

# Tudo o que você sempre quis saber sobre os HDs (e não tinha coragem de perguntar :)

por *Carlos E. Morimoto*



Sem dúvida, o disco rígido foi um dos componentes que mais evoluiu na história da computação. O primeiro disco rígido (o IBM 350) foi construído em 1956, e era formado por um conjunto de nada menos que 50 discos de 24 polegadas de diâmetro, com uma capacidade total de 4.36 MB (5 milhões de caracteres, com 7 bits cada um), algo espantoso para a época. Comparado com os discos atuais, este pioneiro custava uma verdadeira fortuna: 35 mil dólares. Porém, apesar de inicialmente, extremamente caros, os discos rígidos foram tornando-se populares nos sistemas corporativos, pois forneciam um meio rápido de armazenamento de dados.

Foto original: Matt & Kim Rudge's photos  
<http://www.flickr.com/photos/mattandkim/>

Foram produzidas cerca de 1000 unidades do IBM 350 entre 1956 e 1961, quando a produção foi descontinuada em favor de versões mais modernas. Esta foto rara, cortesia do museu digital da IBM dá uma idéia das suas dimensões:



IBM 350

Como você pode ver, o IBM 350 não era exatamente um "disco rígido" dentro da concepção que temos hoje em dia. O gabinete tinha 1.70m de altura e quase o mesmo de comprimento e pesava quase uma tonelada. Na época ele era chamado de "unidade de disco" (termo ainda usado hoje em dia por alguns) e podia ser acoplado a diversos computadores produzidos pela IBM. O termo "disco rígido" só surgiu duas décadas depois, junto com os modelos mais compactos.

De lá pra cá, tivemos uma evolução notável. Hoje em dia os HDs já ultrapassaram a marca de 1 TB, utilizam gravação perpendicular e interfaces SATA 300. São brutalmente mais rápidos que os modelos antigos e também mais baratos. Mesmo com o barateamento da memória Flash, os HDs ainda continuam imbatíveis na hora de armazenar grandes quantidades de dados.

## Como um HD funciona

Dentro do disco rígido, os dados são gravados em discos magnéticos, chamados de **platters**. O nome "disco rígido" vem justamente do fato dos discos internos serem extremamente rígidos.

Os platters são compostos de duas camadas. A primeira é chamada de **substrato**, e nada mais é do que um disco metálico, feito de ligas de alumínio. Mais recentemente, alguns fabricantes passaram a utilizar também vidro, que oferece algumas vantagens, como a maior dureza, embora também seja mais difícil de se trabalhar. Os primeiros HDs com discos de vidro foram os IBM Deskstar 75GXP, lançados em 2001.

Independentemente do material usado, o disco precisa ser completamente plano. Como os discos giram a grandes velocidades e as cabeças de leitura trabalham extremamente próximas da superfície magnética, qualquer variação seria fatal.

Para atingir a perfeição necessária, o disco é polido em uma sala limpa, até que se torne perfeitamente plano. Finalmente, vêm a parte final, que é a colocação da superfície magnética nos dois lados do disco.

Como a camada magnética tem apenas alguns microns de espessura, ela é recoberta por uma fina camada protetora, que oferece alguma proteção contra pequenos impactos. Esta camada é importante, pois apesar dos discos serem encapsulados em salas limpas, eles internamente contêm ar, com pressão ambiente.

Os discos são montados em um eixo também feito de alumínio, que deve ser sólido o suficiente para evitar qualquer vibração dos discos, mesmo a altas rotações. Este é mais um componente que passa por um processo de polimento, já que os discos devem ficar perfeitamente presos e alinhados. No caso de HDs com vários discos, eles ao separados usando espaçadores, novamente feitos de ligas de alumínio.

Finalmente, temos o motor de rotação, responsável por manter uma rotação constante. O motor é um dos maiores responsáveis pela durabilidade do disco rígido, pois uma grande parte das falhas graves provém justamente do motor.

Os HDs mais antigos utilizavam motores de 3.600 rotações por minuto, enquanto que atualmente, são utilizados motores de 5.400, 7.200 ou 10.000 RPM.

Nos HDs de notebook ainda são comuns motores de 4.200 RPM, mas os de 5.400 RPM já são maioria. Embora não seja o único, a velocidade de rotação é sem dúvidas o fator que influencia mais diretamente o desempenho.

Para ler e gravar dados no disco, são usadas **cabeças de leitura** eletromagnéticas (heads) que são presas a um **braço móvel** (arm), o que permite seu acesso a todo o disco. O braço de leitura é uma peça triangular, também feita de ligas de alumínio, para que seja ao mesmo tempo leve e resistente. O mecanismo que movimenta o braço de leitura é chamado de **actuator**.

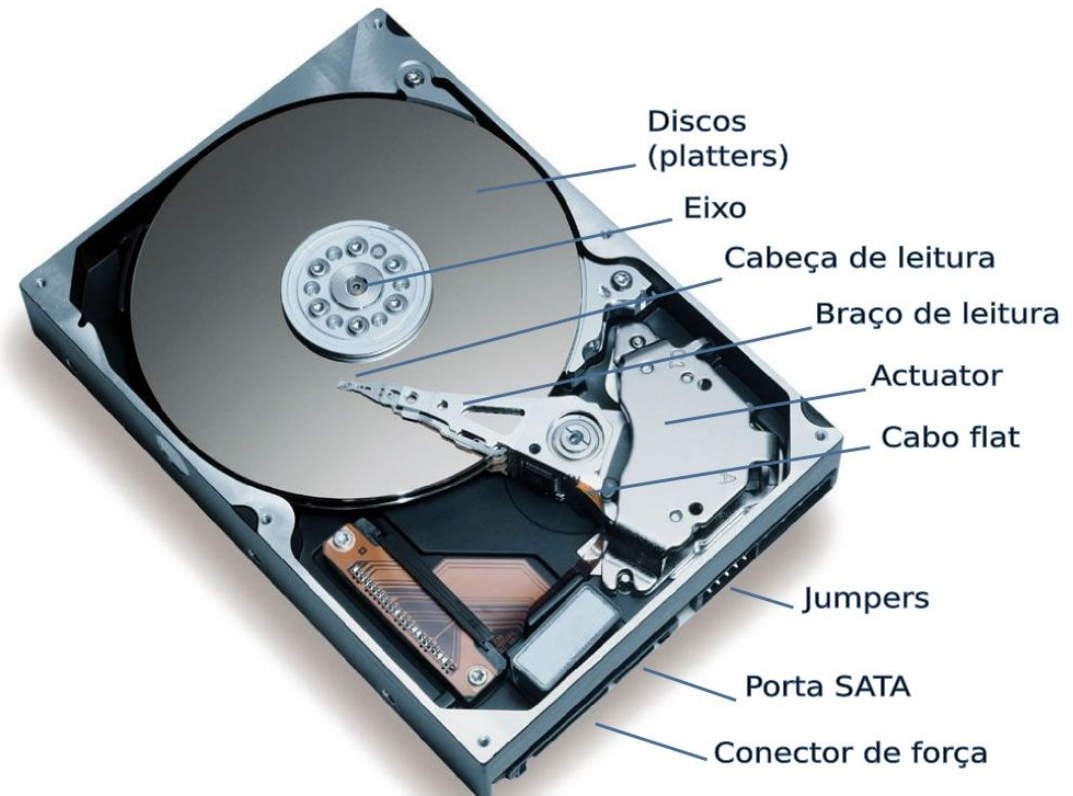
Nos primeiros discos rígidos, eram usados motores de passo para movimentar os braços e cabeças de leitura. Eles são o mesmo tipo de motor usado nos drives de disquete, onde ao receber um impulso elétrico o motor move o braço por uma curta distância, correspondente ao comprimento de uma trilha. O problema é que eles eram muito suscetíveis a problemas de desalinhamento e não permitiam densidades de gravação muito altas.

Os discos contemporâneos (qualquer coisa acima de 80 MB) utilizam um mecanismo bem mais sofisticado para esta tarefa, composto por um dispositivo que atua através de atração e repulsão eletromagnética, sistema chamado de voice coil.

Basicamente temos um eletroímã na base do braço móvel, que permite que a placa controladora o movimenta variando rapidamente a potência e a polaridade do ímã. Apesar de parecer suspeito à primeira vista, esse sistema é muito mais rápido, preciso e confiável que os motores de passo. Para você ter uma idéia, os HDs do início da década de 80, com motores de passo, utilizavam apenas 300 ou 400 trilhas por polegada, enquanto um Seagate ST3750640AS (de 750 GB) atual utiliza nada menos do que 145.000.

Para que o HD possa posicionar a cabeça de leitura sobre a área exata referente à trilha que vai ser lida, existem sinais de feedback gravados nas superfícies do disco, que orientam o posicionamento da cabeça de leitura. Eles são sinais magnéticos especiais, gravados durante a fabricação dos discos (a famosa formatação física), que são protegidos através de instruções de bloqueio incluídas no firmware do HD contra alteração posterior. Estes sinais eliminam os problemas de desalinhamento que existiam nos primeiros HDs.

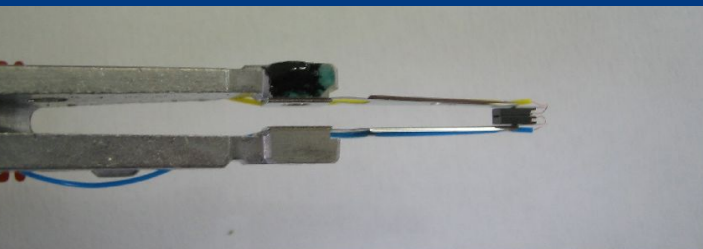
*Aqui temos um diagrama mostrando os principais componentes do HD:*



Ao ler um arquivo, a controladora posiciona a cabeça de leitura sobre a trilha onde está o primeiro setor referente a ele e espera que o disco gire até o setor correto. Este tempo inicial, necessário para iniciar a leitura é chamado de tempo de acesso e mesmo em HDs atuais de 7.200 RPM fica em torno de 12 milésimos de segundo, o que é uma eternidade em se tratando de tempo computacional. O HD é relativamente rápido ao ler setores seqüenciais, mas ao ler vários pequenos arquivos espalhados pelo HD, o desempenho pode cair assustadoramente. É por isso que existem programas desfragmentadores, que procuram reorganizar a ordem dos arquivos, de forma que eles sejam gravados em setores contínuos.

Outro dado interessante é a maneira como as cabeças de leitura lêem os dados, sem tocar na camada magnética. Se você tiver a oportunidade de ver um disco rígido aberto, verá que, com os discos parados, as cabeças de leitura são pressionadas levemente em direção ao disco, tocando-o com uma certa pressão. Aqui temos o braço de leitura de um HD, depois de removido.

*Veja que mesmo sem o disco magnético entre elas, as duas cabeças de leitura pressionam-se mutuamente:*



Apesar disso, quando os discos giram à alta rotação, forma-se uma espécie de colchão de ar, que repele a cabeça de leitura, fazendo com que ela fique sempre a alguns nanômetros de distância dos discos. É o mesmo princípio utilizado na asa de um avião; a principal diferença neste caso é que a cabeça de leitura é fixa, enquanto os discos é que se movem, mas, de qualquer forma, o efeito é o mesmo. Como veremos a seguir, os HDs não são fechados hermeticamente, muito menos a vácuo, pois é necessário ar para criar o efeito.

Esta foto mostra a cabeça de leitura "flutuando" sobre o disco em movimento. A distância é tão curta que mesmo ao vivo você tem a impressão de que a cabeça está raspando no disco, embora na realidade não esteja. Como a cabeça de leitura se movimenta rapidamente durante a operação do disco, é muito difícil tirar fotos. Para conseguir tirar esta, precisei "trapacear", desmontando o actuator e suavemente movendo a cabeça da área de descanso para o meio do disco :).



Os discos magnéticos são montados diretamente sobre o eixo do motor de rotação, sem o uso de correias ou qualquer coisa do gênero. É justamente este design simples que permite que os discos girem a uma velocidade tão grande.

Embora mais potente e muito mais durável, o motor de rotação usado nos HDs é similar aos usados nos coolers. Nos HDs antigos, eram usados motores sleeve bearing, o sistema mais simples e menos durável, que foi usado nos HDs de 3600 RPM. Em seguida, foram adotados motores ball-bearing, onde são usados rolamentos para aumentar a precisão e a durabilidade. Nos HDs modernos, é utilizado o sistema fluid-dynamic bearing, onde os rolamentos são substituídos por um fluido especial, que elimina o atrito, reduzindo o ruído e o nível de vibração.

*Aqui temos o mesmo HD da foto anterior completamente desmontado, mostrando o interior do motor de rotação:*



Assim como a maioria dos modelos de baixa capacidade, este HD utiliza um único disco, mas a maioria dos modelos utiliza dois, três ou quatro, que são montados usando espaçadores. O HD possui duas cabeças de leitura para cada disco (uma para cada face), de forma que um HD com 4 discos utilizaria 8 cabeças de leitura, presas ao mesmo braço móvel.

