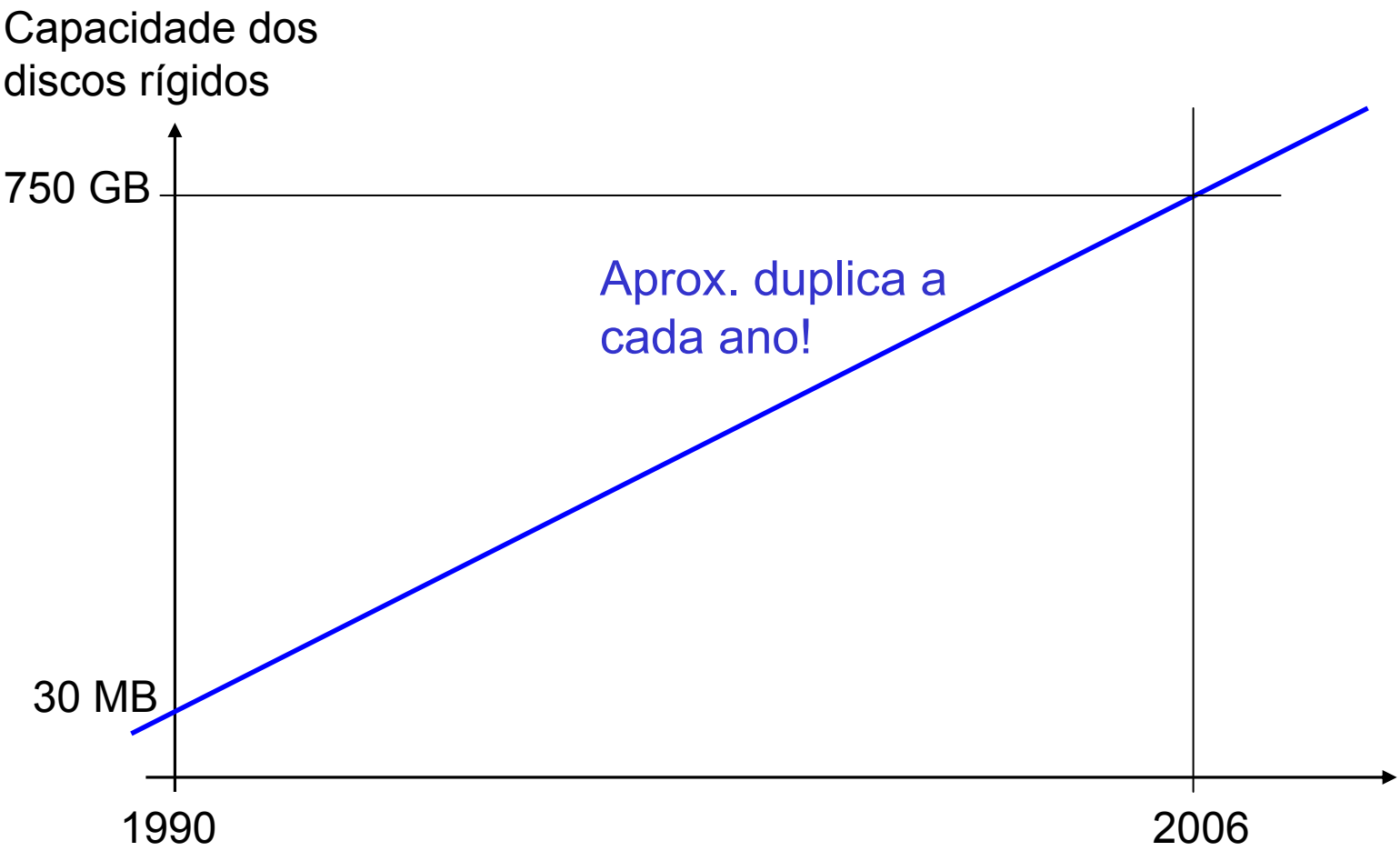
O meio físico

- Os dispositivos de Memória variam dependendo da aplicação: ROM, RAM, Flash memory, SD Card, tapes, floppys, CD, DVD, SM, etc;
- Características de alta capacidade, rapidez de leitura e escrita e baixo custo são fatores normalmente incompatíveis;
- O disco rígido ainda é o elemento que melhor combina características interessantes para funcionar como memória secundária.

Um pouco de história...

- O disco rígido já tem mais de 50 anos!
- O primeiro disco rígido foi o IBM 305 RAMAC (Random Access Method of Accounting and Control). lançado em 1956.
- Tinha uma capacidade de armazenamento de 5 MB.
- Custava cerca de USD 50,000.

