

Tecnologias de Disco

Rogério Alves Cardoso
RA 134078
Instituto de Computação - UNICAMP
Av. Albert Einstein, 1251
Campinas, Brasil
rogerio.cardoso@students.ic.unicamp.br

Murilo Adriano Vasconcelos
RA 134072
Instituto de Computação - UNICAMP
Av. Albert Einstein, 1251
Campinas, Brasil
muriloufg@gmail.com

RESUMO

Com o aumento da capacidade de armazenamento dos discos rígidos, a velocidade de acesso torna-se uma questão cada vez mais crítica. Interfaces de comunicação devem ser cada vez mais rápidas para acompanhar essa evolução.

Além da velocidade de acesso, outras questões como, disponibilidade, confiabilidade e consumo de energia ganharam importância, nos dias de hoje. Interfaces como ATA, SATA, SCSI, SAS e Fusion-io dominam o mercado. Baseado neste contexto, o presente trabalho tem por objetivo explicar o funcionamento, evolução e o desempenho das interfaces, levando em consideração as questões mencionadas.

Categorias e Descrição do Assunto

B.3.2 [Hardware]: Design Styles—*Mass storage*
; B.4.2 [Hardware]: InputOutput Devices—*Channels and controllers*

Termos Gerais

Teoria

Palavras Chaves

Interfaces, tecnologias de Discos, Armazenamento

1. INTRODUÇÃO

O aumento da tecnologia possibilitou a criação e armazenamento de grandes volumes de dados em meio digital, o que exige que os dispositivos de armazenamento sejam cada vez mais rápidos e eficientes. Podemos citar como exemplo os Discos Rígidos (*Hard Disk Drive - HDD*), cuja a capacidade desses dispositivos passou de 5MB, em 1956[17], para a ordem de alguns terabytes nos dias de hoje. Além da capacidade, melhorias na tecnologia, possibilitaram a diminuição do tempo de acesso ao disco e o aumento da taxa de transferência dos dados. Esse aumento também trouxe a tona, questões relacionadas com a disponibilidade e a confiabilidade desses dispositivos. O armazenamento de dados impor-

tantes em servidores tornou crítico o problema da disponibilidade e da confiabilidade em servidores de armazenamento, tornando necessárias uma infraestrutura para a replicação dos dados. Além disso, atualmente, existe uma preocupação cada vez maior em relação ao consumo de energia desses dispositivos.

Soluções para o armazenamento devem levar em conta não apenas a capacidade como, também, o desempenho das aplicações, que podem variar dependendo da interface utilizada. Além disso é necessário considerar outras questões como, orçamento, confiabilidade e a estrutura exigida. Sendo assim escolher dentre as possibilidades disponíveis nem sempre é um processo simples.

O principal objetivo desse trabalho é avaliar as principais interfaces e as tecnologias de discos existentes. O restante do trabalho é organizado como se segue: A seção 2 descreve as principais interfaces existentes (ATA, SCSI, SATA e SAS) e a evolução dessas interfaces. Na seção 3, é apresentada uma maneira de aumentar a confiabilidade e a disponibilidade dos dados; os arranjos de discos (RAID). Na seção 4, é apresentada a nova tecnologia SSD. E por fim, na seção 5, são feitas as considerações finais.

2. INTERFACES DE DISCO

Para que os dados sejam trazidos do disco até a memória principal, é necessário um padrão de comunicação eficiente e confiável. Portanto a interface tem um papel importante na escolha. Podemos então, em um nível abstrato, dividir essas interfaces em dois tipos: 1) uso pessoal (baixo desempenho, baixo custo) e 2) uso empresarial (alto desempenho, alto custo)[16].

Por causa da simplicidade e baixo custo, a interface ATA, é o protocolo de comunicação HDD mais popular em computadores pessoais, enquanto a SCSI destina-se a servidores e armazenamento em larga escala[16].

2.1 Interface ATA/IDE

O padrão foi originalmente concebido como "*PC/AT Attachment*", cuja principal função era realizar uma conexão direta com o barramento de 16 bits ISA introduzido pelo IBM PC/AT. A interface ATA (*Advanced Technology Attachment*), foi desenvolvida, em 1986, pela *Western Digital* em parceria com as empresas *Compaq Computer* e *Control Data Corporation*, inicialmente sobre o nome de IDE (*Integrated*

Drive Electronics), entretanto, devido à confusão entre os fabricantes em que alguns comercializavam a interface sobre o nome IDE e outros sobre o nome AT *Attachment*, o nome ATA/IDE foi adotado[11]. Atualmente que controla o padrão ATA é o INCITS (*InterNational Committee for Information Technology Standards*).

Após o lançamento do primeiro padrão o ATA-1, diversas modificações foram feitas no padrão, com o objetivo de melhorar o suporte da interface para dispositivos como discos óticos (CD-ROM e DVD-ROM), discos flexíveis e dispositivos de fitas em relação às interfaces antecessoras.

