

Tecnologias de Discos Magnéticos

Vinícius Garcia de Oliveira
R.A: 041179
vgarcia@cpqd.com.br

Ricardo Correa
R.A: 107101
correa.ricardo@gmail.com

Resumo

Este artigo objetivo descrever sobre a tecnologia de discos magnéticos e sua evolução ao longo da história. É apresentado o princípio de funcionamento básico de um disco magnético, assim como novas técnicas que estão possibilitando o crescimento contínuo na capacidade de armazenamento. As principais interfaces utilizadas para o acesso e armazenamento das informações também são introduzidas.

Por fim, é realizada uma comparação do disco magnético com o seu principal concorrente, o SSD. De forma a entender se os discos magnéticos continuarão a ser dominantes como meio de armazenamento não volátil nos sistemas de computação nos próximos anos.

Palavras-chave

disco magnético; memória de massa; memória secundária; SCSI; IDE; ATA; Fibre Channel; SSD.

1. Introdução

Discos magnéticos, também conhecidos como discos rígidos, memória de massa ou ainda memória secundária é a parte do computador onde são armazenados os dados de forma persistente [1]. O disco rígido é uma memória não-volátil, ou seja, as informações não são perdidas quando o computador é desligado, sendo considerado o principal meio de armazenamento de dados em massa. Dada essa característica é um meio para realizar, reincidentemente, a execução de programas e carregar arquivos contendo os dados inseridos anteriormente à última inicialização do computador.

Desta forma, os discos magnéticos desempenham dois papéis importantes nos sistemas de computadores. (1) Armazenamento não-volátil de longo prazo para arquivos, mesmo quando nenhum programa esteja em execução. (2) Um nível de hierarquia de memória abaixo da memória principal, usado como apoio para a memória virtual durante a execução de programas.

A tecnologia de discos magnéticos está presente nos sistemas de computação desde meados da década de 50, sendo proposto originalmente pela IBM [2]. Desde então muitas tecnologias tentaram tomar o lugar dos discos magnéticos como principal meio de armazenamento de memória de massa, entre elas destacam-se: as fitas magnéticas, os discos ópticos, e, mais recentemente, as memórias de estado sólido (SSD – Solid State Disc). Com exceção desta última, que se tornou dominante nos sistemas embarcados, e em aplicações de mais baixa capacidade, nenhuma tecnologia ameaçou historicamente a hegemonia dos discos magnéticos na função de memória secundária nos sistemas de computação. Isso se deve principalmente a dois fatores: (1)

constante evolução na capacidade de armazenamento, (2) redução do custo por bit a cada nova geração da tecnologia.

É importante destacar que essa evolução não foi constante ao longo das últimas décadas. Por exemplo, em 1990 os discos magnéticos aumentavam a sua capacidade de armazenamento a uma média de 30% ao ano, enquanto as memórias DRAM apresentavam uma média de 60%. Nesta época as memórias DRAM apresentavam um custo por gigabyte 20 vezes mais que os discos magnéticos. Caso essas médias de crescimento se mantivessem ao longo do tempo as memórias DRAM teriam substituído os discos magnéticos no início dos anos 2000. Entretanto, a partir da década de 90 a relação de custo por gigabyte caiu abruptamente, o que impediu que tal substituição ocorresse. O custo por gigabyte dos discos magnéticos no período entre 1983 a 2001 é mostrado na figura 1.

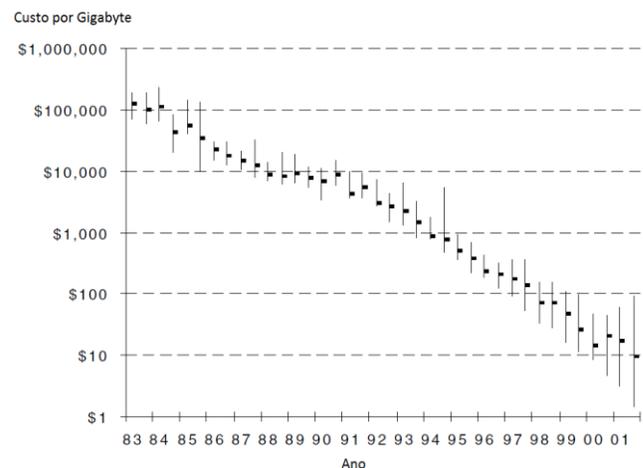


Figura 1. Evolução do custo por gigabyte.

Em função da importância dos discos magnéticos nos sistemas de computadores este trabalho visa apresentar essa tecnologia. O trabalho é dividido como descrito a seguir. A seção 1 realiza uma introdução sobre o assunto e apresenta a estrutura do trabalho. A seção 2 mostra a estrutura básica de um disco magnético e o seu princípio de funcionamento. A seção 3 apresenta as interfaces mais comuns dos discos. A seção 4 discursa sobre o futuro dos discos magnéticos. A seção 5 apresenta as conclusões do trabalho e a seção 6 as referências.