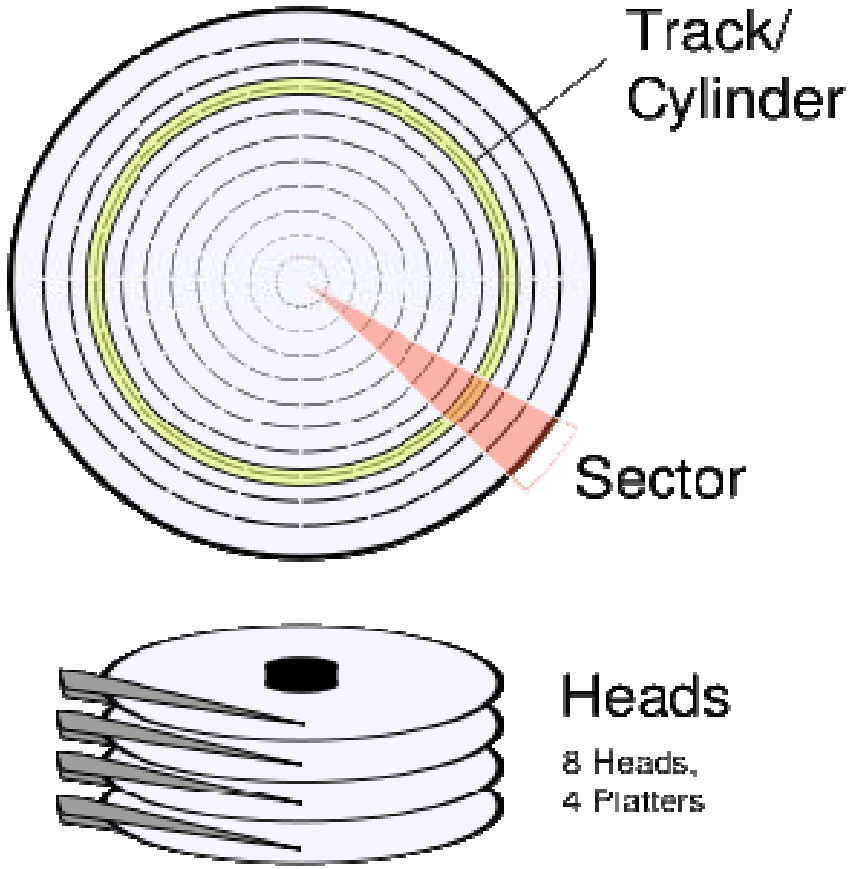
Evolução dos dispositivos de memória



Constituição de um disco rígido

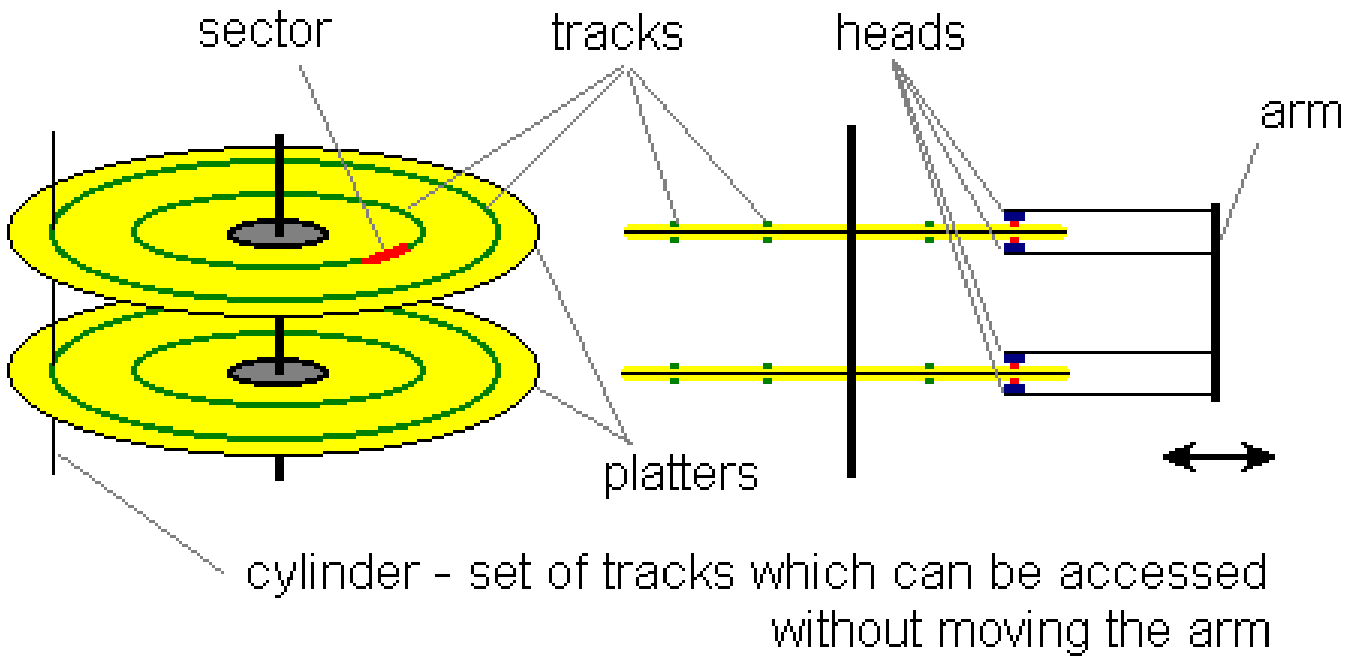
Capacidade de um disco rígido

= n° heads
x n° cilindros
x n° setores.



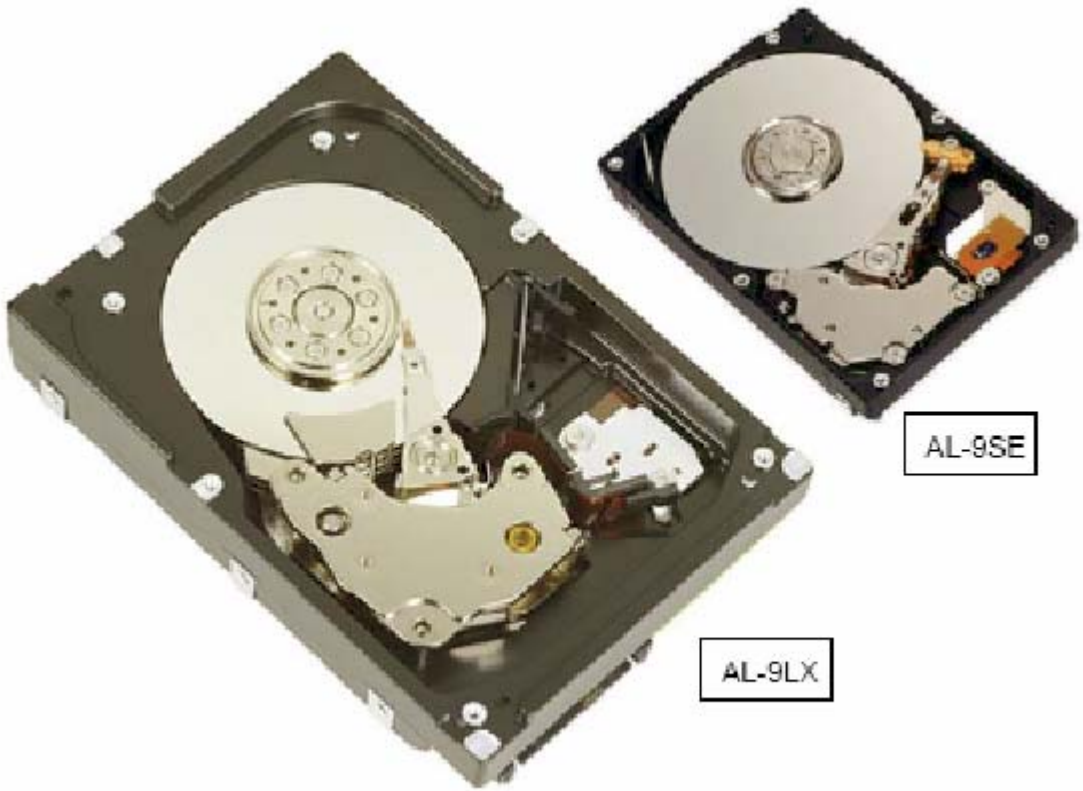
Características de acesso

Tempo médio de acesso = Tempo para encontrar uma trilha específica + ½ rotação do disco.



modelos de disco atuais

Formatos de 3,5“ e 2,5“



Parâmetros de um disco atual

Fujitsu AL-9 series	MAT3300	MAV2073
Formato	3,5”	2,5”
Interface	FC	SAS
dimensões	25,4 mm	15 mm
Capacidade	300 GB	73,5 GB
Velocidade	10.025 rpm	10.025 rpm
Diâmetro do disco	84 mm	65 mm
N° de platters	4	2
N° de Heads	8	4
Densidade de bits	28.500 b/mm	26.500 b/mm
Densidade de Tracks	4.100 T/mm	4.100 T/mm
Mean Seek Time	4,2 ms	4,0 ms
Potência	9,5 W	4,5 W
MTBF	1.200.000	1.400.000

Algumas continhas...