Embora usar mais discos permita construir HDs de maior capacidade, não é comum que os fabricantes utilizem mais de 4, pois a partir daí torna-se muito difícil (e caro) produzir componentes com a precisão necessária para manter todos os discos alinhados. Antigamente, era comum que HDs de alta capacidade (e alto custo :), sobretudo os destinados a servidores, possuíssem 6, ou até mesmo 12 discos, mas eles saíram de moda a partir da década de 90, devido à baixa demanda. Desde então, os fabricantes padronizaram a produção em torno dos HDs com até 4 discos e quem precisa de mais capacidade compra vários e monta um sistema RAID.

Naturalmente, qualquer HD aberto fora de uma sala limpa acaba sendo impregnado por partículas de poeira e por isso condenado a começar a apresentar bad-blocks e outros defeitos depois de alguns minutos de operação.

Todo HD é montado e selado num ambiente livre de partículas, as famosas salas limpas. Apesar disso, eles não são hermeticamente fechados.

Em qualquer HD, você encontra um pequeno orifício para entrada de ar (geralmente escondido embaixo da placa lógica), que permite que pequenas quantidades de ar entrem e saiam, mantendo a pressão interna do HD sempre igual à do meio ambiente. Este orifício é sempre protegido por um filtro, que impede a entrada de partículas de poeira.

Devido a isso, a pressão do ar tem uma certa influência sobre a operação do HD. Os HDs são normalmente projetados para funcionar a altitudes de até 3.000 metros acima do nível do mar. Em altitudes muito elevadas, a pressão do ar é menor, comprometendo a criação do colchão de ar. Para casos extremos, existem HDs pressurizados, que podem trabalhar a qualquer altitude.

Internamente, o HD possui um segundo filtro, que continuamente filtra o ar movimentado pelos discos. Ele tem a função de capturar as partículas que se desprendam dos componentes internos durante o uso, devido a desgaste ou choques diversos.

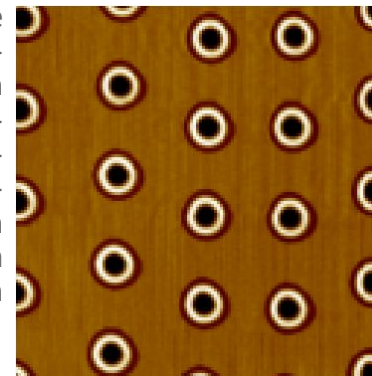
*Aqui temos uma foto de um, preso num dos cantos da parte interna do HD:*



Enquanto o HD está desligado, as cabeças de leitura ficam numa posição de descanso. Elas só saem dessa posição quando os discos já estão girando à velocidade máxima. Para prevenir acidentes, as cabeças de leitura voltam à posição de descanso sempre que não estão sendo lidos dados, apesar dos discos continuarem girando.

É justamente por isso que às vezes, ao sofrer um pico de tensão, ou o micro ser desligado enquanto o HD é acessado, surgem setores defeituosos. Ao ser cortada a energia, os discos param de girar e é desfeito o colchão de ar, fazendo com que as cabeças de leitura possam vir a tocar os discos magnéticos.

Para diminuir a ocorrência deste tipo de acidente, nos HDs modernos é utilizado um sistema que recolhe as cabeças de leitura automaticamente para a área de descanso quando a energia é cortada (tecnologia chamada de auto-parking). A área de descanso é também chamada de "landing zone" e engloba algumas das trilhas mais centrais do disco, uma área especialmente preparada para receber o impacto do "pouso" das cabeças de leitura. Uma das tecnologias mais populares é a LZT (Laser Zone Texture), desenvolvida pela IBM, onde um laser é usado para produzir pequenas cavidades ao longo da zona de pouso, que reduzem o atrito com a cabeça de leitura.



Outra técnica consiste em usar "rampas" feitas de material plástico, posicionadas na área externa dos discos, que suspendem as cabeças de leitura, evitando que elas toquem os discos mesmo quando eles param de girar. Esta tecnologia foi inicialmente usada em HDs de notebook, mas recentemente passou a ser usada também nos de 3.5" para desktops. Ela pode parecer simples, mas na verdade exige bastante tecnologia, devido à precisão necessária.

*Aqui as rampas de material plástico que suspendem as cabeças de leitura:*



Apesar de evitar danos físicos, o auto-parking nada pode fazer para evitar perda de dados ao desligar o micro incorretamente. Mesmo que todos os arquivos estejam salvos, ainda existem dados no cache de disco (criado pelo sistema operacional usando parte de memória RAM) e também no cache do HD, que utiliza memória SDRAM (também volátil). Para acelerar as operações de gravação, todos os arquivos (sobretudo os pequenos) são salvos inicialmente nos caches e depois transferidos para os discos magnéticos em momentos de ociosidade. Quando o micro é desligado abruptamente, os dados em ambos os caches

são perdidos, fazendo com que você sempre perca as últimas alterações, muitas vezes em arquivos que acreditava estarem salvos.

Por causa de tudo isso, é sempre importante usar um nobreak em micros de trabalho. A longo prazo, os dados perdidos e possíveis danos ao equipamento por causa de quedas de energia acabam custando muito mais do que um nobreak popular.

Mesmo assim, por melhores que sejam as condições de trabalho, o HD continua sendo um dispositivo baseado em componentes mecânicos, que tem uma vida útil muito mais curta que a de outros componentes do micro. De uma forma geral, os HDs para desktop funcionam de forma confiável por de dois a três anos (num PC usado continuamente). Depois disso, é melhor substituir o HD por um novo e mover o antigo para outro micro que não armazena informações importantes, pois a possibilidade de defeitos começa a crescer exponencialmente.

Fala-se muito sobre a vulnerabilidade dos HDs com relação a ímãs. Como os HDs armazenam os dados em discos magnéticos, colocar um ímã suficiente forte próximo a ele pode apagar rapidamente todos os dados. Existem inclusive "desmagnetizadores", que são eletroímãs ligados na tomada, que você passa sobre os HDs e outros discos magnéticos, justamente com a intenção de apagar os dados rapidamente.

Entretanto, se você abrir um HD condenado, vai encontrar dois ímãs surpreendentemente fortes instalados dentro do mecanismo que move a cabeça de leitura.

Naturalmente, estes ímãs não danificam os dados armazenados (senão não estariam ali ;). O principal motivo disto é que eles estão instalados numa posição perpendicular aos discos magnéticos. Se você remover os ímãs e colocá-los sobre outro HD, verá que no outro dia uma boa parte dos dados terão sido perdidos.

Se você (como todos nós) é do tipo que não consegue desmontar um micro sem deixar cair parafusos nos locais mais inacessíveis do gabinete, tem dificuldades em colocar os parafusos dos dois lados ao instalar o HD e ainda por cima nunca acha uma chave de fenda magnética para comprar, pode usar estes magnetos "roubados" do HD para transformar qualquer chave de fenda em uma chave magnética. Basta "encaixar" os ímãs nela quando quiser o efeito. Esses magnetos são feitos de uma liga contendo neodímio e, além de parafusos, permitem levantar objetos um pouco mais pesados, como martelos, por exemplo... ;)



Naturalmente, você deve tomar cuidado de não passá-los sobre discos magnéticos, a menos que queira intencionalmente apagá-los. Se você deixar a chave em contato com os ímãs por um longo período, ela continuará magnetizada (por algum tempo) mesmo depois de retirá-los.

Ao contrário da crença popular, chaves magnéticas não são perigosas para os HDs, pois os magnetos usados são muito fracos se comparados aos magnetos usados no mecanismo de leitura e no motor de rotação do HD. Os próprios discos magnéticos são relativamente resistentes a forças magnéticas externas, de forma que ímãs de baixa potência não oferecem grande perigo.

## A placa controladora

A placa lógica, ou placa controladora é a parte "pensante" do HD. Com exceção dela, o HD é um dispositivo relativamente simples, composto por uma série de dispositivos mecânicos. É a controladora que faz a interface com a placa mãe, controla a rotação do motor e o movimento das cabeças de leitura, de forma que elas leiam os setores corretos, faz a verificação das leituras, de forma a identificar erros e se possível corrigi-los usando os bits de ECC disponíveis em cada setor, atualizar e usar sempre que possível os dados armazenados no cache de disco (já que acessá-lo é muito mais rápido do que fazer uma leitura nas mídias magnéticas), e assim por diante.



Veja que a placa possui apenas três chips. O maior, no canto superior é um Samsung K4S641632H-UC60. Você pode notar que ele é muito semelhante a um chip de memória, e na verdade é :). Ele é um chip de memória SDRAM de 8 MB, que armazena o cache de disco. Até pouco tempo, os HD utilizavam chips de memória SRAM, mas os fabricantes passaram a utilizar cada vez mais chips de memória SDRAM convencional para reduzir o custo de produção. Na prática não muda muita coisa, pois apesar de ser mais lenta, a memória SDRAM oferece desempenho suficiente para a tarefa.

Assim como no caso dos processadores, o cache é um componente importante para o desempenho do HD. Ele armazena os dados acessados, diminuindo bastante o número de leituras. Dados armazenado no

cache podem ser transferidos quase que instantaneamente, usando toda a velocidade permitida pela interface SATA ou IDE, enquanto um acesso a dados gravados nos discos magnéticos demoraria muito mais tempo.

Continuando, temos o controlador principal, um chip Marvell 88i6525, que é quem executa todo o processamento. Este chip é na verdade um SOC (system on a chip), pois na verdade é um conjunto de vários chips menores, agrupados dentro do mesmo encapsulamento.

Por exemplo, este HD é um modelo SATA. A controladora da placa mãe se comunica com ele utilizando comandos padronizados, que são comuns a qualquer HD SATA. É por isso que você não precisa instalar um driver especial para cada modelo de HD, precisa apenas de um driver padrão, que sabe se comunicar com qualquer HD. Internamente, estes comandos SATA são processados e convertidos nos comandos que irão moder a cabeça de leitura, fazer girar os discos até o ponto correto e assim por diante. O sistema operacional não gerencia diretamente o cache de disco, quem faz isso é a própria controladora, que esforça para usá-lo da forma mais eficiente possível.

Naturalmente, tudo isso exige processamento, daí a complexidade interna do chip controlador.

Apesar de pequena, a placa controladora de um disco atual é muito mais sofisticada do que um micro antigo inteiro (um 286 por exemplo). Elas possuem mais poder de processamento e até mesmo mais memória, na forma do cache. Os HDs atuais usam de 8 a 32 MB de cache de disco, mais memória do que era usada em micros 386 e 486 e ainda por cima muito mais rápida! :)

Uma curiosidade é que muitos HDs antigos utilizavam um processador Intel 186 como controlador de discos. O 186 é, como você pode imaginar, o "elo perdido" entre o 8088 usados no PC XT e o 286. Ele é um chip que acabou não sendo usado nos micros PCs, mas fez um grande sucesso como microcontrolador para funções diversas.

Concluindo, temos um terceiro chip, escondido na parte inferior esquerda da foto. Ele é um Hitachi HA13645, um chip especializado, que controla o movimento das cabeças de leitura e também a rotação do motor. O chip principal envia comandos a ele, dizendo que quer acessar o setor X, ou que o motor deve entrar em modo de economia de energia, por exemplo, e ele os transforma nos impulsos elétricos apropriados. Estas funções mudam de um modelo de HD para o outro, por isso os fabricantes preferem usar um chip de uso geral como o Marvell 88i6525 como controlador principal, mudando apenas o controlador, que é um chip menor e mais barato.

A placa controladora é um componente "externo" do HD, que pode ser rapidamente substituído caso necessário. Grande parte (talvez até a maioria) dos casos onde o HD "queima" devido a problemas na rede elétrica, ou defeitos diversos, podem ser solucionados através da troca da placa controladora, permitindo recuperar os dados sem ter que recorrer aos caros serviços de uma empresa especializada.

O grande problema é justamente onde encontrar outra placa. Os fabricantes vendem placas avulsas em pequenas quantidades para empresas de recuperação, mas o fornecimento é muito restrito. Para técnicos autônomos e pequenas empresas, a única solução é usar placas doadas por outros HDs. Se o HD for um modelo recente, você pode simplesmente comprar outro, pegar a placa emprestada para fazer a recuperação dos dados e depois devolvê-la ao dono. Mas, no caso de HDs mais antigos, a única forma é procurar nos sites de leilão e fóruns em busca de uma placa usada. Existe um verdadeiro mercado paralelo de venda de placas avulsas, já que existem muitos casos de HDs inutilizados por problemas na mídia magnética, onde a placa ainda é utilizável.