2. Estrutura do disco magnético

Um disco magnético consiste em uma coleção de lâminas, que variam em geral de 1 a 12 unidades, girando em um eixo em 3.600 a 15.000 rotações por minuto (RPM). Essas lâminas, que recebem o nome de pratos, são discos de metal ou vidro

recobertos com um material magnético de gravação em ambos os lados, de forma que 10 pratos apresentam 20 superfícies para a gravação. Em 2001, os diâmetros dos pratos variavam por quase 4 vezes, indo desde 1,0 até 3,5 polegadas, embora mais de 95% das vendas sejam de discos com pratos de 2,5 ou 3,5 polegadas. Os discos de 2,5 visam o mercado de computadores portáteis enquanto os de 3,5 os computadores de mesa e servidores. Tradicionalmente, os discos com pratos de maior diâmetro têm desempenho mais alto, enquanto os discos menores têm preço mais baixo. Com frequência, o preço por gigabyte se refere aos discos vendidos com maior volume, que hoje são os de 3,5 polegadas. A superfície do prato se divide em círculos concêntricos, denominados trilhas. Em geral, existem de 5.000 a 30.000 trilhas em cada superfície. Ao conjunto de trilhas sobrepostas em superfícies diferentes se dá o nome de cilindro. Por sua vez, cada trilha é dividida em setores que contêm as informações. Uma trilha pode ter de 100 a 500 setores. Um setor é a menor unidade que pode ser lida ou gravada. Os *mainframes* da IBM, por exemplo, permitem o usuário selecionar o tamanho dos setores, embora a maioria dos sistemas fixe seu tamanho, normalmente em 512 bytes de dados. A sequência registrada na mídia magnética é um número de setor, um intervalo, as informações para esse setor incluindo código de correção de erros, um intervalo, o número de setor do próximo setor e assim por diante. A figura 2 apresenta esse relacionamento.

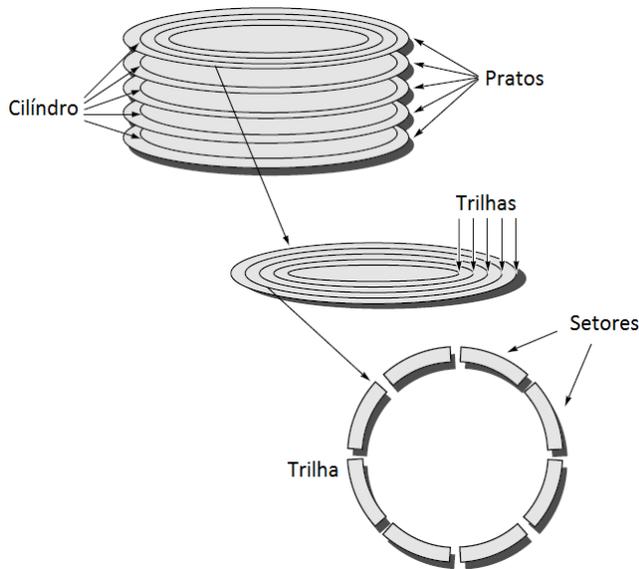


Figura 2. Relacionamento entre pratos, cilindros, trilhas e setores.

No passado, todas as trilhas tinham o mesmo número de setores; as trilhas mais externas, mais longas, registravam informações a uma densidade muito mais baixa que as trilhas mais internas. A gravação de mais setores nas trilhas mais externas que nas trilhas mais internas, o que se chama de densidade de bits constante, é o padrão atual. Esse nome é equivocado pois na realidade a densidade de bits não é constante. Normalmente as trilhas mais internas são gravadas a densidades mais altas e as trilhas mais externas a densidades mais baixas. Por exemplo, uma trilha mais externa pode possuir 2 vezes o tamanho de uma trilha mais interna, enquanto a sua capacidade pode ser apenas 1,7 maior.

Nos primeiros discos rígidos, eram usados motores de passo para movimentar os braços e cabeças de leitura [3]. Porém, além de muito lentos, eles eram muito susceptíveis a problemas de

desalinhamento, além de não serem muito confiáveis. Os discos atuais utilizam um mecanismo mais sofisticado para esta tarefa. O atuador, composto por um dispositivo que atua através de atração e repulsão eletromagnética possui dois eletroímãs, um de cada lado do braço móvel. Alterando-se a intensidade da corrente elétrica e, conseqüentemente, a potência de cada eletroímã, o braço e as cabeças de leitura se movimentam. A figura 3 apresenta uma foto de um disco rígido onde estes componentes são apresentados.



Figura 3. Estrutura física de um disco rígido.

Todo o funcionamento do disco rígido, a movimentação da cabeça de leitura, a velocidade de rotação, a leitura e gravação dos dados, o envio e recebimento de dados através da interface é coordenado pela placa controladora, ou controlador. Nos HDs mais antigos a placa controladora era uma placa separada, conectada a um *slot* ISA e ligada ao HD por dois cabos de dados. Este arranjo era muito ineficiente pois a distância tornava a comunicação muito susceptível a interferências e corrupção de dados. A partir do advento dos discos IDE, a placa controladora passou a fazer parte do próprio disco rígido. Apesar de pequena, a placa controladora de um disco atual é mais sofisticada do que um computador relativamente antigo, como os computadores com processador Intel 286 da década de 90.