Quando a quarta geração, o ATA-4, surgiu, uma nova nomenclatura foi adotada, ATAPI (*ATI Packet Interface*) ou ATA/ATAPI. Desde então, foram surgindo diversos padrões, a última versão do ATA/ATAPI é o ATA/ATAPI-8, de 2004, que não é um padrão inteiramente ATA pois a sua documentação define especificações da interface SATA. O último padrão inteiramente ATA, é o ATA/ATAPI-6, de 2002. Historicamente, a evolução da interface, pode ser resumida da seguinte forma:

- **ATA-1** – Este padrão possui modo de endereçamento de 28 bits. Este padrão introduziu os modos de transferência PIO (*Programmed I/O*), 0, 1 e 2, que tem taxas de transferência que variam de 3,3 à 8,3 MB/s. No ATA-1, também foram definidos os modos DMA (*Direct Access Memory single word*) 0, 1 e 2 e modo DMA *multiword* 0[25][26].
- **ATA-2** - Devido a impaciência das empresas, que começaram a lançar extensões proprietárias, esse padrão também é conhecido como EIDE (*Enhanced IDE*) ou *Fast ATA-2*. Em 1996 o padrão oficial, X3.279-1996, foi publicado e introduziu modos de alta velocidade de transferência de dados, PIO 3 e 4 além de uma implementação completa do DMA, apresentando taxas de transmissão de 16,67 MB/s[25]. Esse padrão também introduziu novos comandos para permitir transferência por blocos e suporte a endereçamento por blocos lógicos (LBA). Nessa especificação também foi introduzido o suporte a gerenciamento de energia, permitindo a reduzindo rotação dos discos quando estes estão ociosos.
- **ATA-3** - A especificação ATA-3, publicada no padrão ANSI X3.298-1997, é uma revisão dos padrões ATA a ATA-2. Foram feitas melhorias na confiabilidade em relação aos padrões antecessores acrescentando uma tecnologia de autodetecção de condições adversas conhecida como: *Self-Monitoring Analysis and Reporting Technology* (SMART). Também foi introduzida uma nova funcionalidade de segurança permitindo que os periféricos fossem protegidos por senha[28]. O padrão manteve suporte aos modos de transferência (PIO 0,1,2,3 e 4 e DMA 0,1 e 2) das suas antecessoras. Houve uma enorme confusão nessa época, onde era possível encontrar descrito em algumas placas o modo PIO 5, que supostamente foi implementado no padrão ATA-3, esse modo foi sugerido por alguns fabricantes mais nunca implementado de fato.[25]
- **ATA/ATAPI-4** - Nessa geração uma nova nomenclatura foi adotada, ATAPI (*ATI Packet Interface*).

O padrão ATA/ATAPI-4 ou Ultra-ATA/33, foi definido em 1998, sob o padrão ANSI NCITS 317-1998. Além dos modos de transferência especificados pelos padrões ATA anteriores, o padrão introduziu novos modos de transferência que utilizam UltraDMA 0, 1 e 2 (UDMA)[10]. O UDMA otimiza a taxa de transmissão utilizando tanto o *clock* de subida quanto de descida, aumentando a velocidade de transferência em 100% (passando de 16,6 MB/s para 33,3 MB/s). O padrão especificava os modos UDMA 0, 1 e 2 modos, no entanto, os modos UDMA 0 e 1 nunca chegaram a ser implementados pelos fabricantes de discos. Outra melhoria introduzida foi o acréscimo de códigos CRC (*Cyclic Redundancy Check*) para detecção de erros nas transmissões[29].