É comum que os fabricantes utilizem a mesma placa lógica e os mesmos discos magnéticos em vários HDs da mesma família, variando apenas o número de discos usados. Assim, o modelo de 500 GB pode ter 4 discos, enquanto o modelo de 250 GB possui apenas dois, por exemplo. Nestes casos, é normal que a placa controladora de um funcione no outro.

Remover a placa é simples, basta usar uma chave torx para remover os parafusos e desencaixar a placa com cuidado. Na maioria dos HDs atuais, a placa é apenas encaixada sobre os contatos, mas em outros ela é ligada através de um cabo flat, que precisa ser desconectado com cuidado.

*Remoção de uma placa lógica*



Mais uma curiosidade é que os primeiros PCs utilizavam HDs com interfaces MFM ou RLL. Eles utilizavam controladoras externas, instaladas em um slot ISA e ligadas ao HD por dois cabos de dados. Este arranjo era muito ineficiente, pois a distância tornava a comunicação muito suscetível a interferências e corrupção de dados. Estes HDs possuíam várias peculiaridades com relação aos atuais, como a possibilidade de fazer uma "formatação física", onde as trilhas de dados eram realmente regravadas, o que permitia recuperar HDs com problemas de alinhamento.

Estes HDs jurássicos foram usados nos micros XT, 286 e sobreviveram até os primeiros micros 386, quando foram finalmente substituídos pelos HDs IDE, que por sua vez foram substituídos pelos HDs SATA que usamos atualmente, onde a controladora é parte integrante do HD.

Hoje em dia, a "formatação física" sobrevive apenas como um vício de linguagem. Muitos dizem que "fizeram uma formatação física" ao reparticionar o HD ou usar um programa que apaga os dados gravados (como o "zero-fill", ou o "dd" do Linux), embora uma coisa não tenha nada a ver com a outra.

## Os discos

A capacidade de um HD é determinada por basicamente dois fatores: a tecnologia utilizada, que determina sua densidade e o diâmetro dos discos, que determina a área útil de gravação.

A densidade de gravação dos HDs tem aumentado de forma surpreendente, com a introdução de sucessivas novas técnicas de fabricação. Para você ter uma idéia, no IBM 350 os discos eram simplesmente pintados usando uma tinta especial contendo limalha de ferro, um processo bastante primitivo.

Com o passar do tempo, passou a ser usado o processo de eletroplating, que é semelhante à eletrólise usada para banhar bijuterias à ouro.

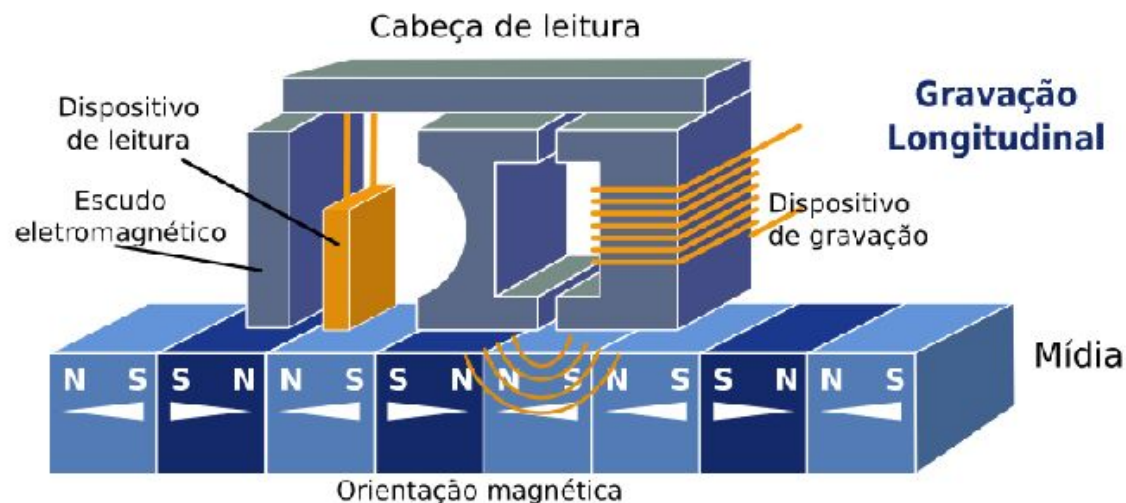
Esta técnica não permite uma superfície muito uniforme e justamente por isso só funciona em discos de baixa densidade. Ela foi usada até o final da década de 80.

A técnica usada atualmente (chamada de sputtering) é muito mais precisa. Nela a superfície magnética é construída depositando grãos microscópicos de forma incrivelmente uniforme. Quanto menores os grãos, mais fina e sensível é a superfície, permitindo densidades de gravação mais altas.

A densidade de gravação de um HD é medida em gigabits por polegada quadrada. Os HDs fabricados na segunda metade de 2006, por exemplo, utilizavam em sua maioria discos com densidade de 100 gigabits (ou 12.5 GB) por polegada quadrada. Neles, cada bit é armazenado numa área magnética com aproximadamente 200x50 nanômetros (uma área pouco maior que a de um transistor nos processadores fabricados numa técnica de 0.09 micron), e é composta por apenas algumas centenas de grãos magnéticos. Estes grãos medem apenas alguns nanômetros e são compostos por ligas de cobalto, cromo, platina, boro e outros materiais raros, muito longe da limalha de ferro utilizada pelos pioneiros.

Considerando que os discos giram a 7200 RPM e a cabeça de leitura lê os dados a mais de 50 MB/s (quando lendo setores seqüenciais), atingir densidades como as atuais é simplesmente impressionante.

*Este esquema mostra como funciona o processo de escrita e gravação em um HD:*



Como você pode ver, a cabeça é composta por dois dispositivos separados, um para gravação e outro para leitura. O dispositivo de gravação é similar a um eletroímã, onde é usada eletricidade para criar o capo magnético usado para realizar a gravação. Nos primeiros HDs, tínhamos um filamento de cobre enrolado sobre um corpo de ferro. Nos HDs atuais, os materiais usados são diferentes, mas o princípio de funcionamento continua o mesmo.

O dispositivo de leitura, por sua vez, faz o processo oposto. Quando ele passa sobre os bits gravados, capta o campo magnético emitido por eles, através de um processo de indução (no HDs antigos) ou resistência (nos atuais), resultando em uma fraca corrente, que é posteriormente amplificada.

O dispositivo de gravação é protegido por um escudo eletromagnético, que faz com que ele capte apenas o campo magnético do bit que está sendo lido, e não dos seguintes. Você pode notar que não existe isolamento entre os dispositivos de leitura e gravação. Isso acontece por que apenas um deles é usado de cada vez.

Note que esta divisão existe apenas nos HDs modernos, que utilizam cabeças de leitura/gravação MR ou GMR. Nos antigos, que ainda utilizavam cabeças de leitura de ferrite, o mesmo dispositivo fazia a leitura e a gravação.

O grande problema é que, assim como em outras áreas da informática, a tecnologia avançou até o ponto em que se começou a atingir os limites físicos da matéria. Num HD, a área referente a cada bit armazenado funciona como um minúsculo ímã, que tem sua orientação magnética alterada pela cabeça de leitura. Quando ela é orientada em um sentido temos um bit 1 e no sentido oposto temos um bit 0. A área da superfície utilizada para a gravação de cada bit chamada de "magnetic element", ou elemento magnético.

*A partir de um certo ponto, a área de gravação torna-se tão pequena que a orientação magnética dos bits pode ser alterada de forma aleatória pela própria energia térmica presente no ambiente (fenômeno de chamado de superparamagnetismo) o que faz com que a mídia deixe de ser confiável.*

A tecnologia usada nos HDs fabricados até a primeira metade de 2007 é chamada de gravação longitudinal (longitudinal recording), onde a orientação magnética dos bits é gravada na horizontal, de forma paralela à mídia.

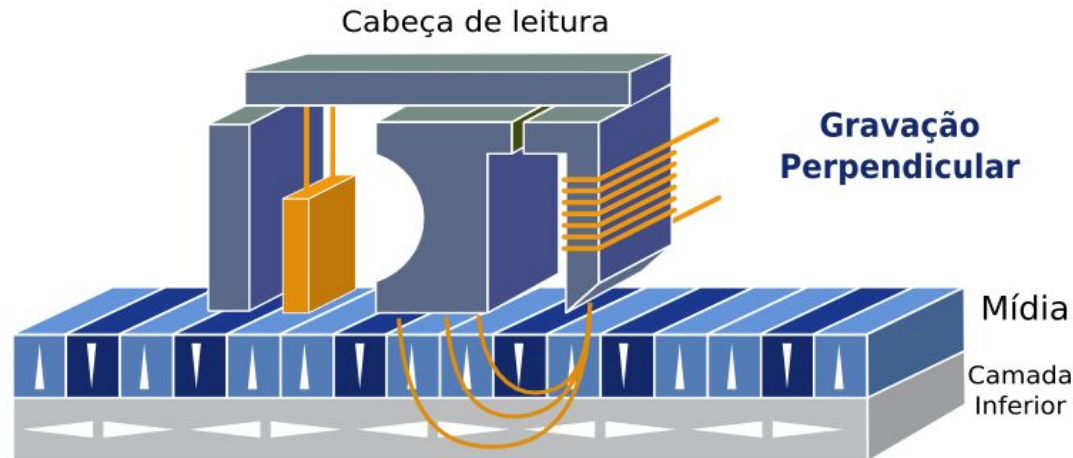
O problema é que a partir dos 100 gigabits por polegada quadrada, tornou-se muito difícil aumentar a densidade de gravação, o que acelerou a migração para o sistema de gravação perpendicular (perpendicular recording), onde a orientação magnética passa a ser feita na vertical, aumentando muito a densidade dos discos.

Estima-se que utilizando gravação longitudinal, seria possível atingir densidades de no máximo 200 gigabits por polegada, enquanto que utilizando gravação perpendicular seja possível atingir até 10 vezes mais. Isso significa que os fabricantes ainda terão margem para produzir HDs de até 10 terabytes antes de esgotar as possibilidades oferecidas pela nova tecnologia.

Na gravação perpendicular, a mídia de gravação é composta de duas camadas. Inicialmente temos uma camada de cromo, que serve como um indutor, permitindo que o sinal magnético gerado pelo dispositivo de gravação "atravesse" a superfície magnética, criando um impulso mais forte e, ao mesmo tempo, como uma espécie de isolante entre a superfície de gravação e as camadas inferiores do disco.

Ele poderia (até certo ponto) ser comparado à camada extra usada nos processadores fabricados com tecnologia SOI (silicon on insulator), onde uma camada isolante é criada entre os transistores e o wafer de silício, reduzindo a perda de elétrons e, conseqüentemente, o consumo elétrico do processador.

Sobre a camada de cromo, são depositados os grãos magnéticos. A diferença é que agora eles são depositados de forma que a orientação magnética seja vertical, e não horizontal. A cabeça de leitura e gravação também é modificada, de forma a serem capazes de lidar com a nova orientação:



Embora pareça uma modificação simples, o uso da gravação perpendicular em HDs é uma conquista técnica notável. Em termos comparativos, seria como se a NASA conseguisse enviar uma missão tripulada até Marte.

O processo de gravação perpendicular foi adotado rapidamente por todos os principais fabricantes. O primeiro foi a Fujitsu, que lançou um HD de 1.8" com gravação perpendicular ainda em 2005. O próximo foi a Seagate, que em Abril de 2006 anunciou o Barracuda 7200.10, um disco de 3.5" com 750 GB. Em Agosto de 2006 a Fujitsu anunciou um HD de 2.5" com 160 GB e em Janeiro de 2007 a Hitachi anunciou o Deskstar 7K1000, um HD de 3.5" com 1 TB que utiliza um design incomum, com 5 platters ao invés dos 4 comumente usados.

Em seguida, temos a questão do diâmetro dos discos. Como vimos, os primeiros HDs eram gigantescos, e utilizavam discos de até 24 polegadas de diâmetro. Com o passar das décadas, os discos foram encolhendo, até chegar ao que temos hoje.

Mas, como tudo na vida, existem receitas ideais para o tamanho dos discos magnéticos, de acordo com a área onde eles serão utilizados.

O problema em produzir discos muito compactos é que a superfície de gravação fica exponencialmente menor, permitindo gravar menos dados. Apesar disso, os demais componentes continuam custando quase o mesmo (ou até mais, dependendo da escala de miniaturização necessária). Isso faz com que o custo por megabyte cresça, conforme

o tamanho físico do HD diminui. Uma exemplo prático disso é a diferença no custo dos HDs de 2.5" para notebooks e os modelos de 3.5" para desktops.