Para ler ou gravar um setor, o controlador de disco envia um comando para mover o braço sobre a trilha apropriada. Essa operação é chamada busca, ou seek, e o tempo necessário para mover o braço até a trilha desejada denomina-se tempo de busca. O tempo médio de busca é assunto de consideráveis inequívocos. Os fabricantes de discos divulgam o tempo mínimo de busca, o tempo máximo de busca e o tempo médio de busca em seus manuais. Os dois primeiros são fáceis de medir, mas o tempo médio permitia ampla interpretação. A indústria decidiu calcular o tempo médio de busca como o somatório do tempo para todas as buscas possíveis, dividido pelo número de buscas possíveis. Os tempos médios de busca são anunciados em geral entre 5ms a 12ms. Porém, dependendo da aplicação e do sistema operacional, o tempo médio de busca real pode ser apenas 25% a 33% do número anunciado. A explicação é a localidade das referências de disco. O tempo para o setor solicitado girar sob a cabeça é chamado latência de rotação ou retardo rotacional. Já o tempo de transferência é o tempo necessário para transferir um bloco de bits, em geral um setor, sob a cabeça de leitura/gravação. Esse tempo é uma função do tamanho do bloco, do tamanho do disco, da velocidade de rotação, da densidade de gravação da trilha e da velocidade dos componentes eletrônicos que conectam o disco ao computador.

Existem muitos parâmetros que caracterizam um disco rígido. A tabela 1, retirada de [4], exibe as características disponíveis comercialmente do disco ST173404LC Ultra 160 do fabricante Seagate. Não são apresentadas nesta tabela as informações detalhadas referentes à interface de acesso. Essa questão será abordada em detalhes na seção 3.

Tabela 1. Principais parâmetros dos discos rígidos.

Características	Unidade SCSI Seagate Cheetah ST173404LC Ultra 160
Diâmetro do disco (polegadas)	3,5
Capacidade de dados formatados (GB)	73,4
Clindros	14.100
Discos	12
Superfícies de gravação (ou cabeças)	24
Bytes por setor	512 a 4.096
Setores médios por trilha (512 bytes)	~ 424
Densidade máxima de área (Gb/pol. quad.)	6,0
Velocidade de rotação (RPM)	10.033
Tempo médio de busca aleatória de cilindro para cilindro (leitura/gravação) (ms)	5,6/6,2
Tempo mínimo de busca (leit./grav.) (ms)	0,6/0,9
Tempo máximo de busca (ms)	14,0/15,0
Taxa de transferência de dados (MB/s)	27 a 40
Veloc. de link para buffer de disco (MB/s)	160
Energia, ociosa/em operação (W)	16,4/23,5
Tamanho do buffer (MB)	4,0
Tamanho: altura x largura x prof. (polegadas)	1,6 x 4,0 x 5,8
Peso (libras)	2,00
MTBF nominal (horas ligadas)	1.200.000
Porcentagem de horas ligadas por mês (POH)	100%
Porcentagem de POH de busca, leit. e grav.	90%
Erros de leitura irrecuperáveis por bits lidos	< 1 por 10^{15}
Erros de busca	< 1 por 10^7
Toler. a choques: operando e não-operando	10G / 175G
Toler. a vibrações: operando e não-operando (curva senoidal, de 0 a pico)	5 a 400Hz @ 0,5 G / 22 a 400Hz @ 2G

2.1 Gravação dos dados

Como descrito anteriormente, o processo de gravação dos bits ocorre através da magnetização do material ferromagnético depositado na superfície dos discos. Discos mais antigos utilizam óxido férrico como superfície de gravação, enquanto discos mais novos utilizam liga à base de cobalto. Mudanças sequenciais na direção da magnetização representam padrões de bits que carregam a informação armazenada. A cabeça de leitura e escrita é capaz de detectar esses padrões no caso de operações de leitura e mudar a magnetização no caso de escritas.

A superfície magnética de cada disco é conceitualmente dividida em regiões chamadas de domínios magnéticos que possuem área menor que 1 micrometro quadrado. Nos discos mais antigos as regiões eram magnetizadas de forma horizontal ou paralela à superfície do disco. No início de 2005 os discos começaram a ser gravados com uma orientação perpendicular, que permite domínios magnéticos menores, aumentando assim a densidade de bits por área. A figura 4 apresenta o conceito.

Devido a natureza policristalina do material magnético, cada região é composta de algumas centenas de grãos magnéticos ou cristalites, que possuem tamanho típico de 10nm. Assim, uma região magnética é composta de grãos que quando alinhados

formam um dipolo magnético que carrega a informação armazenada.