- **ATA/ATAPI-5** - Publicada em 2000, no padrão NCITS 340-2000. Esta atualização trouxe novos modos de transmissão UDMA 3 e 4 (taxas máximas de 44.4MB/s e 66.7MB/s, respectivamente), o modo 4 também é comercializado com o nome de Ultra ATA/66 ou Ultra DMA/66. Porém, para conseguir maiores velocidades, opera com cabos (*Ribbon*) de 80 ou 40 fios. Mesmo duplicando fios e desenvolvendo outras técnicas para operar em maiores frequências, o efeito *crossstalk* ainda impõe grande degradação no desempenho da interface.
- **ATA/ATAPI-6** - Publicada em 2001 no padrão NCITS 347-2001. Essa versão suporta o modo UDMA 5 (com taxa máxima de 100MB/s) e introduziu uma nova funcionalidade, AAM (*Automatic Acoustic Management*). Essa funcionalidade permite que os dispositivos ajustem a velocidade de rotação para reduzir o ruído. O padrão ATA-6 possui um modo de endereçamento, LBA, de 48 bits, chamado LBA48. Graças ao LBA48 é possível endereçar até 144 petabytes[7].
- **ATA/ATAPI-7** - No padrão ATA/ATAPI-7 a transmissão é feita em 16 bits em paralelo, em modo *half-duplex*, utilizando um cabo *Ribbon*. O padrão foi lançado em 2005, NCITS 397-2005, e não constitui em um padrão inteiramente ATA, que passou a ser chamado PATA, pois possui definições do padrão SATA na sua especificação. Sua taxa de transmissão chegaram a 133MB/s. Na prática, nem sempre se consegue esses valores e muitas vezes, o valor real é muito inferior.
- **ATA/ATAPI-8** - O padrão ATA/ATAPI-8 começou a ser desenvolvido em 2006, e engloba não somente especificações do padrão SATA, como algumas especificações para discos sólidos (SSD), um rascunho do conjunto de comandos foi publicado em 2006[36]. Uma das novas funcionalidades trazidas para a especificação, é a tecnologia *Hybrid Drive*, que utiliza uma memória cache para aumentar a velocidade de acesso para arquivos críticos do SO. No entanto o padrão oficial ainda não foi liberado, e não há muitas informações disponíveis sobre esse padrão.

Para permitir taxas de transmissão mais altas é preciso aumentar a frequência de transmissão, agravando o problema do *crossstalk*. Uma solução para o problema é realizar a transmissão bit a bit de forma sequencial, aumentando o clock. Surgiram então as interfaces SATA (*Serial ATA*).

2.2 Interface SATA

A interface ATA realiza as transmissões dos dados de forma paralela. Um cabo contendo diversos fios é responsável por fazer a ponte entre a placa controladora e o disco. Teoricamente, transmissões paralelas são mais rápidas que as transmissões seriais. Porém, transmitir dados em alta frequência, em um barramento com fios muito próximos, pode causar diversas complicações, como o efeito *crossstalk*. Essas complicações levaram os projetistas a focarem seus estudos em transmissões seriais.

A interface SATA (Serial ATA), teve como objetivo serializar a interface ATA/IDE. Seu desenvolvimento foi iniciado em 2002, pelo SATA *Workgroup*, e trouxe resultados promissores com relação à taxa de transmissão. A interface SATA, pode trabalhar com altas frequências e em baixa tensão, evitando os problemas da sua sucessora, permitindo taxas de transmissões mais altas.

Hoje o mercado de computadores pessoais é dominado por essas interfaces, praticamente não se houve mais falar em interfaces PATA, nesses computadores[12]. Atualmente a entidade que controla o padrão SATA chama-se, atualmente, SATA-IO (SATA *International Organization*).

O padrão SATA-1 possuía uma velocidade de transmissão de 150MB/s e foi comercializado com o nome de SATA-150. Apesar da taxa de transmissão ser de 150MB/s, as transmissões SATA utilizam uma codificação 8/10 para efetuar as transmissões. Para cada 8 bits de dados reais, são transmitidos 10 bits, sendo assim uma taxa de 150MB/s de bits efetivo, representam uma taxa de 1,5Gb/s[16]. Além disso nessa versão do SATA foi implementado o recurso NCQ (*Native Command Queuing*), que utiliza políticas de controle de requisições, as quais exploram o processo de *seek* do disco, priorizando solicitações mais próximas, ou seja, leituras e gravações mais próximas da cabeça de leitura são executadas primeiro[30].

Houve uma grande confusão em relação ao nome do segundo padrão SATA, o SATA 3Gb/s. Devido a rápida evolução da interface e não havendo um controle sobre a nomenclatura por parte da organização SATA-IO, o padrão foi comercializado por algumas empresas com o nome de SATA-2, sendo que segundo a SATA-IO, esse era o nome do time que desenvolveu a especificação e não o nome da especificação[1]. O padrão SATA 3Gbs/s, possui como principal característica a velocidade de transmissão de dados de 300MB/s, o dobro do SATA-1, utilizando a mesma codificação 8/10B. Esse aumento na velocidade se deve principalmente ao clock dessa versão que é de 3GHz.

No ano de 2009, foi lançado do conjunto final de especificações da terceira versão da tecnologia Serial ATA, chamada de SATA-3 (SATA 6 Gb/s, SATA 3.0 ou SATA 600). Este padrão permite, teoricamente, taxas de transferências de até 600 MB por segundo. O SATA-3 também utiliza uma versão melhorada da tecnologia NCQ, possui melhor gerenciamento de energia. O padrão SATA-3 se mostra especialmente interessante para uso em unidades SSD, que por utilizarem memória do tipo Flash podem alcançar taxas de transferência mais elevadas que os discos rígidos. A especificação SATA-3 trabalha com frequência de até 6 GHz, também fazendo uso

Table 1: Comparação entre as interfaces[13]

Interface	Taxa de Transferência	Custo GB(2009)
PATA	66, 100, 133 MB/s	-
SCSI	160, 320 MB/s	-
<i>Fibre Channel</i>	1, 2, 4, (8) Gb/s	\$16.22
SAS	3, 6 Gb/s	\$0.90
SATA	1.5, 3, 6 Gb/s	\$0.19

da codificação 8B/10B.