A partir de um certo ponto de miniaturização, o custo por megabyte se torna mais alto que o dos cartões de memória flash e os HDs deixam de ser viáveis. O melhor exemplo é o HD de 0.85" apresentado pela Toshiba em 2005, que tinha como objetivo atender o mercado de palmtops e smartphones. Ele era tão pequeno que podia ser produzido no formato de um cartão SD e possuía um consumo elétrico baixíssimo:



O problema é que ele seria lançado em versões de apenas 2 e 4 GB, com preços a partir de US\$ 150. Com a rápida queda no custo da memória flash, logo surgiram cartões de 2 e 4 GB que custavam menos, de forma que o mini-HD acabou não encontrando seu lugar no mercado e foi descontinuado silenciosamente.

O interessante é que o oposto também é verdadeiro. HDs com discos muito grandes também acabam sendo inviáveis, pois acabam sendo bem mais lentos e mais passíveis de problemas, o que se deve a vários fatores.

O primeiro é a questão da rotação, já que discos maiores são mais pesados e demandam um maior esforço do motor de rotação, consumindo mais energia e gerando mais calor e mais barulho. Discos maiores também acabam sendo menos rígidos, o que impede que sejam girados a velocidades muito altas e torna todo o equipamento mais sensível a impactos. Dobrar o diâmetro dos discos, faz com que a rigidez seja reduzida em até 75%.

O segundo é a dificuldade de produção. Com o avanço da tecnologia, a mídia de gravação precisa ser cada vez mais fina e uniforme. Quanto maior os discos, mais difícil é recobrir toda a superfície sem que haja um grande número de pontos defeituosos.

Como se não bastasse, temos o terceiro motivo, que é o maior tempo de acesso, já que com uma superfície maior, as cabeças de leitura demoram muito mais tempo para conseguir localizar os dados (justamente devido à maior distância a ser percorrida). Se combinarmos isso com a velocidade mais baixa de rotação, acabamos tendo uma redução muito grande no desempenho.

Isso explica por que os HDs com discos de 5.25" usados nos primeiros PCs foram rapidamente substituídos pelos de 3.5". O pico evolutivo dos HDs de 5.25" foram os Quantum Bigfoot, produzidos até 1999, em capacidades de até 18 GB. Embora eles armazenassem um maior volume de dados por platter, a velocidade de rotação era bem mais baixa (apenas 3600 RPM), os tempos de acesso eram maiores e, ainda por cima, a durabilidade era menor.

Os HDs de 3.5" e de 2.5" atuais parecem ser o melhor balanço entre os dois extremos. Os HDs de 3.5" oferecem um melhor desempenho, mais capacidade de armazenamento e um custo por megabyte mais baixo (combinação ideal para um desktop), enquanto os HDs de 2.5" são mais compactos, mais silenciosos, consomem menos energia e são mais resistentes a impactos, características fundamentais no caso dos notebooks.

Temos ainda os HDs de 1.8" (mais finos e do tamanho de um cartão PCMCIA), que são usados em notebooks ultra-portáteis, além de mp3-players e alguns dispositivos de armazenamento portátil.

Uma quarta categoria são os microdrives, que utilizam discos de 1" (pequenos o suficiente para serem produzidos no formato de cartões compact flash) e podem ser utilizados em palmtops e mp3-players.

Eles foram utilizados no Palm Life Drive (4 GB) e também no iPod Nano (4 e 8 GB), mas acabaram perdendo seu espaço para os cartões de memória flash. A Hitachi chegou a anunciar o desenvolvimento de microdrives de 20 GB, utilizando tecnologia de gravação perpendicular, mas a produção em série acabou sendo cancelada, pois o preço de venda seria mais alto que o da mesma quantidade de memória flash.



Para organizar o processo de gravação e leitura dos dados, a superfície dos discos é dividida em trilhas e setores. As trilhas são círculos concêntricos, que começam no final do disco e vão se tornando menores conforme se aproximam do centro. É diferente de um CD-ROM ou DVD, onde temos uma espiral contínua.

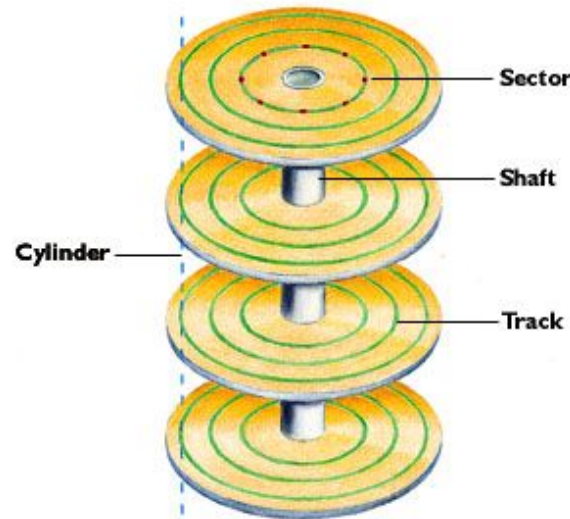
Cada trilha recebe um número de endereçamento, que permite sua localização. A trilha mais externa recebe o número 0 e as seguintes recebem os números 1, 2, 3, e assim por diante. Para facilitar ainda mais o acesso aos dados, as trilhas se dividem em setores, que são pequenos trechos de 512 cada um, onde são armazenados os dados.

Além das trilhas e setores, temos também as faces de disco. Como vimos, os HDs atuais possuem de 1 a 4 discos. Como são utilizadas ambas as faces de cada disco, temos um total de 2 a 8 faces e o mesmo número de cabeças de leitura.

Como todas as cabeças de leitura estão presas no mesmo braço móvel, elas não possuem movimentos independentes. Para acessar um dado contido na trilha 199.982 da face do disco 3, por exemplo, a controladora do disco ativa a cabeça de leitura responsável pelo disco 3 e a seguir, ordena ao braço de leitura que se dirija à trilha correspondente. Não é possível que uma cabeça de leitura esteja na trilha 199.982 ao mesmo tempo que outra esteja na trilha 555.631 de outro disco, por exemplo.

Já que todas as cabeças de leitura sempre estarão na mesma trilha de seus respectivos discos, deixamos de chamá-las de trilhas e passamos a usar o termo "cilindro". Um cilindro nada mais é do que o conjunto de trilhas com o mesmo número nos vários discos. Por exemplo, o cilindro 1 é formado pela trilha 1 de cada face de disco, o cilindro 2 é formado pela trilha 2 de cada face, e assim por diante. A próxima ilustração mostra como funciona esta divisão.

### Tracks, Cylinders, and Sectors



*Esta antiga ilustração da Quantum mostra como funciona esta divisão*

A trilha mais externa de um disco rígido possui mais que o dobro de diâmetro da trilha mais interna e, conseqüentemente, possui capacidade para armazenar muito mais dados. Porém, nos primeiros discos rígidos, assim como nos disquetes, todas as trilhas do disco, independentemente de seu diâmetro, possuíam o mesmo número de setores, fazendo com que nas trilhas mais externas, os setores ocupassem um espaço muito maior do que os setores das trilhas mais internas.

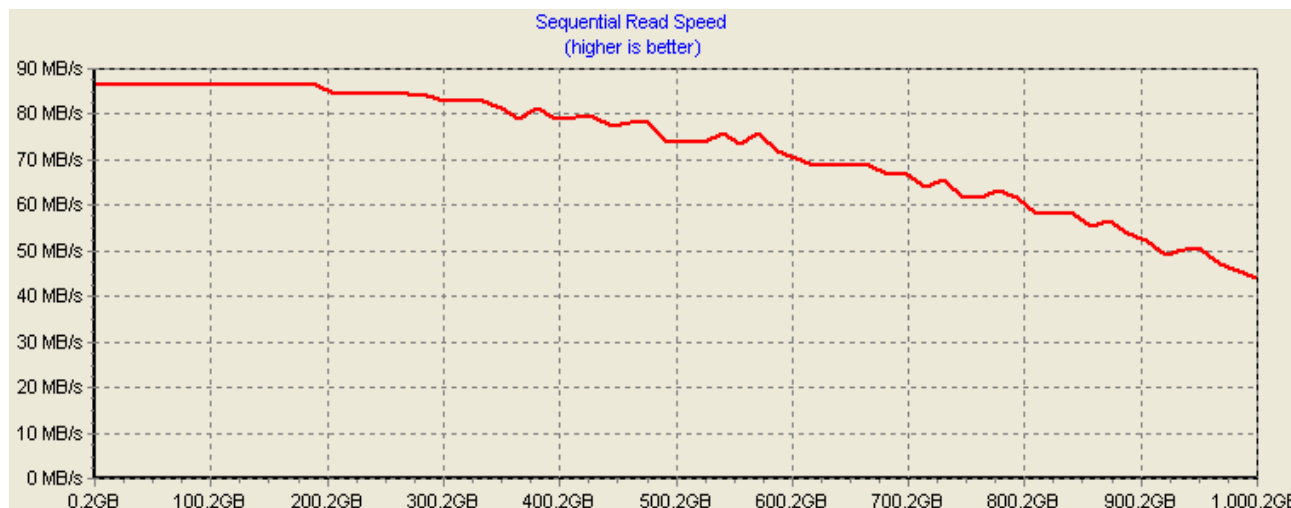
Tínhamos então um grande espaço desperdiçado, pois era preciso nivelar por baixo, fazendo com que todas as trilhas possuísem o mesmo número de setores permitido pelas trilhas mais internas, acabando por desperdiçar enormes quantidades de espaço nas primeiras trilhas do disco.

Atualmente, os HDs utilizam o Zoned bit Recording (ZBR), que permite variar a quantidade de setores por trilha, de acordo com o diâmetro da trilha a ser dividida, permitindo uma organização mais racional do espaço em disco e, conseqüentemente, uma maior densidade de gravação.

O HD pode ter então 1584 setores por trilha na área mais externa dos discos e apenas 740 na área mais interna, por exemplo. Como os discos giram sempre na mesma velocidade, isso causa um pequeno efeito colateral, que é uma considerável variação no desempenho de acordo com a área do disco que está sendo lida, proporcional ao número de setores por trilha.

Tocando em miúdos, o desempenho ao ler as trilhas mais externas acaba sendo mais que o dobro do obtido ao ler as mais internas. É por isso que em geral se recomenda colocar a partição com a instalação do sistema, ou com a partição swap no início do disco (que corresponde às trilhas mais externas) para obter o melhor desempenho.

Usando um programa de benchmark que permita realizar uma leitura seqüencial de toda a superfície do HD, como o HD Tach, você obterá sempre um gráfico similar ao da próxima página, onde a taxa de leitura começa num nível alto (trilhas externas) e vai decaindo até atingir o ponto mais baixo no final do teste (ao ler o conteúdo das trilhas mais internas).



*Queda na taxa de leitura variando de acordo com a localização da trilha (da mais externa para a mais interna)*

Um dos principais motivos do desempenho dos HDs não ter crescido na mesma proporção da capacidade ao longo das últimas décadas é que a densidade das trilhas aumentou numa escala muito maior que a dos setores dentro destas. Ou seja, as trilhas foram ficando mais "finas", mas o número de setores por trilha passou a aumentar em escala incremental. Aumentar o número de trilhas permite aumentar a área de armazenamento, mas é o número de setores por trilha, combinado com a velocidade de rotação do HD que determina a performance.

Um antigo Maxtor 7040A, de 40 MB, por exemplo, possuía uma taxa de leitura média em torno de 700 KB/s, o que permitia ler todo o conteúdo do disco em cerca de um minuto.

Um Seagate Barracuda 7200.10 atual, de 750 GB, é bem mais rápido, com uma taxa média de leitura de 64 MB/s, mas, apesar disso, como a capacidade é brutalmente maior, ler todos os dados do disco demoraria pelo menos 3:15 horas!

No futuro, esta tendência deve se manter, pois é muito mais simples para os fabricantes produzirem cabeças de leitura e sistemas de codificação capazes de lidar com trilhas mais finas, do que espremer mais dados dentro de cada trilha, já que elementos magnéticos mais curtos correspondem a um sinal magnético mais fraco e mais difícil de ser captado pela cabeça de leitura. Como um agravante, temos o problema do superparamagnetismo, que vimos a pouco.

## Correção de erros e bad blocks

Concluindo, temos a questão da detecção e correção de erros, que faz parte do processo de leitura e gravação.

Por melhor que seja sua qualidade, uma mídia magnética nunca é 100% confiável (como pode confirmar quem já teve o prazer de trabalhar com disquetes ;). Pequenas falhas na superfície da mídia podem levar a erros de leitura, sobretudo quando ela possui uma densidade de gravação de mais de 100 gigabits por polegada quadrada e gira a 7.200 RPM ou mais, como nos HDs atuais.

Isso não significa que o seu HD vá pifar amanhã, mas que são comuns erros na leitura de um setor ou outro. Obviamente, como todos os nossos dados importantes são guardados no disco rígido, a possibilidade de erros na leitura de "um setor ou outro" não seria aceitável, principalmente no caso de máquinas destinadas a operações críticas. Imagine se neste "setor ou outro" do servidor de um grande banco, estivessem gravados os dados referentes à conta bancária de um cliente importante, por exemplo.