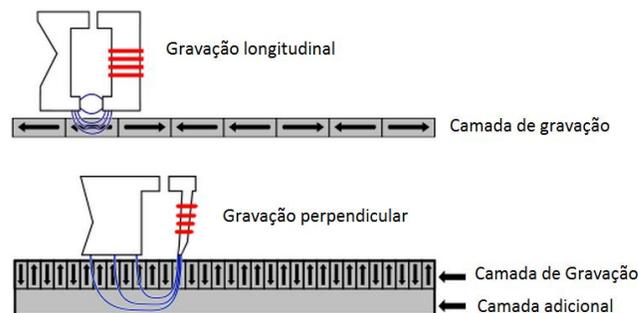


Figura 4. Formas de gravação no disco magnético

Para um armazenamento confiável, a material de gravação precisa resistir ao efeito de auto-desmagnetização, que ocorre quando os domínios magnéticos se repelem uns aos outros. Esse efeito limita o tamanho mínimo da região de magnetização e, por consequência, a quantidade máxima de bits por área. Além disto, variações na temperatura também podem levar a desmagnetização do material. A seção 4 realiza uma discussão sobre a atual densidade de bits dos discos magnéticos e os seus limites.

A figura 5 apresenta uma possível forma de codificação de bits. O tamanho da região é conhecido a priori, de forma que quando a cabeça se aproxima da borda será analisado se a orientação do campo magnético é a mesma ou se é invertida. Neste exemplo, uma reversão representa um bit 1, enquanto a não reversão um bit zero.

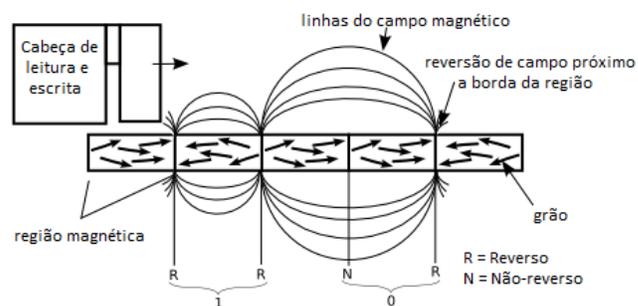


Figura 5. Processo de identificação dos bits armazenados

3. Interfaces

Com o vertiginoso aumento da capacidade de armazenamento dos discos magnéticos a interface de transferência de dados com o sistema de computação precisa evoluir constantemente para permitir que os dados, cada vez em maior densidade, possam fluir entre o disco e as hierarquias superiores de memória. A seguir são descritas as principais interfaces desenvolvidas para os discos rígidos.

3.1 ATA

A Advanced Technology Attachment (ATA) é um padrão para interligar dispositivos de armazenamento, como discos rígidos e drives de CD-ROMs, que opera como um barramento compartilhado entre os dispositivos. A evolução do padrão ao longo do tempo criou diversas variações de interfaces, que são listadas a seguir.

3.1.1 (E)IDE - (Extended) Integrated Drive Electronics.

Embora o padrão tenha definido a designação ATA desde o início, o mercado inicial divulgou a tecnologia como IDE (e sucessora E-IDE). Embora estas designações fossem meramente comerciais e estes termos aparecem muitas vezes ao mesmo tempo. O termo Integrated Drive Electronics (IDE) refere-se não somente à definição do conector e interface, mas também ao fato do controlador estar integrado à unidade de disco rígido, eximindo a placa-mãe dessa função.

3.1.2 ATAPI - Advanced Technology Attachment Packet Interface.

A interface foi projetada inicialmente apenas para conectar discos rígidos; com o advento de outros tipos de dispositivos de armazenamento como drives de CD-ROM, fitas magnéticas, e disquetes de grande capacidade, como o ZIP drive, foi necessária à introdução de extensões ao padrão inicial. É neste contexto que surge o ATAPI (Advanced Technology Attachment Packet Interface).

3.1.3 UDMA - Ultra DMA.

O modo de transferência também sofreu evoluções que se refletiram na tecnologia de interface dos discos. O modo PIO (Programmed I/O) é uma forma de acesso a disco efetuada com ajuda da CPU. O padrão ATA estabelece cinco modos que diferem nas taxas em que os dados são transferidos. O modo DMA (Direct Memory Access) transfere dados entre disco e memória sem interferência da CPU, que fica liberada para realizar outras tarefas. O padrão original previa um único modo DMA: DMA 0, mais tarde foram incorporados os modos DMA 1 e DMA 2, os quais diferem apenas quanto às taxas máximas de transferência de dados. Muitos modos DMA foram definidos para a interface IDE/ATA. Tais modos foram agrupados em duas categorias. O primeiro conjunto de modos é denominado DMA single word enquanto que o segundo é denominado DMA multiword. Claramente, os discos rígidos estavam ficando cada vez mais rápidos e, obviamente, a taxa máxima do modo DMA multi word (16.7MB/s) tornou-se insuficiente para tais discos. Os engenheiros perceberam que a tarefa de elevar a taxa máxima de transferência não era fácil; aumentar a velocidade da interface poderia provocar problemas relacionados à interferência de transmissão. Então, ao invés de tornar a interface mais rápida, uma abordagem diferente foi desenvolvida. O resultado foi a criação de um novo tipo de modo de transferência DMA, chamada de modo Ultra DMA.