Capacidade armazenada em uma trilha (externa)
84 mm x π x 28.500 b/mm
= 940 KB

Taxa de transferência para esta mesma trilha
84 mm x π x 1025/60 x 28.500 b/mm
= 1,25 Gb/s

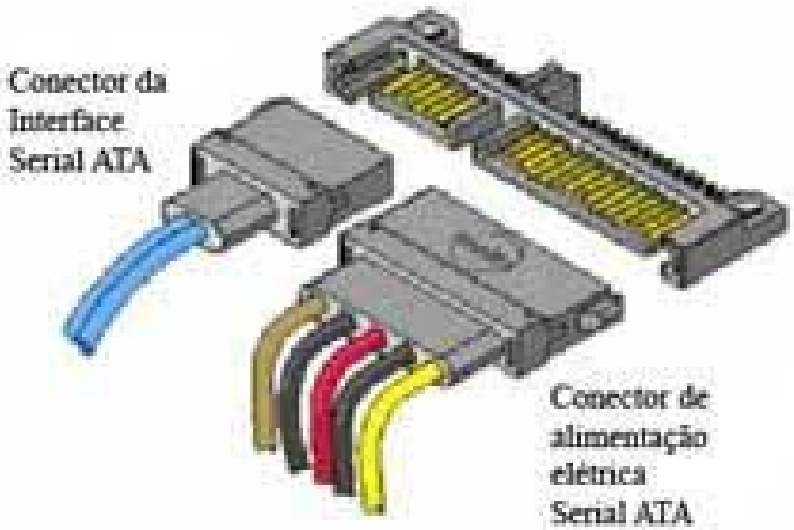
Interfaces de conexão

- IDE x SCSI
- SATA – Serial ATA – 1,5 Gb/s
- SAS – Serial Attached SCSI – 3 Gb/s
- FC – Fibre Chanel – 2 Gb/s