Com o intuito de criar diferenças competitivas, empresas buscam inserir novas funcionalidades a interface. Com o surgimento e a popularização das memórias Flash e SSD, é possível encontrar a denominação mSATA (mini-SATA) ou μ SATA em algum lugar. Não se trata necessariamente de uma especificação nova do SATA, mas sim de um padrão de conexão desenvolvido especialmente para unidades SSD de pequeno porte que pode ser utilizado, por exemplo, em *ultrabooks* (notebooks com espessura pequena) ou mesmo tablets[15]. Outra denominação encontrada hoje em dia é o eSATA (*external SATA*) que é um tipo de porta que permite a conexão de dispositivos externos a uma interface SATA do computador[15]. Esta funcionalidade interessante aos usuários que desejam aproveitar a compatibilidade de HDs externos com a tecnologia SATA obtendo assim maiores taxas de transferência de dados.

A tecnologia SATA se mostrou como uma solução para esta questão do desempenho, que supera o desempenho de interfaces voltadas para uso corporativo como as interfaces SCSI. A interface SATA mostrou-se ser uma interface de alto desempenho, sem ter custos de produção maior como consequência como mostra a Tabela 1. Um dos fatores que foram determinantes para a sua ampla aceitação no mercado.

2.3 A Interface SCSI

A interface SCSI começou a ser desenvolvida em 1979, pela empresa *Shugart* que buscava uma por uma nova interface de discos que suportasse o endereçamento por blocos lógicos e transferências paralelas com alto desempenho. Essa interface foi chamada de SASI (*Shugart Associates System Interface*). Em 1981, a Shugart, em parceria com o NCR *Corporation*, apresentaram ao comitê ANSI a interface SASI, que passou então a ser chamada de SCSI[30][34].

Diversas melhorias foram adicionadas ao SCSI, a principais delas foram a habilidade de competir pelo barramento e libera-lo temporariamente durante a execução de um comando e a inserção de um adaptador *host* no barramento SCSI, responsável por intermediar as transações entre o dispositivo e a memória principal[34].

Em 1986, a documentação oficial do padrão, a X3.131-1986, foi lançada pela ANSI. Ao longo dos anos três padrões oficiais foram lançados o SCSI-1 (1986), o SCSI-2 (1990) e o SCSI-3 (1995). No entanto, ao longo dos anos, muitas modificações foram lançadas sobre diferentes nomes comerciais.

É possível agrupar a evolução dos padrões dessa interface da

seguinte forma:

- **SCSI-1** - A especificação básica, possui um conjunto de comandos de 6 a 10 bytes e um sistema de sinalização. Possui um barramento de 8 bits (sendo 1 bit de paridade), permitindo a conexão simultânea de até 7 dispositivos. Esse barramento é capaz de realizar transmissões de modo síncrono e assíncrono, a uma taxa de 3,5MB/s e 5MB/s respectivamente.
- **SCSI-2** - Antes do lançamento oficial da especificação, em 1985, a interface SCSI-2 começou a ser projetada. Esse padrão especifica que um subconjunto mínimo de comandos, o *Command Common Set* (CCS), que deve ser implementado. Em relação a sua antecessora, a especificação oferece novos comandos para dar suporte a dispositivos como CD-ROM e Scanners, além de oferecer a capacidade para realizar múltiplas requisições de E/S de forma simultânea entre os dispositivos conectados ao barramento, aumentando o desempenho. A interface muitas vezes é chamada de *Fast-SCSI*.
- **SCSI-3** - O SCSI-3 levou, aproximadamente, 8 anos para ser concluído, de 1992 à 2000, devido as dificuldades em formalizar um padrão abrangente consequência das diversas extensões proprietárias presentes no mercado. O padrão foi então dividido em um conjunto de padrões, que poderiam ser desenvolvidos mais rapidamente. O relacionamento entre esses padrões é definido pelo SCSI-3 *Architecture Model* (SAM). Sendo que a forma mais conhecida do SCSI passou a se chamar *Parallel SCSI* (SPI). A interface geralmente é comercializada sobre o nome de Ultra-SCSI. Em sua versão mais recente, a Ultra-320 (SPI-4), permite transmissões de 320 MB/s em modo síncrono, com 16 bits por vez à 80MHz. Vários recursos foram adicionados: detecção de erros CRC, transmissão de dados em pacotes, compensação do efeito de atraso, conexão de até 16 dispositivos, além de novas técnicas para melhorar o desempenho[23].