De modo a tornar os HDs uma forma de armazenamento confiável, os fabricantes utilizam sistemas de ECC para detectar e corrigir erros de leitura eventualmente encontrados. O ECC é o mesmo sistema utilizado em pentes de memória destinados a servidores e também em CD-ROMs, onde são usados alguns bits adicionais para cada bloco de dados.

Num HD, cada setor armazena, além dos 512 bytes de dados, mais algumas dezenas de bytes contendo os códigos ECC. A criação dos bytes de ECC, assim como sua utilização posterior é feita pela placa lógica, um processo automático que é feito de forma completamente transparente ao sistema operacional.

Quando um setor é lido pela cabeça de leitura, juntamente com os dados são lidos alguns dos códigos ECC, que visam apenas verificar se os dados que estão sendo lidos são os mesmos que foram gravados, uma técnica que lembra o sistema de paridade antigamente usado na memória RAM. Caso seja verificado um erro, são usados os demais códigos para tentar corrigir o problema. Na grande maioria dos casos, esta primeira tentativa é suficiente. Estes erros transitórios, que são corrigidos com a ajuda dos códigos ECC são chamados de "soft errors" e não causam nenhum efeito colateral além de um delay de alguns milissegundos na leitura.

Caso não seja possível corrigir o erro usando o ECC, a controladora faz uma nova tentativa de leitura do setor, pois é grande a possibilidade do erro ter sido causado por alguma interferência ou instabilidade momentânea. Caso o erro persista, ela fará várias tentativas sucessivas, reduzindo a velocidade de rotação dos discos e comparando o resultado de várias leituras, de forma a tentar recuperar os dados gravados no setor. Este processo gera aquele ruído característico de HD sendo "mastigado" e quase sempre indica o aparecimento de um bad block.

Por serem defeitos físicos na mídia magnética, não existe muito o que fazer com relação a eles. O jeito é marcar os bad blocks, de forma que eles não sejam mais usados.

Os HDs atuais são capazes de marcar automaticamente os setores defeituosos. A própria controladora faz isso, independentemente do sistema operacional. Existe uma área reservada no início do disco chamada "defect map" (mapa de defeitos) com alguns milhares de setores que ficam reservados para alocação posterior. Sempre que a controladora do HD encontra um erro ao ler ou gravar num determinado setor, ela remapeia o setor defeituoso, substituindo-o pelo endereço de um setor "bom", dentro do defect map. Como a alocação é feita pela própria controladora, o HD continua parecendo intacto para o sistema operacional.

De fato, é normal que os HDs já venham de fábrica com alguns setores remapeados, causados por pequenas imperfeições na superfície da mídia. Como eles não são visíveis para o sistema operacional, nem causam problemas no uso normal, acabam passando despercebidos.

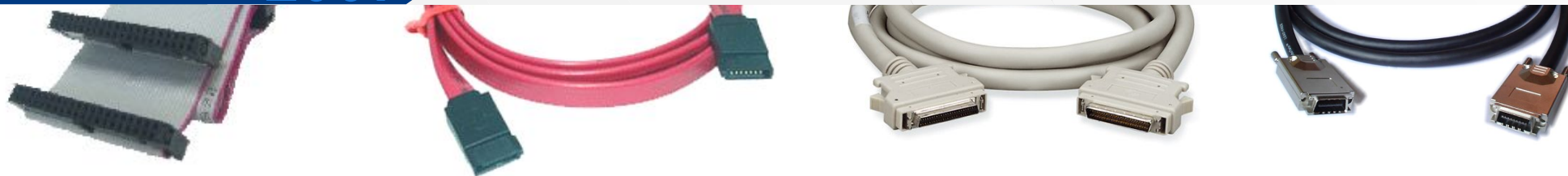
Naturalmente, o defect map é uma área limitada, que corresponde normalmente a uma única trilha. Caso o HD possua algum problema crônico, eventualmente os endereços se esgotarão e os bad blocks realmente passarão a se tornar visíveis.

Deste ponto em diante, entram em cena utilitários como o scandisk (no Windows) e o badblocks (no Linux), que permitem realizar um exame de superfície, marcando os setores defeituosos encontrados. Estes setores não são marcados no defect map, mas sim em uma área reservada da partição.

Um grande número de setores defeituosos são indício de problemas graves, como envelhecimento da mídia, defeitos no mecanismo de leitura ou mesmo contaminação do HD por partículas provenientes do ambiente. O ideal nestes casos é fazer backup de todos os dados e substituir o HD o mais rápido possível.

Entretanto, mesmo para estes HDs condenados, às vezes existe uma solução. É comum a maioria dos setores aparecerem mais ou menos agrupados, englobando uma área relativamente pequena do disco. Se houverem muitos bad clusters em áreas próximas, você pode reparticionar o disco, isolando a área com problemas.

Se, por exemplo, você percebesse que a maioria dos defeitos se encontra nos últimos 20% do disco, bastaria abrir o particionador, deletar a partição atual e criar uma nova, englobando apenas 80% do disco. Neste caso, você perderia uma boa parte da área útil, mas pelo menos teria a possibilidade de continuar usando a parte "boa" do HD (em algum micro usado para tarefas secundárias, sem dados importantes), até que ele dê seus derradeiros suspiros.



## Entendendo as interfaces: IDE, SATA, SCSI e SAS

Assim como outros componentes, as interfaces usadas como meio de conexão para os HDs passaram por um longo caminho evolutivo.

As placas-mãe usadas nos primeiros PCs sequer possuíam interfaces de disco embutidas. Naquela época, as interfaces IDE ainda não existiam, de forma que novas interfaces eram vendidas junto com os HDs e instaladas em slots ISA disponíveis. A primeira interface foi criada pela Seagate, para uso em conjunto com o ST-506, um HD de 5 MB. Em seguida foi lançado o ST-412, de 10 MB. As duas interfaces são chamadas respectivamente de MFM e RLL devido ao método de codificação usado. Além da Seagate, estes HDs e interfaces foram produzidos também por outros fabricantes, como a Quantum e a Maxtor.

Em 1985 a Quantum lançou um produto bastante peculiar, o "Plus HardCard", que era um HD RLL de 20 MB onde tanto o HD, quanto a controladora eram integrados a uma única placa ISA e o "HD" era instalado diretamente no slot, sem ocupar uma das baias do gabinete.

Como você pode imaginar, estes HDs eram um tanto quanto problemáticos, pois a placa era presa por um único parafuso, o que causava problemas de vibração excessiva e barulho. Mesmo assim, estes HDs foram relativamente populares na época:



*Quantum Plus HardCard*

O padrão seguinte foi o ESDI (Enhanced Small Device Interface), criado por um consórcio de diversos fabricantes, incluindo a Maxtor. As interfaces ESDI ainda eram instaladas em slots ISA, mas trabalhavam a uma velocidade muito maior que as MFM e RLL, oferecendo um barramento teórico de 3 MB/s. É bem pouco para os padrões atuais, mas os HDs da época trabalhavam com taxas de transferências muito mais baixas, de forma que a velocidade acabava sendo mais do que satisfatória.





## IDE

Tanto as interfaces MFM e RLL, quanto a ESDI possuem algo em comum, que é o fato da controladora fazer parte da interface, e não ao próprio HD, como temos hoje em dia. Naturalmente, integrar a interface ao HD oferece diversas vantagens, pois elimina os problemas de sincronismo causados pelo uso de cabos longos e simplifica todo o design.

Não demorou para que os fabricantes percebessem isso. Surgiu então o padrão IDE "Integrated Drive Electronics" (que indica justamente o uso da controladora integrada), desenvolvido pela Quantum e Western Digital.

Os primeiros HDs e interfaces IDE chegaram ao mercado em 1986, mas inicialmente não existia um padrão bem definido, o que fez que os primeiros anos fossem marcados por problemas de compatibilidade entre os produtos dos diferentes fabricantes.

Em 1990 o padrão foi ratificado pelo ANSI, dando origem ao padrão ATA. Como o nome "IDE" já estava mais difundido, muita gente continuou usando o termo "IDE" e outros passaram a usar "IDE/ATA" ou simplesmente "ATA", fazendo com que os dois termos acabassem virando sinônimos.

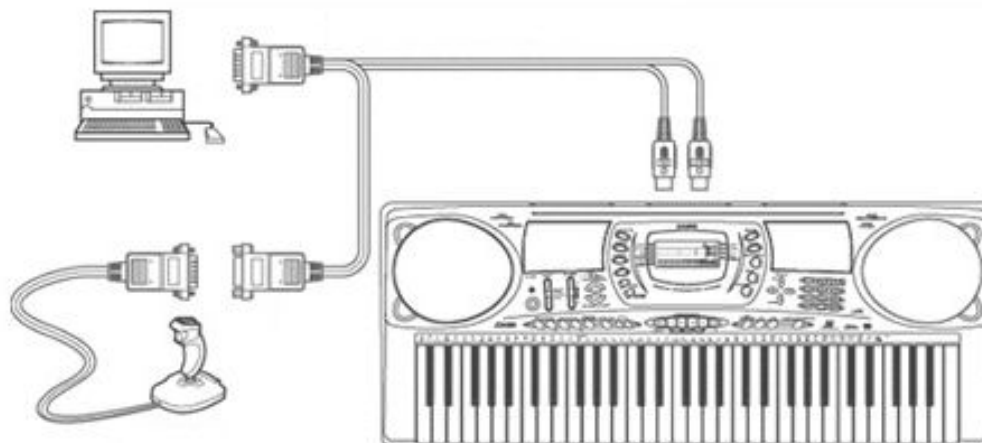
As primeiras placas IDE traziam apenas uma ou duas portas IDE e eram instaladas num slot ISA de 16 bits. Mas, logo os fabricantes passaram a integrar também outros conectores, dando origem às placas "super-ide", que eram usadas na grande maioria dos micros 386 e 486. As placas mais comuns incluíam uma porta IDE, uma porta FDD, duas portas seriais, uma paralela, além do e o conector do joystick.



*Controladora super IDE*

Como você pode ver, estas placas eram configuradas através de um conjunto de jumpers, já que na época ainda não existia plug-and-play :). Os jumpers permitiam configurar os endereços de IRQ, DMA e I/O usados, além de desativar os componentes individualmente. Se você precisasse de duas portas paralelas, por exemplo, utilizaria duas placas e configuraria uma delas para usar o IRQ 5 e endereço de I/O 378 e a outra para usar o IRQ 7 e o endereço de I/O 278.

A partir de um certo ponto, os fabricantes passaram a integrar os controladores diretamente no chipset da placa mãe, dando origem às placas com conectores integrados que conhecemos. A exceção ficou por conta do conector do joystick, que passou a ser integrado nas placas de som. Uma curiosidade é que o conector inclui também os pinos usados por dispositivos MIDI (como teclados musicais), que também são ligados no conector do joystick, através de um adaptador:

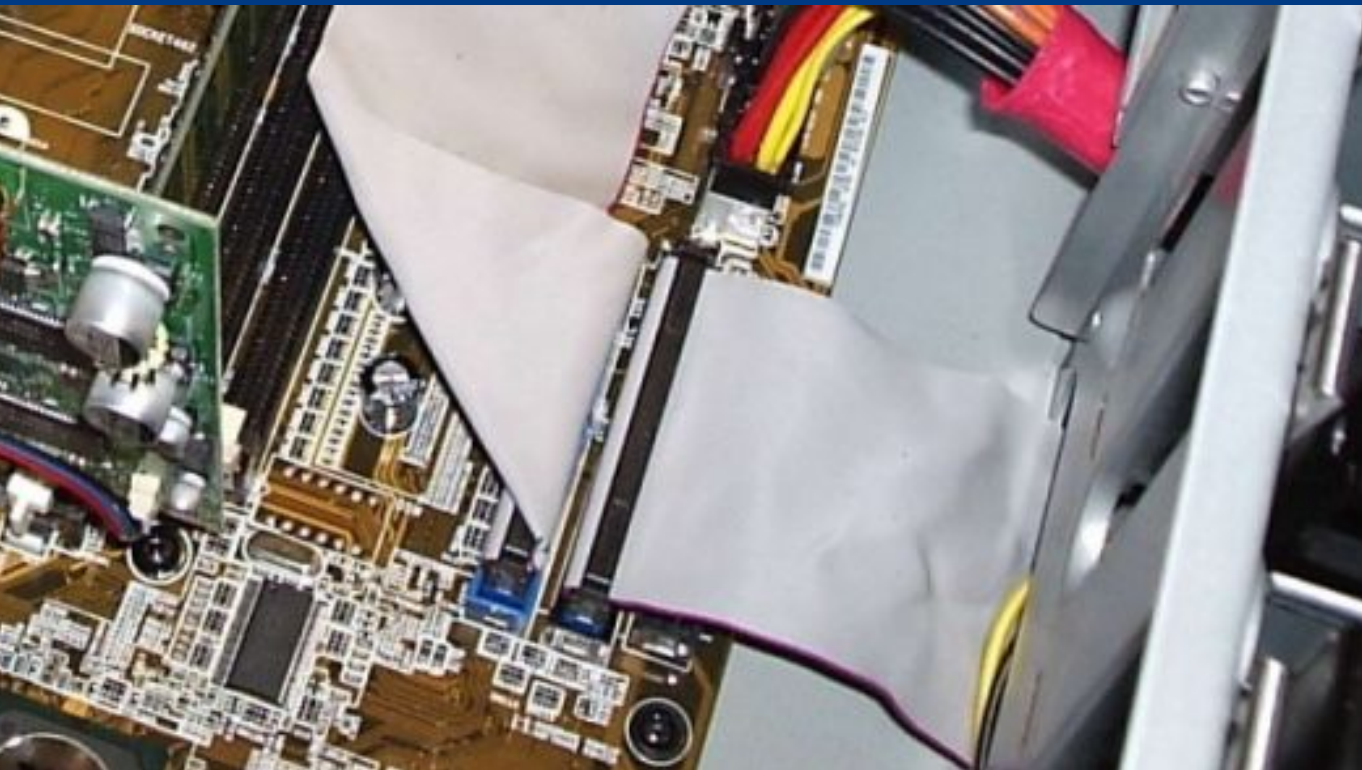


Inicialmente, as interfaces IDE suportavam apenas a conexão de HDs. Devido a isso, os primeiros drives de CD utilizavam interfaces proprietárias, incorporadas à placa de som, ou mesmo controladoras SCSI. Na época eram comuns os "kits multimídia", que incluíam o CD-ROM, placa de som, caixinhas e microfone.