Conectores - SATA

Conectores e cabos

Aparência dos conectores Serial ATA

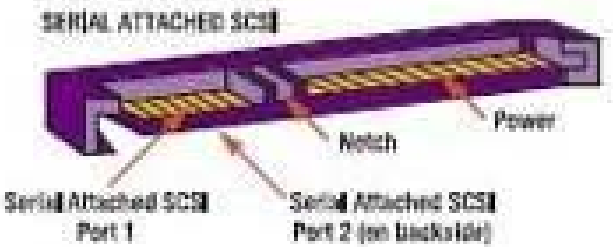


Drawing courtesy of Molex

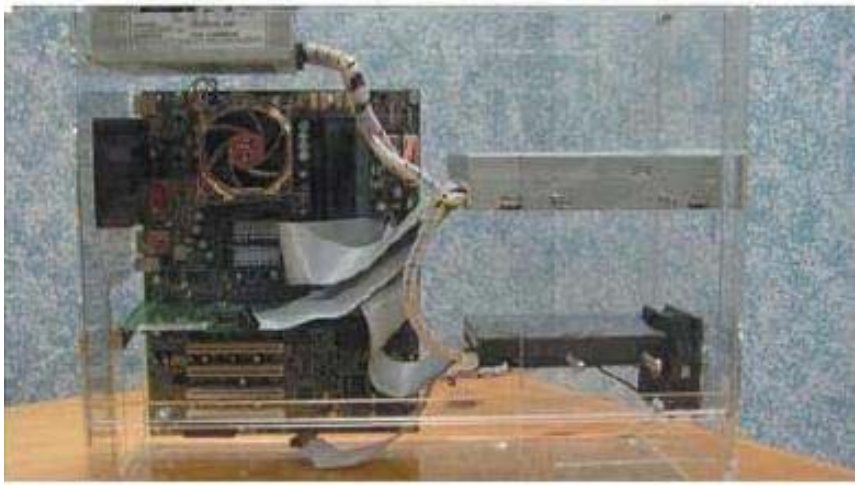
Localização e tamanho dos conectores

Conectores - SAS

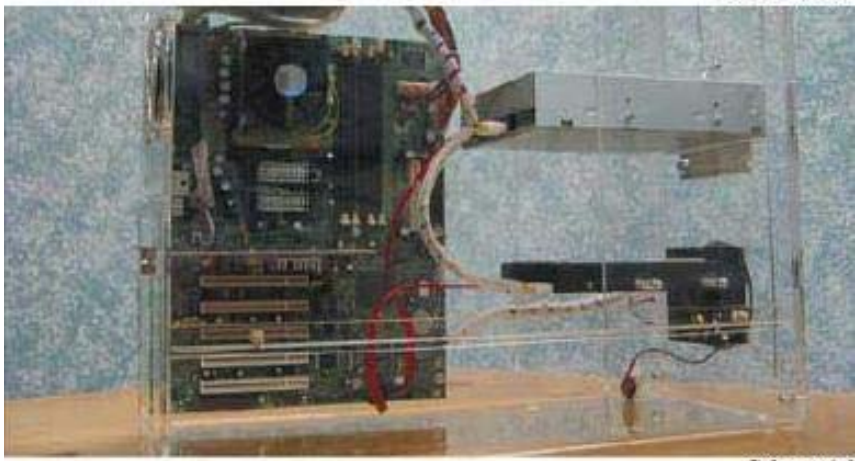
SERIAL ATTACHED SCSI BACKPLANE ACCEPTS SERIAL ATTACHED SCSI AND SATA DISK DRIVES