Com a introdução do Serial ATA em 2003, o padrão ATA original foi retroativamente renomeado para Parallel ATA (ATA Paralelo, ou PATA).

A tabela 2 resume a evolução do padrão ATA, descrevendo as suas principais funcionalidades.

Tabela 2. Evolução dos discos ATA.

Padrão	Outros nomes	Funcionalidades
ATA-1	ATA, IDE	PIO modo 0: 3.3 MBps PIO modo 1: 5.2 MBps PIO modo 2: 8.3 MBps Single-word DMA modo 0: 2.1 MBps Single-word DMA modo 1: 4.2 MBps Single-word DMA modo 2: 8.3 MBps Multi-word DMA modo 0: 4.2
ATA-2	EIDE, Fast ATA, Fast IDE, Ultra	PIO 3,4: 11.1, 16.6 Multi-word DMA 1,2: 13.3, 16,6

	ATA	
ATA-3	EIDE	S.M.A.R.T., Security
ATA-4	ATAPI-4, ATA/ATAPI-4	Ultra DMA/33: UDMA 0,1,2: 16.7, 25.0, 33.3
ATA-5	ATA/ATAPI-5	Ultra-DMA/66: UDMA 3,4: 44.4, 66.7
ATA-6	ATA/ATAPI-6	Ultra-DMA/100: UDMA 5: 100
ATA-7	ATA/ATAPI-7	Ultra-DMA/133: UDMA 6: 133

3.2 SATA

Serial ATA, SATA ou S-ATA é o sucessor da tecnologia ATA. Diferentemente dos discos rígidos ATA, que transmitem os dados em um barramento de quarenta ou oitenta fios paralelos, os discos rígidos SATA transferem os dados em série através de uma conexão ponto a ponto. Os cabos Serial ATA são formados por dois pares de fios (um par para transmissão e outro par para recepção) usando transmissão diferencial, e mais três fios para o aterramento, totalizando 7 fios. Isso permite o uso de cabos com menor espessura que facilitam a ventilação do gabinete.

As principais vantagens sobre a interface ATA são:

- Maior rapidez na transferência de dados devido a maior imunidade à ruído e cross-talk;
- Possibilidade de remover ou acrescentar dispositivos enquanto em operação (hot swapping);
- Utilização de cabos mais estreitos que permitem o resfriamento de forma mais eficiente;
- Capacidade de reconhecer os dispositivos de imediato após serem conectados;
- Ligação de dispositivos externos através do conector eSATA;
- Codificação de linha codificação conhecido como 8B/10B, também usado nas interfaces de rede Ethernet.

O padrão de interface dos controladores SATA é chamada AHCI (Advanced Host Controller Interface) que permite a utilização de recursos avançados, como o já mencionado hot plug e NCQ (Native Command Queuing)[5], que será brevemente apresentado a frente.

3.2.1 SATA 1.5 Gbps

A primeira geração Serial-ATA, também conhecida como SATA/150 ou mesmo SATA I, opera em 1,5 gigahertz. A transferência de dados é de 1,2 gigabits por segundo ou 150 megabytes por segundo, e permite o uso de cabos mais longos do que os antigos cabos IDE ou ATA/133.

Logo após a introdução de SATA/150 uma série de falhas foi observada. A primeira versão do SATA é capaz de tratar uma única transação pendente de cada vez, assim como o antecessor ATA. Já na mesma época discos com interface SCSI tratavam vários pedidos pendentes. Este recurso, chamado de NCQ, foi aprovado como um recurso opcional para a versão SATA-II, aumenta o desempenho do disco rígido quando uma série de comandos de leitura de pontos distantes um do outro é enviada pelo computador. O disco rígido analisa esses comandos e os reordena, de forma que seja possível ler mais dados do disco rígido com apenas uma rotação.

3.2.2 SATA 3.0 Gbps

Os dispositivos da primeira geração SATA é um pouco mais rápida que os dispositivos da última geração ATA, chamado de ATA-7 ou ATA/133. De fato, foi a segunda geração da tecnologia SATA (SATA-II ou SATA/300) que ganhou apelo no que diz respeito à taxa de transmissão de dados e na eficiência no atendimento às solicitações das aplicações.

Para operar em 3 Gbps foi acrescentada uma camada física (PHY) mais robusta, permitindo a duplicação da taxa de transferência para 300 MB/s.

A retro-compatibilidade com os controladores SATA-I permite que os discos SATA-II operem nos sistemas legados, o que facilitou a adoção e o barateamento devido ao alto volume de produção.

No roadmap da tecnologia SATA está prevista a padronização do SATA-III que operará na taxa de 6.0 Gbps. Também é esperada nesta versão a funcionalidade de compartilhamento de uma única conexão SATA entre múltiplos discos.

