

STORAGE

ARMAZENAMENTO DE DADOS

Tutorial

4 out 2003
Fabio Montoro

A cada dia cresce, nas corporações, a necessidade de armazenamento de informações digitais, sejam elas textos, documentos escaneados, fotografias, clips, vídeos de segurança etc.

Os bancos de dados corporativos também lançam mão, cada vez mais, da imagem, fixa ou em movimento, para ilustrar produtos, procedimentos, etc, gerando mais demanda de memória.

O objetivo deste tutorial é somente dar uma introdução sobre a matéria, abordando os principais conceitos envolvidos.

CONCEITOS BÁSICOS

- Uma memória digital armazena bits
- 8 bits = 1 B (Byte)
- 1000 Bytes = 1 kB (kilo)
- 1000 kB = 1 MB (Mega)
- 1000 MB = 1 GB (Giga)
- 1000 GB = 1TB (Tera)
- **dpi** = *dots per inch* (pontos por polegada). Medida da definição de uma imagem digitalizada, linearmente, em pontos por polegada. Um equipamento fax comum (grupo 3) possui definição de 203 dpi na horizontal e 97,5 dpi na vertical. Uma impressora, Laser ou jato de tinta, ou um scanner, normalmente possui a definição de 300 dpi.
- **Pixel** = **pel** = *picture element*. É a quantidade de pontos na imagem. Quanto maior esse número, melhor a resolução (melhor a qualidade) e mais espaço ocupa de memória. A qualidade pode ser medida pela relação pixel/área.

- A memória necessária para armazenar uma imagem vai depender de quantos bits são utilizados para cada ponto de imagem (bit/pixel).
- 1 foto colorida com 640x480 pixels, em baixa definição ocupa cerca de 64 kB
- 1 foto colorida com 640x480 pixels, em média definição ocupa cerca de 144 kB
- **NTSC** = padrão de vídeo analógico, estabelecido em 1953, que utiliza 640 pixels por linha, 525 linhas por quadro, 30 quadros por segundo.
- **SDTV** = padrão de vídeo digital standard. Utiliza 640 pixels na horizontal, 480 pixels na vertical, 24 bits por pixel, e 30 quadros por segundo.

Para se ter ideia da necessidade de memória, seguem alguns exemplos de quanta memória ocupa algumas informações sob a forma digital:

- 64 kByte = Foto colorida 640x480 pixels
- 25 MByte = um livro sem figuras, com 1000 páginas (aproximadamente)
- 900 Mbyte = 24 horas de vídeo de vigilância digitalizado MPEG

Por exemplo, um servidor de Storage com capacidade de 1 Tbyte, conseguiria armazenar, grosso modo, o seguinte:

- 4 mil livros com 1000 páginas + 100 dias de vídeo-monitoração

Quando o storage é acessado pela rede Ethernet recebe o nome de NAS – *Network Attached Storage*.

Uma das características mais importantes de um NAS é a sua estratégia de armazenamento, definida como RAID (*Redundant Array of Independent Disks*):

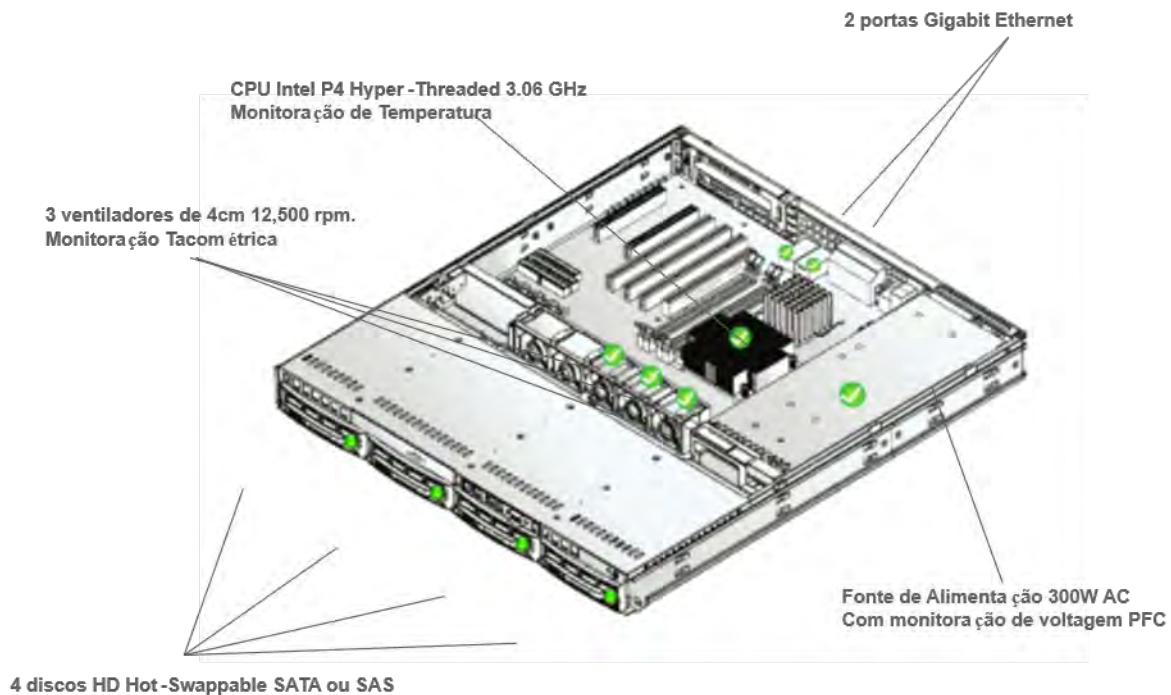
- RAID 0 = armazena em vários HDs mas sem redundância, ficando sujeito a falhas. Excelente desempenho. Usa 100% da capacidade total em disco.
- RAID 1 = armazena a informação duas vezes, em discos diferentes (espelhamento), ganhando boa imunidade a falhas. Usa só 50% da capacidade total em disco. Bom desempenho. Recuperação fácil
- RAID 5 = armazena usando paridades por blocos. Mais eficiente que o RAID 1, mas precisa de pelo menos 3 HDs. Recuperação mais demorada. Com 3 HDs usa 66% da capacidade total em disco. Com 4 HDs (mais comum), usa 75% da capacidade total em disco