cabo serial e cabo paralelo

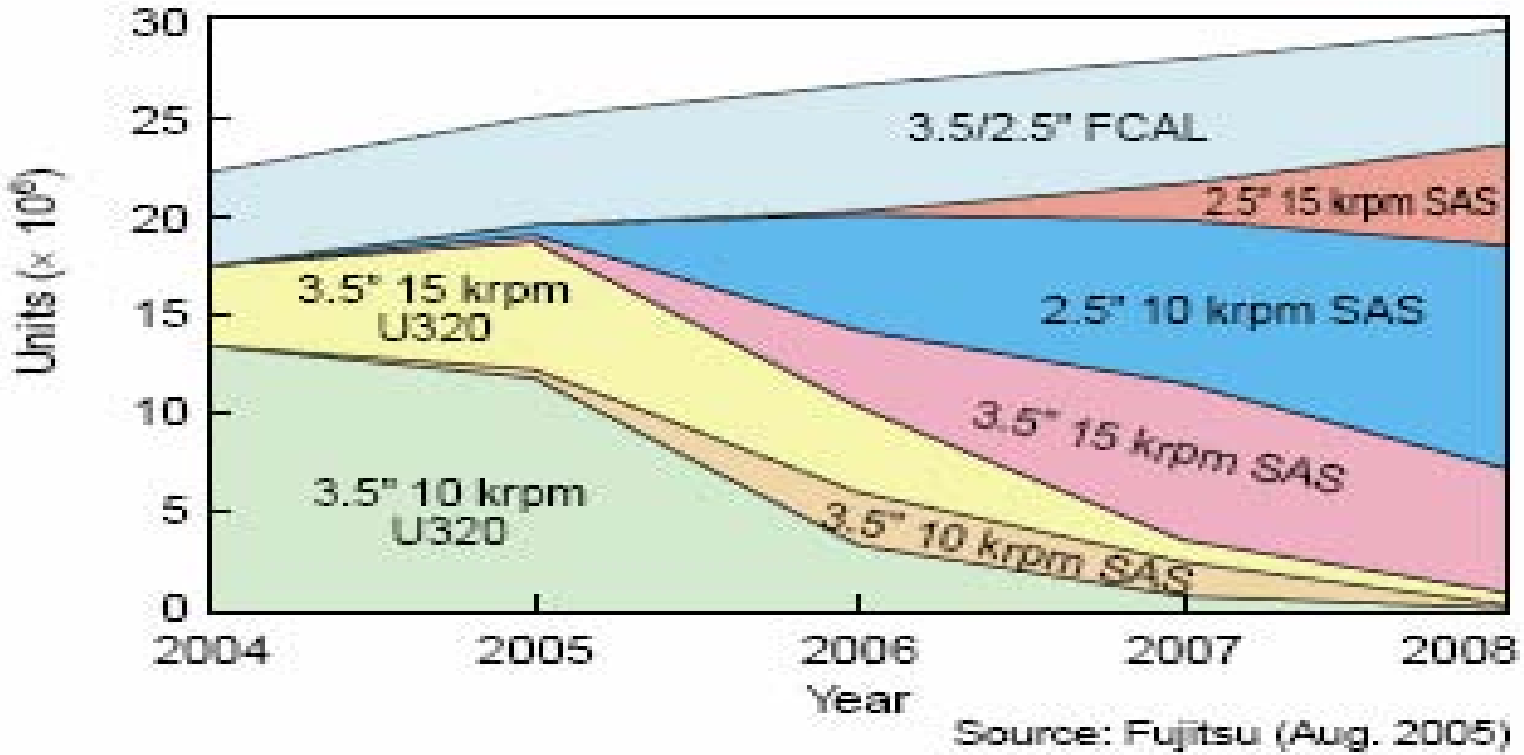


Cabo Paralelo



Cabo serial

Tendência de evolução para os HDs



FCAL: Fibre Channel Arbitrated Loop
SAS: Serial Attached SCSI
U320: Ultra 320 SCSI

RAID Redundant Arrays of Inexpensive Disks

- Técnica que associa vários discos;
- Criada para otimizar o fator custo;
- Atualmente usado para prover maior desempenho e confiabilidade;
- O “I” atualmente pode ser entendido como **independent**, ao invés de **inexpensive**.

Melhoria da confiabilidade por Redundância

- A chance de uma falha em um conjunto de N discos é maior que a de um único disco específico falhar. Ex: um sistema com 100 discos, cada um com MTBF de 100.000 horas (aprox. 11 anos), tem um MTBF de 1000 horas (aprox. 41 dias).
- **Redundância** – duplicar cada disco (cada operação de escrita é feita em ambos os discos);
- **Redundância** - armazenar bits extras de informação (bits de paridade).

Melhoria do Desempenho pelo Paralelismo

- Balancear os acessos para aumentar o throughput;
- Paralelizar acessos longos para reduzir o tempo de resposta
- Aumentar a taxa de transferência distribuindo os dados por vários discos.
- **Bit-level striping** – separar os bits por N discos
- Leitura de dados a uma taxa N vezes a de um único disco.
- O seek/access time é pior do que em um único disco.
- **Block-level striping** – Com N discos, o bloco i de um arquivo vai para o disco (i mod N) + 1.

RAID Levels

- Esquemas para prover redundância a baixo custo usando disk striping combinado com bits de paridade.
- Diferentes RAID levels possuem diferentes custos, desempenhos e características de confiabilidade.

RAID Level 0

- Striping no nível de blocos;
- Não redundante;
- Usado em aplicações de alto desempenho, onde a perda de dados não é fator crítico;
- Mínimo de 2 dois discos.

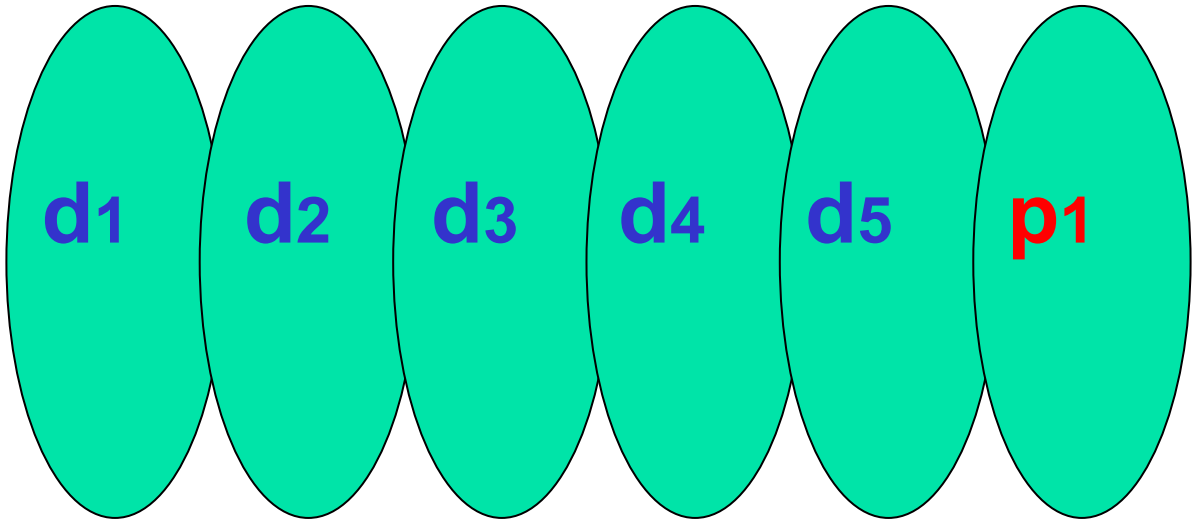
RAID Level 1

- Discos espelhados;
- Oferece o melhor desempenho de escrita;
- Popular para aplicações como armazenamento de arquivos de logs;
- Mínimo de 2 discos;
- Usualmente 2 discos.