3.3 SCSI

O SCSI, Small Computer Systems Interface, é uma interface em forma de barramento amplamente usada em estações de trabalho e servidores de médio a alto desempenho. O SCSI oferece taxas de transferência mais rápidas do que a ATA, que em função do baixo custo é a interface mais comumente usada em computadores pessoais. Em comparação com o ATA o SCSI pode dar suporte a até 16 dispositivos em um único barramento, enquanto o ATA suporta apenas 2, além de oferecer banda passantes consideravelmente maior e usar menos capacidade de CPU durante a operação.

Igualmente ao ATA, o SCSI sofreu evoluções ao longo do tempo que resultaram em novas gerações deste tipo de interface. A evolução das interfaces SCSI é apresentada na tabela 3.

Tabela 3. Evolução dos discos SCSI.

Interface	Barramento (bits)	Clock (Mhz)	Banda (MBps)
SCSI-1	8	5	5
Fast SCSI	8	10	10
Fasta-Wide SCSI	16	10	20
Ultra SCSI	8	20	20
Ultra Wide SCSI	16	20	40
Ultra2 SCSI	8	40	40
Ultra2 Wide SCSI	16	40	80
Ultra3 SCSI	16	40 DDR*	160
Ultra-320 SCSI	16	80 DDR*	320
Ultra-640 SCSI	16	160 DDR*	640

* DDR: Double Data Rate.

Pode-se dizer que SCSI e ATA não são de fato tecnologias concorrentes apesar de realizarem a mesma função. Principalmente devido à diferença de custo cada uma atendeu segmentos de mercado diferentes. Enquanto a tecnologia ATA visou o mercado de computadores pessoais a tecnologia SCSI atendeu o segmento de estações de trabalho de alto desempenho e servidores. A figura 6 apresenta uma linha do tempo descrevendo os lançamentos das gerações das tecnologias SCSI e ATA, assim como a capacidade de transferência de dados de cada uma dessas.

Além da banda os discos SCSI se diferenciam dos discos ATA em outras características como otimização na busca e escrita de dados, menor latência, maior capacidade de dispositivos por barramento, maiores buffers, entre outros. Comparativos mais detalhados entre ATA e SCSI podem ser encontrados em [6].

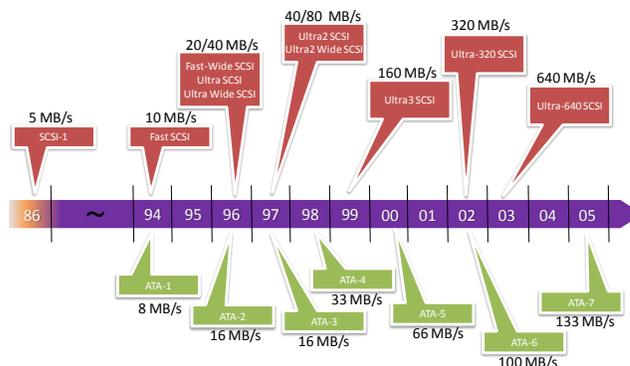


Figura 6. Comparação do SCSI e ATA ao longo do tempo.

3.4 SAS

Seguindo a linha dos discos SATA, em 2003 o padrão SCSI também ganhou uma versão serial, o SAS (Serial Attached SCSI) que é atualmente o sucessor desta tecnologia. O SAS apresenta uma melhor relação custo benefício uma vez que interfaces seriais são mais eficientes do que as interfaces paralelas.

A primeira versão opera a 3 Gbps atingindo uma capacidade real de 300 MBps. Já a segunda versão, lançada no início de 2009 opera a 6 Gbps.

Um das características mais atraentes do SAS é a capacidade de operar com discos SATA-II para aplicações de mais baixo custo. Em contrapartida, o barramento SATA-II não tem a capacidade de operar com discos SAS. Para aplicações de mais alto desempenho é possível agregar 4 interfaces SAS atingindo assim taxas de transferências de até 2.4 GBps.

Outro grande diferencial do SAS é a capacidade de expandir a conectividade para diversos discos em um mesmo barramento. Uma característica que somente surgirá no SATA em sua terceira geração. É possível somar 256 expansores cada um com capacidade de 256 discos, o que resulta em um sistema de até 65 mil dispositivos de armazenamento.

Além disto, um cabo de conexão SAS pode atingir até 10 metros, e diversos servidores podem ser ligados a um cluster SAS. Tal configuração pode substituir demandas hoje atendidas por uma SAN (Storage Area Network) utilizando a tecnologia Fibre Channel. Como será visto na seção 3.5 o Fibre Channel é uma sofisticada e custosa solução que cria uma rede de armazenamento através de uma infra-estrutura óptica ou cabeada. Grandes datacenters necessitam dessa solução, porém para aplicações de menor porte, onde alguns poucos servidores compartilham um conjunto de discos que estão a poucos metros de distância, o SAS é capaz de suprir toda a necessidade de armazenamento com igual desempenho e a um custo muito inferior. A figura 7 apresenta uma proposta da DELL [7] para cluster SAS.