Para solucionar o problema, foi desenvolvido o protocolo ATAPI (AT Attachment Packet Interface) que tornou-se rapidamente o padrão, riscando as interfaces proprietárias do mapa. É graças a ele que você pode comprar um drive de CD ou DVD e instalá-lo diretamente em uma das portas IDE, sem ter que comprar junto uma placa de som do mesmo fabricante :).

Na placa-mãe você encontra duas portas IDE (primária e secundária). Mesmo com a popularização das interfaces SATA, as portas IDE ainda continuam sendo incluídas nas placas recentes e devem demorar ainda mais alguns anos para desaparecerem completamente.

*Cada uma das portas permite instalar dois drives, de forma que podemos instalar um total de 4 HDs ou CD-ROMs na mesma placa*



Existem casos de placas mãe com 4 portas IDE (permitindo usar até 8 drives) e também controladoras IDE PCI, que incluem duas portas adicionais, que podem ser usadas em casos onde você precise usar mais do que 4 drives IDE no mesmo micro.

Para diferenciar os dois drives instalados na mesma porta, é usado um jumper, que permite configurar cada drive como master (mestre) ou slave.

Dois drives instalados na mesma porta compartilham o barramento oferecido por ela, o que acaba sempre causando uma pequena perda de desempenho. Por isso, quando são usados apenas dois drives (um HD e um CD-ROM, por exemplo), é preferível instalar cada um em uma das portas, deixando ambos jumpeados como master. Ao adicionar um terceiro, você poderia escolher entre instalar na primeira ou segunda porta IDE, mas, de qualquer forma, precisaria configurá-lo como slave, mudando a posição do jumper.

Usar cada drive em uma porta separada ajuda principalmente quando você precisa copiar grandes quantidades de dados de um HD para outro, ou gravar DVDs, já que cada drive possui seu canal exclusivo com o chipset.

No Windows, os drives são simplesmente identificados de forma seqüencial. O HD instalado como master da IDE primária apareceria no Windows Explorer como "C:" e o CD-ROM, instalado na IDE secundária como "D:", por exemplo.

Se você adicionasse um segundo HD, instalado como slave da primeira IDE, ele passaria a ser o "D:" e o CD-ROM o "E:".

*No Linux, os drives recebem endereços fixos, de acordo com a posição em que forem instalados:*

<b>IDE primária</b>	Master = /dev/hda
	Slave = /dev/hdb
<b>IDE secundária</b>	Master = /dev/hdc
	Slave = /dev/hdd

O cabo IDE possui três encaixes, um que é ligado na placa mãe e outro em cada dispositivo. Mesmo que você tenha apenas um dispositivo IDE, você deverá ligá-lo no conector da ponta, nunca no conector do meio. O motivo para isto, é que, ligando no conector do meio o cabo ficará sem terminação, fazendo com que os dados venham até o final do cabo e retornem na forma de interferência, prejudicando a transmissão.

*Como de praxe, as interfaces IDE/ATA passaram por um longo caminho evolutivo. As interfaces antigas, usadas em micros 386/486 e nos primeiros micros Pentium suportam (de acordo com seu nível de atualização), cinco modos de operação, que vão do PIO mode 0, ao PIO mode 4:*

<b>PIO mode 0</b>	3.3 MB/s
<b>PIO mode 1</b>	5.2 MB/s
<b>PIO mode 2</b>	8.3 MB/s
<b>PIO mode 3</b>	11.1 MB/s
<b>PIO mode 4</b>	16.6 MB/s

As mais recentes suportam também o Multiword DMA, que é um modo de acesso direto, onde o HD ou CD-ROM podem transferir dados diretamente para a memória, sem que o processador precise se envolver diretamente na transferência. O uso do DMA melhora bastante o desempenho e a responsividade do sistema, evitando que o micro "pare" enquanto um programa pesado está sendo carregado, ou durante a gravação de um CD, por exemplo.

Apesar disso, o Multiword DMA não chegou a ser muito usado, pois não era diretamente suportado pelo Windows 95, e os drivers desenvolvidos pelos fabricantes freqüentemente apresentavam problemas de estabilidade. Para piorar, muitos drives de CD e HDs antigos não funcionavam quando o DMA era ativado.

A solução veio com o padrão ATA-4, ratificado em 1998. Ele nada mais é do que o padrão Ultra ATA/33 (o nome mais popularmente usado) que é usado em placas para micros Pentium II e K6-2 fabricadas até 2000. Nele, a taxa de transferência máxima é de 33 MB/s e é suportado o modo UDMA 33, que permite transferências diretas para a memória também a 33 MB/s. É graças a ele que você pode assistir a filmes em alta resolução e DVDs no seu PC sem falhas

Você pode fazer uma experiência, desativando temporariamente o suporte a UDMA para o seu DVD-ROM para ver o que acontece.

No Linux, use o comando "hdparm -d0 /dev/dvd" (como root). No Windows, acesse o gerenciador de dispositivos, acesse as propriedades do drive e desmarque a opção referente ao DMA.

Tente agora assistir a um DVD. Você vai perceber que tanto o vídeo quanto o som ficam cheios de falhas, tornando a experiência bastante desagradável. Isso acontece por que, com o UDMA desativado, o processador precisa periodicamente parar o processamento do vídeo para ler mais dados no DVD. Quanto mais rápido o processador, mais curtas são as falhas, mas elas persistem mesmo num processador de 2 ou 3 GHz.

Para reverter, use o comando "hdparm -d1 /dev/dvd" ou marque novamente a opção do DMA, no caso do Windows.

Depois que o problema do DMA foi finalmente resolvido, os fabricantes se concentraram em aumentar a velocidade das portas. Surgiram então os padrões ATA-5 (Ultra ATA/66), ATA-6 (Ultra ATA/100) e ATA-7 (Ultra ATA/133), que é o usado atualmente.

Eles suportam (respectivamente), os modos UDMA 66, UDMA 100 e UDMA 133, além de manterem compatibilidade com os padrões anteriores.

No caso do HD, não existe muito com o que se preocupar, pois os fabricantes são os primeiros a adotar novos modos de operação, de forma a manter seus produtos. Se você tem em mãos um HD antigo, que só suporta UDMA 33, por exemplo, pode ter certeza de que a taxa de transferência oferecida por ele é baixa, o que torna desnecessário o uso de uma interface mais rápida em primeiro lugar.

Ao contrário dos HDs, os drivers de CD e DVD ficaram estagnados no UDMA 33, pois como eles trabalham com taxas de transferência muito mais baixas, os padrões mais rápidos também não trazem vantagens. É possível que alguns fabricantes eventualmente passem a lançar drives "ATA/133", usando a interface mais rápida como ferramenta de marketing, mas isso não faria diferença alguma no desempenho.

Como de praxe, devo insistir na idéia de que a velocidade da interface determina apenas o fluxo de dados que ela pode transportar e não a velocidade real do dispositivo ligado a ela. Um CD-ROM de 52x lerá as mídias a no máximo 7.8 MB/s, independentemente da velocidade da interface. Funciona como numa auto-estrada: se houver apenas duas pistas para um grande fluxo de carros, haverão muitos congestionamentos, que acabarão com a duplicação da pista. Porém, a mesma melhora não será sentida caso sejam construídas mais faixas.

Modo de Operação	Taxa de transferência:
<b>ATA-4 (Ultra ATA/33, UDMA 33)</b>	33 MB/s
<b>ATA-5 (Ultra ATA/66, UDMA 66)</b>	66 MB/s
<b>ATA-6 (Ultra ATA/100, UDMA 100)</b>	100 MB/s
<b>ATA-7 (Ultra ATA/133, UDMA 133)</b>	133 MB/s

As portas ATA/133 usadas nas placas atuais são uma necessidade por dois motivos. O primeiro é que os HDs atuais já superam a marca dos 70 ou 80 MB/s de taxa de transferência ao ler setores contínuos e a interface precisa ser substancialmente mais rápida que o HD, para absorver também as transferências feitas a partir do cache, que são bem mais rápidas. O segundo motivo é que só a partir das interfaces ATA/100 foi introduzido o suporte a HDs IDE com mais de 137 GB (decimais) de capacidade, como veremos em detalhes a seguir.

Para que os modos mais rápidos sejam utilizados, é necessário que exista também suporte por parte do HD e que o driver correto esteja instalado.

Continuando, junto com as interfaces Ultra ATA/66, veio a obrigatoriedade do uso de cabos IDE de 80 vias, substituindo os antigos cabos de 40 vias. Eles são fáceis de distinguir dos antigos, pois os fios usados no cabo são muito mais finos, já que agora temos o dobro deles no mesmo espaço:



A adição dos 40 fios adicionais é uma história interessante, pois eles não se destinam a transportar dados. Tanto os conectores, quanto os encaixes nos drives continuam tendo apenas 40 pinos, mantendo o mesmo formato dos cabos anteriores. Os 40 cabos adicionais são intercalados com os cabos de dados e servem como terras, reduzindo o nível de interferência entre eles. Este "upgrade" acabou sendo necessário, pois os cabos IDE de 40 vias foram introduzidos em 1986, projetados para transmitir dados a apenas 3.3 MB/s!

Os cabos de 80 vias são obrigatórios para o uso do UDMA 66 em diante. A placa mãe é capaz de identificar o uso do cabo de 80 vias graças ao pino 34, que é ligado de forma diferente.

Ao usar um cabo antigo, de 40 vias, a placa baixa a taxa de transmissão da interface, passando a utilizar o modo UDMA 33.

Veja que no caso dos CD-ROMs e DVDs, ainda é comum o uso dos cabos de 40 vias, simplesmente por que, como vimos, eles ainda utilizam o modo UDMA 33. Entretanto, se você precisar instalar um HD junto com o drive óptico, é interessante substituir o cabo por um de 80 vias, caso contrário o desempenho do HD ficará prejudicado.

Outra exigência trazida pelos novos padrões é o uso de cabos com no máximo 45 centímetros de comprimento, já que acima disso o nível de interferência e atenuação dos sinais passa a prejudicar a transmissão dos dados. O padrão ATA original (o de 1990) permitia o uso de cabos de até 90 centímetros (!) que não são mais utilizáveis hoje em dia, nem mesmo para a conexão do drive de CD/DVD.

Mais uma mudança introduzida pelos cabos de 80 vias é o uso de cores para diferenciar os três conectores do cabo. O conector azul deve ser ligado na placa mãe, o conector preto é ligado no drive configurado com master da interface, enquanto o conector do meio (cinza) é usado para a conexão do segundo drive, caso presente.

Os cabos de 80 vias também suportam o uso do sistema cable select (nos de 40 vias o suporte era opcional), onde a posição dos drives (master/slave) é determinada por em qual conector do cabo eles estão ligados, eliminando a possibilidade de conflitos, já que instalar dois drives configurados como master na mesma interface normalmente faz com que ambos deixem de ser identificados no setup.

*Para usar o cable select é preciso colocar os jumpers dos dois drives na posição "CS". Consulte o diagrama presente no topo ou na lateral do drive para ver a posição correta*



Os HDs IDE de 2.5", para notebooks utilizam um conector IDE miniaturizado, que possui 44 pinos. Os 4 pinos adicionais transportam energia elétrica, substituindo o conector da fonte usado nos HDs para desktop.