RAID Level 2

- Similar aos códigos de correção de erros em Memoria, com bit striping.
- Um disco exclusivo para bits de paridade;
- Mínimo de 3 discos.

Esquema de paridade - detecção



0 + 0 = 0

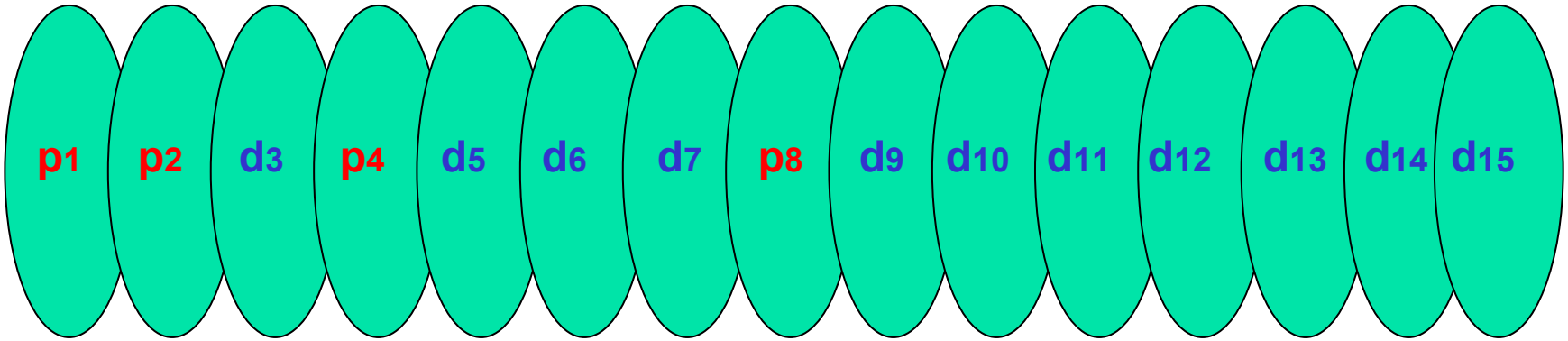
0 + 1 = 1

1 + 0 = 1

1 + 1 = 0

$p_1 = d_1 + d_2 + d_3 + d_4 + d_5$

Esquema de paridade - correção



$$p_1 = d_3 + d_5 + d_7 + d_9 + d_{11} + d_{13} + d_{15}$$

$$p_2 = d_3 + d_6 + d_7 + d_{10} + d_{11} + d_{14} + d_{15}$$

$$p_4 = d_5 + d_6 + d_7 + d_{12} + d_{13} + d_{14} + d_{15}$$

$$p_8 = d_9 + d_{10} + d_{11} + d_{12} + d_{13} + d_{14} + d_{15}$$

RAID Level 3

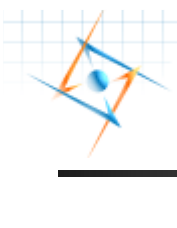
- Bit-Interleaved Parity; um único bit de paridade pode ser usado para a correção de erros, não apenas a deteção;
- Na escrita, os bits de paridade devem ser computados e gravados;
- Transferência de dados mais rápida que um único disco, mas menor I/Os por segundo;
- Superior ao Level 2;
- Mínimo de 3 discos

RAID Level 4

- Usa striping a nível de bloco;
- Mantém um disco exclusivo para os bits de paridade;
- Taxas de I/O para leituras de blocos independentes mais altas que Level 3;
- Altas taxas de transferência para leituras de blocos múltiplos;
- Blocos de paridade se tornam um gargalo para escrita de blocos independentes.