A especificação SCSI-3 (SPI-5), também conhecido como Ultra-640 começou a ser desenvolvido em 2003 e ainda não foi concluída. Um dos grandes motivos foram os avanços tecnológicos e os benefícios obtidos nas transmissões seriais, que oferecem altas taxas de transmissões a custo baixos. A primeira evolução no sentido de serializar a interface SCSI foi a utilização de uma interface *Fibre Channel*. Um disco *Fibre Channel*, é definido como parte da especificação SCSI, onde protocolos definem uma interface entre o SCSI e o *Fibre Channel*. Além disso a composição de camadas do *Fibre Channel* permite a aplicação de vários outros protocolos além do SCSI[2].

Recursos de conexões entre os dispositivos de longa distância fornecidos pelo *Fibre Channel* permitiu a sua utilização em redes de grande compartilhamento, as *Storage Area Network* (SAN)[8]. No entanto o principal problema da tecnologia *Fibre Channel* está no modo de comunicação síncrono, o que gera perda de desempenho para manter a comunicação entre dispositivos heterogêneos onde a taxa de transferência é dada pelo dispositivo mais lento[33].

Por causa desses problemas, quando se trata de conexões entre dispositivos de longa distância, um outro padrão sugerido é o iSCSI (internet SCSI)[33], que preserva o paradigma básico do SCSI, especialmente no que diz respeito ao conjunto de comandos, que é praticamente o mesmo. O iSCSI é uma incorporação do padrão SCSI-3 sobre TCP/IP, o argumento é que as taxas de transmissão de dados via *Ethernet* crescem mais rapidamente que as taxas de transmissão via *Fibre Channel*. Além disso, as interfaces *Fibre Channel*, tem custo bastante elevado, alguns servidores optam por utilizar os chamados FATA, que são simplesmente discos SATA equipado com algum conversor externo, que muda a interface para *Fibre Channel*.

2.4 A Interface SAS

Em 2001 as empresas Compaq/HP, LSI, Logic, Maxtor e Seagate Technologies, iniciaram o desenvolvimento de uma nova interface de alto desempenho, denominada SATA (*Serial Attached SCSI*). A interface SAS é considerada a sucessora lógica das interfaces SCSI paralelas. A primeira especificação formal foi lançada em 2004[18] e já estabelecia taxas de transferências de 3GB/s. Ao contrário do padrão SCSI, onde há uma confusão de nomenclaturas, o SAS possui uma documentação simples e bem definida, a qual é mantida pela *SCSI Trade Association*[5].

O SAS é uma interface rápida, provendo taxas de transmissão bastante altas, podendo ser usadas em transmissões a longa distância quando fibras óticas são utilizadas. Além de ser expansível através de conectores de expansão e possuir compatibilidade com o padrão SATA[18]. O que torna a interface SAS atraente foi que ela herdou o conjunto de comandos das SCSI Paralelas e o modo *Full Duplex* do *Fibre Channel*. Sua grande vantagem em relação ao padrão *Fibre Channel* é que ele permite o uso de discos de várias taxas de transmissão, usando a taxa máxima de cada dispositivo[22].

A segunda geração do SAS, o SAS-2, manteve a compatibilidade com as interfaces SAS-1 e dobrou a taxa de transmissão para 6Gb/s e adicionou expansores autos configuráveis. Além de incluir gerenciadores de tráfego. Um rascunho da documentação[6][19], foi liberado em 2008, contendo as possíveis especificações SAS-2.

Os equipamentos SAS ainda são bem caros comparados com os equipamentos SATA, mas essa diferença vem caindo gradativamente. O padrão SCSI Paralelo, vem sendo substituído, mas ainda é cedo para definir quais das interfaces irão substituí-lo. Devido a compatibilidade entre os equipamentos SAS e SATA algumas empresas optam pela criação de servidores de armazenamento híbridos, balanceando, assim, custo e desempenho.

3. ARRANJOS DE DISCOS

Uma inovação tecnológica que trouxe melhorias para a dependabilidade e a performance de sistemas de armazenamento foram os arranjos de disco. Inicialmente a ideia era que a vazão potencial poderia ser aumentada, mantendo muitos discos ao invés de simplesmente obter discos de maior capacidade. Espalhando os dados sobre múltiplos discos isso automaticamente forçaria o acesso à diversos discos se o con-

junto de dados fosse grande o suficiente, aumentando a vazão do sistema.

Um questão importante, que deve ser considerada é que um grande número de discos em um arranjo podem apresentar mais falhas do que um pequeno número de discos de maior capacidade, quando a confiabilidade nos discos é a mesma. Sendo assim a confiabilidade pode ser melhorada adicionando discos redundantes de forma a tolerar falhas. O único perigo é que um outro disco falhe durante o tempo médio de reparo (*Mean Time To Repair* - MTTR). No entanto o tempo médio de falhas (*Mean Time To Failure* - MTTF) é medido em anos enquanto o MTTR é medido em horas. Sendo assim a redundância é uma medida confiável para muitos discos em relação ao uso de um único disco[21]. Um estudo quantitativo detalhado[32], sobre as falhas de disco em arranjos, realizado pela Google Inc., mostra estatisticamente essa afirmação.