Existem ainda adaptadores que permitem instalar drives de 2.5" em desktops. Eles podem ser usados tanto em casos em que você precisa recuperar dados de um notebook com defeito, quanto quando quiser usar um HD de notebook no seu desktop para torná-lo mais silencioso.

Estes adaptadores são muito simples e baratos de fabricar, embora o preço no varejo varie muito, já que eles são um item relativamente raro:





## SATA

As interfaces IDE foram originalmente desenvolvidas para utilizar o barramento ISA, usado nos micros 286. Assim como no barramento ISA, são transmitidos 16 bits por vez e utilizados um grande número de pinos. Como é necessário manter a compatibilidade com os dispositivos antigos, não existe muita margem para mudanças dentro do padrão, de forma que, mesmo com a introdução do barramento PCI e do PCI Express, as interfaces IDE continuam funcionando fundamentalmente da mesma forma.

Mesmo quando foram introduzidas as interfaces UDMA, a única grande mudança foi a introdução dos cabos de 80 vias, desenvolvidos de forma a permitir taxas de transmissão maiores, sem contudo mudar o sistema de sinalização, nem mudar os conectores.

A partir de um certo ponto, ficou claro que o padrão IDE/ATA estava chegando a seu limite e que mudanças mais profundas só poderiam ser feitas com a introdução de um novo padrão. Surgiu então o SATA (Serial ATA).

Assim como o PCI Express, o SATA é um barramento serial, onde é transmitido um único bit por vez em cada sentido. Isso elimina os problemas de sincronização e interferência encontrados nas interfaces paralelas, permitindo que sejam usadas frequências mais altas.

Graças a isso, o cabo SATA é bastante fino, contendo apenas 7 pinos, onde 4 são usados para transmissão de dados (já que você precisa de 2 fios para fechar cada um dos dois circuitos) e 3 são terras, que ajudam a minimizar as interferências.



*Cabo e conector em um HD com interface SATA*

Os cabos SATA são bem mais práticos que os cabos IDE e não prejudicam o fluxo de ar dentro do gabinete. Os cabos podem ter até um metro de comprimento e cada porta SATA suporta um único dispositivo, ao contrário do padrão master/slave do IDE/ATA. Por causa disso, é comum que as placas mãe ofereçam 4 portas SATA (ou mais), com apenas as placas de mais baixo custo incluindo apenas duas.

No final, o ganho de desempenho permitido pela maior frequência de transmissão acaba superando a perda por transmitir um único bit por vez (ao invés de 16), fazendo com que, além de mais simples e barato, o padrão SATA seja mais rápido.

Existem três padrões de controladoras SATA, o SATA 150 (também chamado de SATA 1.5 Gbit/s ou SATA 1500), SATA 300 (SATA 3.0 Gbit/s ou SATA 3000) e também o padrão SATA 600 (ou SATA 6.0 Gbit/s), que ainda está em desenvolvimento. Como o SATA utiliza dois canais separados, um para enviar e outro para receber dados, temos 150 ou 300 MB/s em cada sentido, e não 133 MB/s compartilhados, como no caso das interfaces ATA/133.

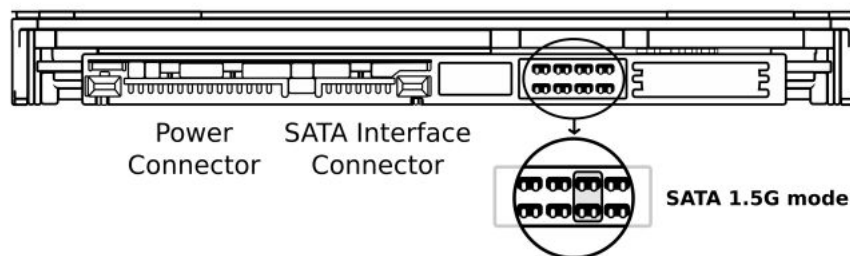
Os nomes SATA 300 e SATA 3000 indicam, respectivamente, a taxa de transferência, em MB/s e a taxa "bruta", em megabits. O SATA utiliza o sistema de codificação 8B/10B, o mesmo utilizado pelo barramento PCI Express, onde são adicionados 2 bits de sinalização para cada 8 bits de dados. Estes bits adicionais substituem os sinais de sincronismo utilizados nas interfaces IDE/ATA, simplificando bastante o design e melhorando a confiabilidade do barramento. Desta forma, a controladora transmite 3000 megabits, que, devido à codificação correspondem a apenas 300 megabytes. Ou seja, não é um arredondamento :).

As controladoras SATA 300 são popularmente chamadas de "SATA II", de forma que os dois termos acabaram virando sinônimos. Mas, originalmente, "SATA II" era o nome da associação de fabricantes que trabalhou no desenvolvimento dos padrões SATA (entre eles o SATA 300) e não o nome de um padrão específico.

Da mesma forma, o padrão de 600 MB/s chama-se SATA 600, e não "SATA III" ou "SATA IV". Mesmo os próprios fabricantes de HDs não costumam usar o termo "SATA II", já que ele é tecnicamente incorreto.

Outra curiosidade é que muitas placas mãe antigas, equipadas com controladoras SATA 150 (como as baseadas no chipset VIA VT8237 e também nas primeiras revisões dos chipsets SiS 760 e SiS 964), apresentam problemas compatibilidade com HDs SATA 300. Por causa disso, a maioria dos HDs atuais oferecem a opção de usar um "modo de compatibilidade" (ativado através de um jumper), onde o HD passa a se comportar como um dispositivo SATA 150, de forma a garantir a compatibilidade.

*Veja a s instruções impressas na etiqueta de um HD da Samsung:*



If your S-ATA Controller or motherboard cannot support S-ATA 3.5 Gbps, change jumper pin setting with reference to the above picture

Inicialmente, os HDs e placas mãe com interfaces SATA eram mais caros, devido ao tradicional problema da escala de produção. Todo novo produto é inicialmente mais caro que a geração anterior simplesmente por que a produção é menor. A partir do momento em que passa a ser produzido em quantidade, os preços caem, até o ponto em que a geração anterior é descontinuada.

A partir do momento em que os HDs SATA se popularizaram, o preço caiu em relação aos IDE. Atualmente os HDs IDE são produzidos em escala cada vez menor e por isso se tornaram mais caros e mais difíceis de encontrar do que os HDs SATA.

No caso dos micros antigos, uma opção é instalar uma controladora SATA. As mais baratas, com duas portas e em versão PCI, já custam menos de 20 dólares no exterior e tendem a cair de preço também por aqui, tornando-se um item acessível, assim como as **controladoras USB**. Note que o uso do barramento PCI limita a velocidade da controladora a 133 MB/s (um pouco menos na prática, já que o barramento PCI é compartilhado com outros dispositivos), mas isso não chega a ser um problema ao utilizar apenas um ou dois HDs.



Existem ainda conversores (chamados de bridges), que permitem ligar um HD IDE diretamente a uma porta SATA, mas eles são mais difíceis de encontrar e geralmente mais caros que uma controladora SATA PCI:



Com o lançamento do SATA, os HDs e controladoras IDE/ATA passaram a ser chamadas de "PATA", abreviação de "Parallel ATA", ressaltando a diferença.



## SCSI

As controladoras SCSI (pronuncia-se "iscâzi") são as tradicionais concorrentes das interfaces IDE. O primeiro padrão SCSI (SCSI 1) foi ratificado em 1986, na mesma época em que os primeiros HDs IDE chegaram ao mercado e consistiam em controladoras de 8 bits, que operavam a 5 MHz, oferecendo um barramento de dados de até 5 MB/s.

Em 1990, foi lançado o padrão Wide SCSI (SCSI 2). A frequência continuou a mesma, mas as controladoras passaram a utilizar um barramento de 16 bits, que dobrou a taxa de transmissão, que passou a ser de 10 MB/s.

Em seguida surgiram os padrões Fast SCSI (8 bits) e Fast Wide SCSI (16 bits), que operavam a 10 MHz e ofereciam taxas de transferência de, respectivamente 10 MB/s e 20 MB/s.

A partir daí, surgiram os padrões Ultra SCSI (8 bits, 20 MHz = 20 MB/s), Wide Ultra SCSI (16 bits, 20 MHz = 40 MB/s), Ultra2 SCSI (8 bits, 40 MHz = 40 MB/s) e Wide Ultra2 SCSI (16 bits, 40 MHz = 80 MB/s). Veja que até a evolução foi bastante previsível, com um novo padrão simplesmente dobrando a frequência e, conseqüentemente, a taxa de transferência do anterior.

A partir daí, o uso de controladoras de 8 bits foi abandonado e surgiram os padrões Ultra160 SCSI, onde a controladora operava a 40 MHz, com duas transferências por ciclo, resultando num barramento de 160 MB/s e no Ultra 320 SCSI, que mantém as duas transferências por ciclo, mas aumenta a frequência para 80 MHz, atingindo 320 MB/s.

Modelo	Controladora de 8 Bits (Narrow SCSI)	Controladora de 16 Bits (Wide SCSI)
<b>SCSI 1</b>	5 MB/s	10 MB/s
<b>Fast SCSI (SCSI-2)</b>	10 MB/s	20 MB/s
<b>Ultra SCSI (SCSI-3)</b>	20 MB/s	40 MB/s
<b>Ultra2 SCSI (SCSI-4)</b>	40 MB/s	80 MB/s
<b>Ultra160 SCSI</b>	-	160 MB/s
<b>Ultra320 SCSI</b>	-	320 MB/s

Além da diferença na velocidade, as antigas controladoras de 8 bits permitiam a conexão de apenas 7 dispositivos, enquanto as atuais, de 16 bits, permitem a conexão de até 15.

Diferentemente do que temos numa interface IDE, onde um dispositivo é jumpeado como master e outro como slave, no SCSI os dispositivos recebem números de identificação (IDs) que são números de 0 a 7 (nas controladoras de 8 bits) e de 0 a 15 nas de 16 bits. Um dos IDs disponíveis é destinado à própria controladora, deixando 7 ou 15 endereços disponíveis para os dispositivos.

O ID de cada dispositivo é configurado através de uma chave ou jumper, ou (nos mais atuais), via software. A regra básica é que dois dispositivos não podem utilizar o mesmo endereço, caso contrário você tem um conflito similar ao que acontece ao tentar instalar dois HDs jumpeados como master na mesma porta IDE :).

A maioria dos cabos SCSI possuem apenas 3 ou 4 conectores, mas existem realmente cabos com até 16 conectores, usados quando é realmente necessário instalar um grande número de dispositivos.



No barramento SCSI temos também o uso de terminadores, que efetivamente "fecham" o barramento, evitando que os sinais cheguem à ponta do cabo e retornem na forma de interferência. Na maioria dos casos o terminador é encaixado no dispositivo, mas em alguns casos basta mudar a posição de uma chave. Também existem casos de cabos que trazem um terminador pré-instalado na ponta.



*Terminador SCSI*

Note que estou usando o termo "dispositivos" e não "HDs", pois (embora raro hoje em dia) o padrão SCSI permite a conexão de diversos tipos de dispositivos, incluindo CD-ROMs, impressoras, scanners e unidades de fita.

Os gravadores de CD SCSI foram populares nos anos 90, pois o barramento SCSI oferece transferências mais estáveis que as antigas portas ATA-2 e ATA-3, usadas até então. Naquela época ainda não existia burn-free, de forma que qualquer interrupção no fluxo de dados causava a perda da mídia. Com o surgimento das interfaces IDE com suporte a UDMA, a briga se equilibrou e os gravadores de CD IDE invadiram o mercado.

As impressoras e scanners SCSI também ganharam algumas batalhas, mas acabaram perdendo a guerra para os dispositivos USB.

As unidades de fita já foram o meio mais popular para fazer backup de grandes quantidades de dados, utilizando as famosas fitas DAT. Como a fita precisa ser gravada e lida seqüencialmente, o mais comum é gerar um arquivo compactado em .tar.gz, .tar.bz2, ou mesmo em .rar, contendo todos os arquivos do backup e gravá-lo na fita, de forma seqüencial. Um arquivo muito grande pode ser dividido em vários volumes e gravado em fitas separadas. O grande problema é que é preciso ler e descompactar todo o arquivo para ter acesso aos dados.

O problema com as unidades de fita é que, embora as fitas sejam relativamente baratas, as unidades de gravação são vendidas por preços salgados. Conforme os HDs foram crescendo em capacidade e caindo em custo, eles passaram a oferecer um custo por megabyte mais baixo, fazendo com que os sistemas RAID e servidores de backup se popularizassem roubando o mercado das unidades de fita.

Um drive VXA-320 da Exabyte, por exemplo, custa US\$ 1.250 e utiliza fitas de apenas 160 GB. É comum que os fabricantes dobrem a capacidade, dizendo que as fitas armazenam "320 GB comprimidos", mas a taxa compressão varia de acordo com o tipo de dados.