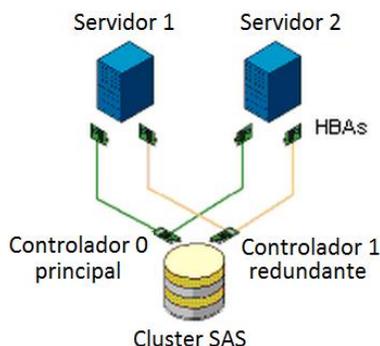


Figura 7. Solução SAS proposta pela DELL.

3.5 Fibre Channel

Fiber Channel, ou FC, é uma tecnologia de rede que opera a taxa de gigabits por segundo e é usada para a construção de redes de armazenamento (SAN). O FC foi utilizado primariamente no campo dos supercomputadores, mas atualmente é o padrão de conexão para SAN em datacenters e ambientes cooperativos. Apesar do nome, o FC pode operar tanto em meio óptico como em cobre através de pares trançados. Através da infra-estrutura óptica uma SAN FC pode ter cobertura de até 10km entre duas portas.

O protocolo FC é utilizado para o transporte de dados de forma análoga ao protocolo TCP em redes IP, e predominantemente transporta comandos SCSI ao longo da rede FC. A tabela 4 apresenta os padrões das interfaces FC desde 1997, quando a primeira versão gigabit se tornou disponível.

Tabela 4. Padrões da interface FC gigabit.

Interface	Banda (MBps)	Ano de lançamento
1FC	200	1997
2GFC	400	2001
4GFC	800	2005
8GFC	1.600	2008
10GFC	2.550	2004
16GFC	3.200	2011
20GFC	5.100	2008

Existem três formas de interconexão para as redes FC: ponto a ponto, loop arbitrated e com a utilização de um switch fabric. A seguir cada uma delas é descrita em maiores detalhes.

- Conectividade ponto a ponto.

A FC-P2P (Fibre Channel Point to Point) permite a interconexão direta de dois dispositivos. Esta é a topologia mais simples possível.

- Arbitrated loop.

A FC-AL (Fiber Channel Arbitrated loop) é projetada de forma que todos os dispositivos sejam interconectados por um anel de rede. A adição ou remoção de dispositivos causa a interrupção de todo o anel. Apenas um par de portas podem transmitir concorrentemente no anel. Essa configuração opera com portas 8GFC e inferior, todas as portas do anel devem operar na mesma velocidade. Esse solução é a forma mais barata de interconectar diversas portas FC.

- Conectividade por meio de switch.

A FC-SW (Fibre Channel Switched Fabric) é a configuração de conectividade de maior desempenho. Para tal os dispositivos são conectados em um switch fabric, similar a uma rede Ethernet gigabit. O tráfego entre duas portas não atinge a outras como no caso do FC-AL, permitindo múltiplas comunicações simultâneas, e portas com diferentes velocidades podem operar na mesma rede. Falhas em uma porta não causam a parada da rede, e novos dispositivos podem ser adicionados e removidos sem interrupção no serviço. Em contrapartida essa é a solução FC de maior custo.

4. Futuro dos discos magnéticos

A tecnologia de discos magnéticos continuará a ser dominante como memória secundária dos sistemas de computação enquanto as tecnologias concorrentes não apresentarem densidade de armazenamento e custo por gigabyte semelhantes.

Entretanto essa tecnologia apresenta pontos fracos que ameaçam a sua hegemonia, entre estes os três principais são:

- Baixa confiabilidade.
- Elevado consumo de energia.
- Alta latência na leitura e escrita de dados.

Esses fatores são decorrentes da estrutura dos discos rígidos, que foi detalhada na seção 2. As partes mecânicas são muito mais susceptíveis a defeitos, consomem consideráveis quantias de energia e limitam o tempo de operação.

Os discos rígidos apresentam como grande vantagem a facilidade de expansão da capacidade de armazenamento. Para criar um disco rígido de maior capacidade, pode-se utilizar mais pratos no mesmo drive, usar pratos de maior diâmetro, ou aumentar a densidade de gravação dos discos. Simplesmente aumentar a quantidade de pratos, de 3 para 6 discos por exemplo, resulta na duplicação da capacidade de armazenamento.

Em contrapartida, com exceção do aumento na densidade de gravação as demais alternativas resultam sempre em uma queda de desempenho, principalmente com relação ao aumento da latência na aquisição de dados. De fato, pode-se considerar que a evolução real da tecnologia se dá através do aumento do número de bits por área.

Os primeiros discos magnéticos, introduzidos pela IBM em 1956 no seu computador RAMAC, possuíam capacidade de 2 mil bits por polegada. Hoje, mais de 50 anos depois, discos com gravação longitudinal apresentam capacidade de 150 bilhões de bits por polegada quadrada, um aumento de 75 milhões de vezes na capacidade de armazenamento. Já os discos com gravação perpendicular apresentam densidade de 421 bilhões de bits por polegada quadrada em sistemas experimentais.