Arranjos de discos redundantes ficaram conhecido pelo acrônimo RAID, que originalmente significava *Redundant Array of Inexpensive Disks* (Conjunto Redundante de Discos Baratos), no entanto alguns preferem o uso de *Redundant Array of Independent Disks* (Conjunto Redundante de Discos Independentes).

A tolerância a falhas e a alta vazão tornaram o RAID uma solução atrativa. Quando combinado com as vantagens de discos de menor tamanho e conseqüentemente com o baixo consumo desses discos, especialmente nos dias de hoje, onde o consumo de energia vem se tornando uma preocupação crescente, fizeram com que o RAID dominasse o mercado de sistemas de armazenamento de larga escala.

A tecnologia RAID é classificada por níveis. De fato o artigo[31] que introduziu o termo RAID utilizou uma classificação numérica que tornou-se bastante popular. Os níveis RAID padrões são bem documentados, e são classificados da seguinte forma:

- **RAID 0** - Não há redundância, alguns autores apelidam de JBOD (*Just a Bunch Of Disks*), no entanto os dados podem ser espalhados pelo arranjo de discos. Serve apenas como uma medida padrão para comparar outros níveis de RAID em termos de custo, desempenho e confiabilidade. Alguns autores nem ao menos consideram esse nível como um nível RAID.
- **RAID 1** - Também chamado de espelhamento (*mirroring* ou *shadowing*), existe duas cópias de cada dado. É o mais antigo esquema de redundância e no entanto possui custo elevado também. Como os dados são copiados (espelhados) para um outro disco, é necessário duas escritas para cada dado a ser armazenado no disco o que degrada o desempenho. Algumas controladoras otimizam o desempenho para as leituras, permitindo que o disco leia de apenas um dos discos de forma independente, no entanto essa otimização pode degradar ainda mais o desempenho das escritas no disco[21].
- **RAID 2** - Essa organização foi inspirada nas correções de código de erro das memórias, o ECC (*Error Correcting Codes*)[31]. Isso possibilita certa proteção

adicional, porém o RAID 2 ficou obsoleto pois as novas tecnologias de disco já possuem este tipo de correção internamente.

- **RAID 3** - Nessa organização foi adicionado um disco de paridade extra, contendo a paridade das informações contidas nos discos de dados, um único disco permite recuperar uma falha dos discos. Os dados são organizados em N blocos de dados e um bloco de paridade. Quando ocorre uma falha, os dados de todos os discos são subtraídos os discos de paridade; as informações restantes são aquelas que foram perdidas. Para discos rígidos, essa organização exige que todos os eixos das unidades estejam sincronizados para evitar a latência rotacional[21].
- **RAID 4 e RAID 5** - Utilizam a mesma razão entre os discos de dados e verificação do RAID 3, mas acessam os dados de maneira diferente. A paridade é armazenada sobre a forma de blocos e associada a um conjunto. Esse modo permite reduzir a sobrecarga nas gravações, considerando que a paridade é apenas um somatório das informações; ao gravar novas informações apenas os bits correspondentes a paridade são alterados. A principal diferença entre os dois modos é a distribuição dos dados, sendo que no RAID 5 a paridade associada a cada linha do bloco não é mais restrita à um único disco, permitindo que várias gravações ocorram simultaneamente em vários discos, desde que as unidades de faixa não estejam em um mesmo disco[21].
- **RAID P + Q** - Também chamado de RAID 6, ao contrário dos outros esquemas de RAID, que protegem sobre uma única falha auto-indentificada, quando uma única correção de falha não é suficiente, a paridade pode ser generalizada para se ter um segundo cálculo sobre os dados juntamente com um outro disco de paridade, permitindo a recuperação de uma segunda falha[21]. Existem diversas variações desse esquema, como o RAID 50[9] e o RAID 100, que são implementações híbridas (muitas vezes combinando técnicas de *hardware* e *software*)

A organização de discos RAID é bastante atraente por sua habilidade de recuperação de falhas e a alta vazão do sistema. Quando combinado com discos menores onde o consumo de energia, especialmente nos dias de hoje onde a preocupação com o consumo de energia, principalmente nos grandes *Data Centers*, nos leva a assim chamada "computação verde".