Como exemplo de equipamento, a seguir estão as especificações de um Storage fabricado pela AMI, modelo StoreTrends® 1100i:



- sistema operacional baseado em LINUX. Muitos especialistas consideram o Linux mais rápido e mais seguro que o Windows. De fato, esse equipamento é duas vezes mais rápido que os similares concorrentes, segundo testes da fábrica.
- próprio para ser instalado em rack de 19 polegadas e ocupa apenas uma unidade de altura (1 U), que corresponde a 44,45 mm.
- quantidade de Discos Rígidos (HD): 4 drivers do tipo ATA (IDE). Os discos são encaixados pela parte frontal do equipamento.
- capacidade máxima de memória: 1 Tbyte com 4 discos de 250 Gbyte e 4 Tbyte com discos de 1 Tbyte.
- processador: Intel Pentium 4 @ 3 GHz
- duas portas Gigabit (1 Gbps)
- redundância: RAID 0, 1 e 5.
- protocolos de rede suportados: TCP/IP, IPX e Apple Talk
- protocolos de arquivo suportados: Microsoft CIFS, Unix NFS, Apple AFP, NCP, HTTP 1.0 (somente leitura), FTP (leitura e escrita).
- Permite definir espaços máximos de uso por usuário.





Necessidade de armazenamento para um sistema de vídeo SDTV

$$M = Pixels_{hor} \cdot Pixels_{vert} \cdot \frac{bits}{Pixel} \cdot \frac{Quadros}{s} \cdot T(horas) \cdot \frac{60^2}{8 \cdot 10^9}$$

$$M = Pixels_{hor} \cdot Pixels_{vert} \cdot \frac{bits}{Pixel} \cdot \frac{Quadros}{s} \cdot T(horas) \cdot \frac{450}{10^9}$$

M = memória [GB]

Utilizando os dados da introdução, constatamos que a taxa de transferência seria de 221 Mbps e concluímos que seriam necessários 99,5 GB de memória para armazenar uma hora de vídeo digital SDTV.

Não se armazena vídeo no formato SDTV.

Essa quantidade enorme de memória pode ser reduzida quando se utiliza sinal de vídeo codificado, com compressão, segundo padrões internacionais como o MPEG e gravação somente se houver movimento na cena.

O padrão MPEG possui diversos níveis de qualidade, para atender desde a TV de alta definição até a vídeo-monitoração dos sistemas de vigilância.

O padrão MPEG-1 oferece a mesma qualidade do VHS com uma taxa de apenas 1,5 Mbps e foi criado com o objetivo de armazenar vídeos em CDs.

O padrão MPEG-2 oferece mais definição, permitindo escolher entre 2 e 100 Mbps.

O padrão MPEG-4 foi criado para armazenar vídeos de menor qualidade e ocupar menos memória.

Necessidade de armazenamento para um sistema de vigilância

$$M = \frac{L_{frame} (kB)}{10^6} \cdot V_{frame} (frames / s) \cdot T (horas) \cdot 60^2$$

$$M = L_{frame} \cdot V_{frame} \cdot T \cdot 0,0036$$

M = memória [GB]

L_{frame} = tamanho do quadro [kB]

V_{frame} = velocidade de quadros [frames por seg]

T = tempo de gravação [horas]

Exemplo: quanta memória deve ter um sistema de armazenamento para gravar imagens, durante 10 dias, de um sistema de vigilância MPEG, com 16 câmeras digitais, cada uma enviando 2 quadros por segundo, cada quadro com 5 kBytes:

$$M = L_{frame} \cdot V_{frame} \cdot T \cdot 0,0036 = (5) \cdot (16 \cdot 2) \cdot (24 \cdot 10) \cdot 0,0036 = \underline{\underline{138,24 GBytes}}$$

Parâmetros de performance de um sistema de armazenamento

Tempo médio de rotação do HD:

$$T_r = \frac{0,5}{\frac{V}{60}} = \frac{30}{V}$$

V = velocidade de rotação do HD [RPM]

Tempo de transferência dos dados:

$$T_{tx} = \frac{B}{v_{tx}}$$

Tempo médio de acesso ao disco:

$$T_a = T_r + T_b + T_{tx} + T_{ov}$$

T_r = tempo médio de rotação

T_b = tempo médio de busca

T_{tx} = tempo de transferência dos dados

T_{ov} = tempo médio de overhead do controlador

Exemplo: qual o tempo médio de acesso a um disco com as seguintes características:

- 10.000 RPM
- Tempo de busca = 5 ms
- Blocos de 512 kBytes
- Overhead do controlador = 0,1 ms

$$T_a = \frac{30}{10000} + 5 \cdot 10^{-3} + \frac{512}{40 \cdot 10^6} + 0,1^{-3} = \underline{\underline{8,11ms}}$$

- 0 - 0 - 0 -