RAID Level 5

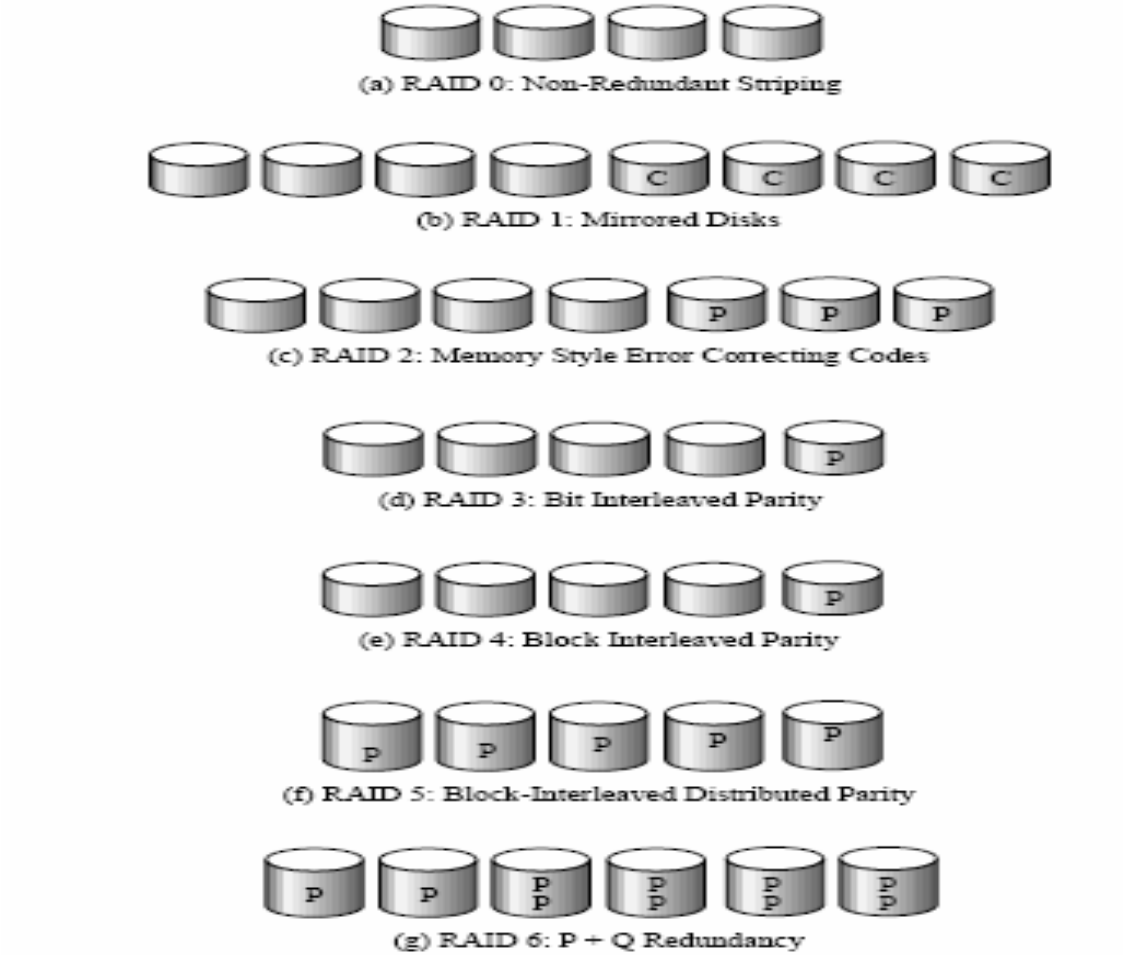
- Distribui dados e paridade por N + 1 discos, ao invés de armazenar dados em N e paridade em 1.
- Ex. com 5 discos: bloco de paridade é armazenado no disco (n mod 5) + 1, com os blocos de dados armazenados nos demais 4.
- Maior taxa de I/O que o Level 4. (Escrita de bloco ocorre em paralelo se os blocos de dados e seus blocos de paridades estiverem em discos diferentes.
- Superior ao Level 4



RAID Level 6

- Esquema de Redundância P+Q;
- Similar ao RAID Level 5;
- Armazena redundância extra como prevenção para múltiplas falhas;
- Maior confiabilidade que o Level 5;
- Não muito usado;
- Mínimo de 4 discos.

Sumário dos vários tipos de RAID



(a) RAID 0: Non-Redundant Striping

(b) RAID 1: Mirrored Disks

(c) RAID 2: Memory Style Error Correcting Codes

(d) RAID 3: Bit Interleaved Parity

(e) RAID 4: Block Interleaved Parity

(f) RAID 5: Block-Interleaved Distributed Parity

(g) RAID 6: P + Q Redundancy

Sistemas de Armazenamento

- DAS – Directed Attached Storage
- NAS – Network Attached Storage
- SAN – Storage Area Network