4. OUTRAS TECNOLOGIAS

A crescente demanda por taxas de transmissão cada vez mais altas levaram os fabricantes a concorrer pelo título de dispositivo mais rápido. Outro problema é que cada vez mais o poder computacional dos dispositivos móveis vem aumentando necessitando cada vez mais de um espaço maior de armazenamento. Nesses dispositivos, discos magnéticos comuns são indesejáveis devido às partes eletromecânicas que compõem esses discos, que além do consumo de energia ser muito grande, o que torna inviável para dispositivos móveis como *smartphones* e *tablets*, ainda existe o problema deles serem sensíveis a choques e vibrações.

4.1 Solid State Drives SSD

Solid-State Drive (SSD), é uma unidade de armazenamento não volátil, semelhante a memórias EEPROM (*Electrically-Erasable Programmable Read-Only Memory*), apresentando desempenho bem superior comparados com discos magnéticos. Atualmente esses dispositivos encontraram um nicho no mercado em computadores portáteis e dispositivos móveis. Embora o custo por Gigabyte seja maior comparado ao disco magnético, \$2.42 por GB em 2011[14], ele apresenta um consumo de energia menor, além de ser mais resistente a vibrações e choques. Além disso, o adjetivo que mais tem sido empregado para descrever as novas tecnologias é “verde”. Os dispositivos de memória SSD fazem parte disso, pois são criados para reduzir o consumo energético e garantir melhor eficiência com os recursos que utiliza. Devido a esse fato a tecnologia encontrou um nicho no mercado de dispositivos móveis, onde o consumo de energia é uma questão crítica.

Atualmente existem dois tipos básicos de SSD, os baseados em memórias SLC (*Single Level Cell*) e os baseados em memórias MLC (*Multi Level Cell*). Nas memórias utilizadas nos SSDs, cada célula é composta de um único transistor. A diferença chave entre as memórias SLC e MLC, é que as SLC guardam apenas 1 bit em cada, enquanto as do tipo MLC conseguem guardar 2 bits em cada célula[24][27]. No entanto as SLC são mais confiáveis pois resistem a ciclos de escritas 10 vezes maiores que as MLC, e são mais rápidas também. Porém essas memórias são bem mais caras, sendo mais voltadas para o mercado corporativo.

Devido aos valores cobrados por esses dispositivos, produzir discos com capacidades elevadas, para o usuário comum, ainda não é vantajoso para as fabricantes. No entanto já existem alguns SSDs que contam com 256 GB de capacidade e alguns fabricantes, como a Apple, estão substituindo os discos mecânicos em seus *notebooks*[20][35]. Outro problema é a limitação dos discos: enquanto HDs podem ser sobrescritos infinitas vezes, SSDs podem parar de apagar dados em determinados setores da memória. O fato é que a tecnologia vem ganhando um espaço cada vez maior em dispositivos móveis como *notebooks* e *tablets*, embora os preços estejam caindo gradativamente e mais rápido que o preços dos HDD comuns[14], ainda é muito cedo para afirmar que a tecnologia SSD vai de fato vir a substituir o armazenamento em disco definitivamente.

4.2 Fusion-io

Utilizando sistemas paralelos de discos SSD, a empresa, na primeira geração dos dispositivos, conseguiu alcançar taxas de transferências de 700MB/s para leitura e 500MB/s para escrita[3]. Recentemente a empresa Fusion-io anunciou uma nova memória SSD, o ioFX, com capacidade de 420GB, e taxas de transferências de 700MB/s para escrita e 1.5Gb/s para leitura. Cada unidade funciona com um disco SSD independente. Todas as unidades são acessadas, como um arranjo de discos, isso permite que os discos atinjam esse desempenho.

A empresa também fornece dispositivos voltados para armazenamento de grande escala e servidores. O Fusion ioDrive Duo[3], têm capacidades que variam de 320GB à 1.8TB, consegue alcançar a impressionante marca dos 1.5Gb/s para

leitura e 1.0Gb/s para escrita, com uma latência de acesso médio de 30 μ s (512 bytes). Para alcançar essa velocidade esses dispositivos utilizam o barramento PCI-Express que é amplamente utilizado em placas gráficas. Uma outra solução disponibilizada pela empresa são os ioDrive Octal[4], com capacidades que variam de 5.12TB à 10.24TB e possuem uma largura de banda de aproximadamente 6.0Gb/s para leitura e 4.4Gb/s para escrita utilizando pacotes de 64KB.

No entanto existe algumas deficiências em relação a essa tecnologia. O dispositivos utilizam barramentos PCI-Express, o que torna a quantidade de interfaces que podem ser instaladas dependentes da quantidade de barramentos PCI-Express disponíveis. Além disso o custo desses dispositivos é muito maior do que os discos comuns.

5. CONSIDERAÇÕES FINAIS

Ao avaliar as interfaces e dispositivos disponíveis, é possível observar que o mercado é dividido entre duas classes: uso doméstico e uso corporativo. Acompanhando a evolução das interfaces, é possível observar que interfaces inicialmente voltadas para uso doméstico, como a interface SATA, ultrapassou, tecnologicamente, o desempenho de interfaces paralelas de alto desempenho de uso corporativo, como as interfaces SCSI Paralelas.