A velocidade de gravação também é relativamente baixa, em torno de 12 MB/s (cerca de 43 GB reais por hora) e cada fita custa US\$ 80, o que dá um custo de US\$ 0.50 por GB. Como hoje em dia um HD de 300 GB custa (no Brasil) menos de R\$ 250, a unidade de fita simplesmente perde em todos os quesitos, incluindo confiabilidade e custo por megabyte. Ao invés de utilizar a unidade de fita, acaba sendo mais prático, rápido e barato fazer os backups usando HDs externos.



*Exabyte VXA-320*

Chegamos então na questão dos cabos. O SCSI permite tanto a conexão de dispositivos internos, quanto de dispositivos externos, com o uso de cabos e conectores diferentes para cada tipo. As controladoras de 8 bits utilizam cabos de 50 vias, enquanto as de 16 bits utilizam cabos de 68 vias. Este da foto é um **HD Ultra320 SCSI**, que utiliza o conector de 68 pinos.



As controladoras SCSI são superiores às interfaces IDE em quase todos os quesitos, mas perdem no mais importante, que é a questão do custo. Como a história da informática repetidamente nos mostra, nem sempre o padrão mais rápido ou mais avançado prevalece. Quase sempre, um padrão mais simples e barato, que consegue suprir as necessidades básicas da maior parte dos usuários, acaba prevalecendo sobre um padrão mais complexo e caro.

De uma forma geral, o padrão IDE tornou-se o padrão nos desktops e também nos servidores e estações de trabalho de baixo custo, enquanto o SCSI tornou-se o padrão dominante nos servidores e workstations de alto desempenho. Em volume de vendas, os HDs SCSI perdem para os IDE e SATA numa proporção de mais de 30 para 1, mas ainda assim eles sempre representaram uma fatia considerável do lucro líquido dos fabricantes, já que representam a linha "premium", composta pelos HDs mais caros e de mais alto desempenho.

É comum que novas tecnologias sejam inicialmente usadas em HDs SCSI sendo somente utilizadas nos discos IDE depois de tornarem-se mais baratas. Isto acontece justamente por causa do mercado de discos SCSI, que prioriza o desempenho muito mais do que o preço.

Além do custo dos HDs, existe também a questão da controladora. Algumas placas destinadas a servidores trazem controladoras SCSI integradas, mas na grande maioria dos casos é necessário comprar uma controladora separada.

As controladoras Ultra160 e Ultra320 seriam subutilizadas caso instaladas em slots PCI regulares (já que o PCI é limitado a 133 MB/s), de forma que elas tradicionalmente utilizam slots PCI-X, encontrados apenas em placas para servidores. Isto significa que mesmo que você quisesse, não poderia instalar uma controladora Ultra320 em seu desktop. Apenas mais recentemente passaram a ser fabricadas **controladoras PCI-Express**.



Como de praxe, vale lembrar que a velocidade da interface não corresponde diretamente à velocidade dos dispositivos a ela conectados. Os 320 MB/s do Ultra320 SCSI, por exemplo, são aproveitados apenas ao instalar um grande número de HDs em RAID.

Existem muitas lendas com relação ao SCSI, que fazem com que muitos desavisados comprem interfaces e HDs obsoletos, achando que estão fazendo o melhor negócio do mundo. Um HD não é mais rápido simplesmente por utilizar uma interface SCSI. É bem verdade que os HDs mais rápidos, de 15.000 RPM, são lançados apenas em versão SCSI, mas como os HDs ficam rapidamente obsoletos e tem uma vida útil limitada, faz muito mais sentido comprar um HD SATA convencional, de 7.200 ou 10.000 RPM, do que levar pra casa um HD SCSI obsoleto, com 2 ou 3 anos de uso.



## SAS

Com a introdução do Serial ATA, o barramento SCSI perdeu grande parte de seus atrativos, já que o SATA oferece uma grande parte das vantagens que antes eram atribuídas ao SCSI e, ao mesmo tempo, oferece um sistema de cabeamento mais simples.

Para preencher a lacuna, surgiu o SAS (Serial Attached SCSI), um barramento serial, muito similar ao SATA em diversos aspectos, que adiciona diversas possibilidades interessantes voltadas para uso em servidores. Ele preserva o mesmo conjunto de comandos e por isso é compatível a nível de software. Não estou falando aqui do Windows e programas como os que utilizamos em desktops, mas sim de aplicativos personalizados, complexos e caros, utilizados em grandes servidores.

Assim como o SCSI conviveu com o padrão IDE por mais de duas décadas, o SAS está destinado a concorrer com o SATA, com cada um entrincheirado em seu respectivo nicho: o SATA nos micros domésticos e servidores de baixo custo e o SAS em servidores maiores e estações de trabalho.

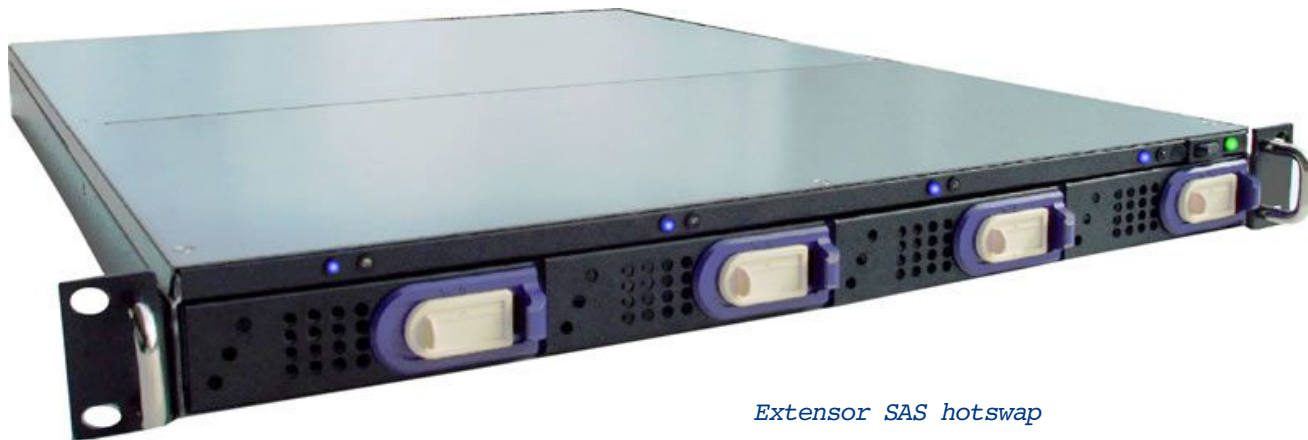
As versões iniciais do SAS suportavam taxas de transferência de 150 e 300 MB/s. Recentemente foi introduzido o padrão de 600 MB/s e passou a ser desenvolvido o padrão seguinte, de 1.2 GB/s.

A evolução é similar à do padrão SATA (note que as velocidades são as mesmas), porém o SAS tende a ficar sempre um degrau à frente.

A maior velocidade é necessária, pois o SAS permite o uso de extensores (expanders), dispositivos que permitem ligar diversos discos SAS a uma única porta. Existem dois tipos de extensores SAS, chamados de "Edge Expanders" e "Fanout Expanders". Os Edge Expanders permitem ligar até 128 discos na mesma porta, enquanto os Fanout Expanders permitem conectar até 128 Edge Expanders (cada um com seus 128 discos!), chegando a um limite teórico de até 16.384 discos por porta SAS.

Este recurso foi desenvolvido pensando sobretudo nos servidores de armazenamento. Com a popularização dos webmails e outros serviços, o armazenamento de grandes quantidades de dados tornou-se um problema. Não estamos falando aqui de alguns poucos gigabytes, mas sim de vários terabytes ou mesmo petabytes de dados. Imagine o caso do Gmail, por exemplo, onde temos vários milhões de usuários, cada um com mais de 2 GB de espaço disponível.

Os extensores SAS normalmente possuem a forma de um gabinete 1U ou 2U, destinados a serem instalados nos mesmos racks usados pelos próprios servidores. Em muitos, os discos são instalados em gavetas removíveis e podem ser trocados "a quente" (hotswap), com o servidor ligado. Isto permite substituir rapidamente HDs defeituosos, sem precisar desligar o servidor.



*Extensor SAS hotswap*

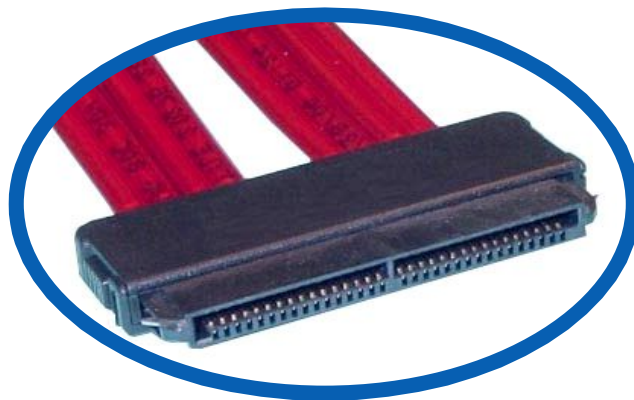
Nestes casos, seria utilizado um sistema RAID, onde parte do espaço e armazenamento é destinado a armazenar informações de redundância, que permitem restaurar o conteúdo de um HD defeituoso assim que ele é substituído, sem interrupção ou perda de dados. Ao contrário das controladoras RAID de baixo custo, encontradas nas placas mãe para desktop, que executam suas funções via software, as controladoras SAS tipicamente executam todas as funções via hardware, facilitando a configuração (já que deixa de ser necessário instalar drivers adicionais) e oferecendo um maior desempenho e flexibilidade.

Outra pequena vantagem é que o SAS permite o uso de cabos de até 6 metros, contra apenas 1 metro no SATA. A maior distância é necessária ao conectar um grande número de extensores, já que eles são grandes e os últimos tendem a ficar fisicamente afastados do servidor.

As **controladoras SAS** incluem normalmente 4 ou 8 portas e são instaladas num slot PCI-X, ou PCI Express. Nada impede também que você instale duas ou até mesmo três controladoras no mesmo servidor caso precise de mais portas. Algumas placas-mãe destinadas a servidores já estão vindo com controladoras SAS onboard, reduzindo o custo.



Assim como a maioria das controladoras SAS, esta Adaptec da foto utiliza um único conector SFF 8484, ao invés de 4 conectores separados. Ele simplifica um pouco a instalação, mas na prática não muda muita coisa, pois o conector dá origem aos 4 cabos separados da mesma forma:



Um detalhe interessante é que o padrão SAS oferece compatibilidade retroativa com os HDs SATA, permitindo que você use HDs SATA convencionais como uma forma de cortar custos, sem ter que abrir mão da possibilidade de usar os extensores.

A relação, entretanto, não é recíproca: embora o conector seja o mesmo, HDs SAS não são reconhecidos caso instalados numa porta SATA convencional, pois eles utilizam comandos específicos, que vão bem além do conjunto suportado pelas controladoras SATA.

De qualquer forma, os HDs SAS são mais caros e não oferecem vantagens em termos de desempenho, de forma

que você dificilmente iria querer utilizar um em seu desktop, de qualquer forma. O principal diferencial é que eles são certificados para operação contínua e possuem garantias maiores, geralmente de 5 anos.



A maior parte dos HDs de alto desempenho, com rotação de 15.000 RPM, que antes só existiam em versão SCSI, estão sendo lançados também em versão SAS. Nos próximos anos é de se esperar que o SAS substitua gradualmente o SCSI, assim como o SATA já substituiu o IDE quase que completamente nos micros novos.

Não existe nada de fundamentalmente diferente, que impeça que estes drives de alto desempenho sejam lançados também em versão SATA, o problema reside unicamente na questão da demanda.

Por serem caros e possuírem capacidades reduzidas (devido ao uso de discos de 2.5"), os HDs de 15.000 RPM acabam não sendo muito adequados para o público doméstico. Você dificilmente pagaria R\$ 1500 por um HD de 73 GB (como Seagate Cheetah 15K.4), por mais rápido que ele fosse, quando pode comprar um HD SATA de 300 GB por menos de R\$ 250. Esta brutal diferença de custo acaba sendo justificável apenas no mercado de servidores de alto desempenho e workstations, onde, literalmente, "tempo é dinheiro".



Carlos E. Morimoto.

É editor do site [www.guiadohardware.net](http://www.guiadohardware.net), autor de mais de 12 livros sobre Linux, Hardware e Redes, entre eles os títulos: "Redes e Servidores Linux", "Linux Entendendo o Sistema", "Linux Ferramentas Técnicas", "Entendendo e Dominando o Linux", "Kurumin, desvendando seus segredos", "Hardware, Manual Completo" e "Dicionário de termos técnicos de informática". Desde 2003 desenvolve o Kurumin Linux, uma das distribuições Linux mais usadas no país.