Especula-se que o limite máximo com a gravação perpendicular seja de 1 trilhão de bits por polegada quadrada, o que permitirá que os discos continuem a crescer a uma média de 50% ao ano pelos próximos 5 anos, considerando os discos comercialmente disponíveis que possuem capacidade em torno de 180 bilhões de bits por polegada quadrada. Dentro deste período dificilmente alguma tecnologia será uma ameaça para os discos magnéticos.

Passado esse período, novos métodos de gravação de maior densidade precisam surgir para manter essa mesma taxa de crescimento. Uma nova técnica, chamada de gravação com assistência térmica [8] promete atingir densidades na ordem de dezenas de bilhões de bits por polegada quadrada.

A principal tecnologia que ameaça os discos rígidos é a de SSD. Um SSD é um dispositivo de armazenamento que utiliza memórias de estado sólido para armazenar dados de forma persistente. O acesso é provido através das mesmas interfaces desenvolvidas para os discos rígidos, como SATA e SAS, de forma que a substituição de um disco rígido por um SSD é direta. Existem dois principais grupos de SSDs, os baseados em memória flash e os baseados em memória RAM. O segundo, apesar de ter mais rápido acesso necessita de alimentação contínua para manter a persistência dos dados.

O SSD apresenta inúmeras vantagens sobre a tecnologia de disco rígido, entre elas destacam-se: tempo de acesso 50 vezes menor, não necessidade de defragmentação de arquivos, nenhuma emissão de ruído, altíssima confiabilidade, menos suscetibilidade a fatores ambientais, menor peso e tamanho e um terço do consumo de energia.

Em contrapartida, enquanto o custo por gigabyte do SSD é de 1,2 dólares o do disco magnético é de 0,05. Ao mesmo tempo, discos de estado sólido são comercializados com capacidade inferior a 1TB enquanto discos magnéticos de 3TB já estão disponíveis no mercado. Em termos de crescimento, o SSD apresenta uma taxa de aumento de 50% a cada dois anos, enquanto os discos magnéticos apresentam 40% a cada ano.

Por causa destes fatores o SSD conseguiu conquistar o segmento dos sistemas embarcados, devida a alta confiabilidade e a baixa capacidade de armazenamento requeridas. O SSD também conquistou o segmento dos sistemas de armazenamento de mídias removíveis, substituindo disquetes e mídias ópticas. É provável que o segmento de computadores portáteis também passe a ser dominado pelo SSD, uma vez que esses computadores possuem demandas de tamanho, peso e consumo cada vez mais difíceis de serem atendidas.

Entretanto, pode-se dizer que a tecnologia de discos magnéticos continuará a ser dominante, pelo menos dentro dos próximos dez anos, nos sistemas de alta capacidade de armazenamento, como os existentes em datacenters, nos computadores de mesa e estações de trabalho.

5. Conclusões

Esse trabalho objetivou apresentar a tecnologia de discos magnéticos e sua evolução ao longo da história. Foi apresentado o princípio de funcionamento básico de um disco magnético, assim como novas técnicas que estão possibilitando a continuação do

crescimento da capacidade de armazenamento desta tecnologia. Também foi apresentado os principais tipos de discos de acordo com as diferentes interfaces existentes e de uso mais comum, tanto para aplicações em computadores pessoais quanto para as de mais alto desempenho.

Por fim, foi realizada uma comparação do disco magnético com o seu principal concorrente, o SSD. Conclui-se que esta nova tecnologia tende a ganhar espaço em segmentos de mercado onde não há grande demanda de armazenamento, porém são necessários requisitos como alta confiabilidade, baixo consumo e desempenho, em especial para o uso em dispositivos portáteis e mídias removíveis. Entretanto, o mercado de armazenamento em massa continuará a ser dominado pela tecnologia de discos magnéticos pelas próximas décadas.

6. Referências

- [1] Wikipedia – Hard Disk Drive
http://en.wikipedia.org/wiki/Hard_disk_drive#cite_note-1
- [2] Computador RAMAC
<http://www.cedmagic.com/history/ibm-305-ramac.html>
- [3] Sanderson Loubet Izidre; Tecnologia de Discos. Faculdade de Computação – Universidade Federal de Mato Grosso do Sul (UFMS)
- [4] Hennessy J. L., Patterson D. A.; Arquitetura de Computadores – Uma Abordagem Quantitativa. Tradução para o português da 3ª edição americana. Editora Campus 2003.
- [5] Wikipedia – Native Command Queuing
http://en.wikipedia.org/wiki/Native_Command_Queueing
- [6] Anderson, D.; Dykes, J.; Riedel, E. More Than an Interface – SCSI vs. ATA. In Proceedings of the 2nd Annual Conference on File and Storage Technology (FAST).
- [7] Dell SAS cluster
<http://www.delltechcenter.com/page/SAS+Cluster>
- [8] Heat-assisted Disk Drives
<http://www.zdnetasia.com/a-divide-over-the-future-of-hard-drives-39393818.htm>