As interfaces SAS, embora mais caras, mantém compatibilidade com padrões mais baratos (SATA), onde é possível essas duas tecnologias é possível atingir níveis altos de segurança, disponibilidade e confiabilidade a custos mais acessíveis. Especialmente quando se trata de servidores de armazenamento em larga escala, onde é necessário empregar técnicas para a redundância dos dados (RAID), em que muitas vezes é necessário um equilíbrio entre desempenho e custo.

Novas tecnologias, como as memórias SSD, ultrapassaram as limitações mecânicas dos discos rígidos e apresentaram um desempenho muito superior ao desses discos, ao passo que deram um passo muito grande em relação a diminuição no consumo de energia, que tem se tornado uma preocupação cada vez maior nos tempos da assim chamada “computação verde”. No entanto essa ainda é uma tecnologia bastante cara e pouco expansível, no sentido em que ela depende de barramentos de alta velocidade como o PCI-Express. Mas é possível observar que essa tecnologia encontrou um mercado bastante promissor nos dispositivos móveis e vem ganhando cada vez mais espaço nesse mercado.

Em relação ao uso da tecnologia SSD para servidores, com a introdução do Fusion-io Drive grandes mudanças são esperadas, embora essa tecnologia apresente algumas deficiências e um custo bastante elevado. Ainda é cedo para dizer se a tecnologia SSD irá substituir os discos magnéticos, especialmente no mercado de servidores.

6. REFERÊNCIAS

- [1] SATA International Organization.
<http://www.sata-io.org>. [Accessed Jun. 08, 2012].
- [2] Fibre Channel Arbitrated Loop (FC-AL-2).
www.t11.org/ftp/t11/member/fc/al-2/99-524v0.pdf, 1999.[Online]. [Accessed Jun. 02, 2012].

- [3] Fusion-io. www.fusionio.com/, 20-. [Online]. [Accessed Jun. 07, 2012].
- [4] Fusion-io - ioDrive Octal. www.fusionio.com/platforms/iodrive-octal, 20-. [Online]. [Accessed Jun. 07, 2012].
- [5] STA - SCSI Trade Association. <http://www.scstita.org/>, 20-. [Online]. [Accessed Jun. 03, 2012].
- [6] T10 Working Drafts. <http://www.t10.org/drafts.htm>, 20-. [Online]. [Accessed Jun. 03, 2012].
- [7] 48-bit LBA Technology. <http://www.48bitlba.com/>, 2005. [Online]. [Accessed May. 28, 2012].
- [8] *Introduction to Storage Area Networks*. Vervante, 2006.
- [9] RAID 50 Interleave White Paper. www.lacie.com/download/whitepaper/WPRAIDen.pdf, 2006. [Online]. [Accessed Jun. 02, 2012].
- [10] ATA, IDE and EIDE. <http://en.kioskea.net/contents/pc/ide-ata.php3>, Apr. 2009. [Online]. [Accessed May. 28, 2012].
- [11] The Brief History of ATA and ATAPI. www.ata-atapi.com/hist.html, June 2009. [Online]. [Accessed May. 23, 2012].
- [12] Serial ATA: Meeting Storage Needs Today and Tomorrow. <http://www.serialata.org/documents/SATA-Rev-30-Presentation.pdf>, jun 2009. [Online]. [Accessed Jun. 05, 2012].
- [13] SSD versus Enterprise SAS and SATA disks. <http://www.anandtech.com/show/2739/2>, mar 2009. [Online]. [Accessed Jun. 02, 2012].
- [14] SSD Prices Falling Faster Than HDD Prices. <http://www.tomshardware.com/news/ssd-hdd-solid-state-drive-hard-disk-drive-prices,14336.html>, dec 2011. [Online]. [Accessed Jun. 10, 2012].
- [15] E. Alecrin. Hardware Interface SATA (Serial Advanced Technology Attachment). <http://www.infowester.com/serialata.php>, may 2012. [Online]. [Accessed May. 30, 2012].
- [16] D. Anderson, J. Dykes, and E. Riedel. More than an interface - SCSI vs. ATA. In *2nd Conference on File and Storage Technologies*, volume 1, pages 245–257, 2003.
- [17] B. Chen, T. Lee, K. Peng, and V. Venkataramanan. *Hard Disk Drive Servo Systems (Advances in Industrial Control)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [18] R. Elliott. *Information technology - Serial Attached SCSI - 1.1 (SAS-1.1)*. ANSI, revision 10 edition, sep 2005.
- [19] R. Elliott. *Information technology - Serial Attached SCSI - 2 (SAS-2)*. ANSI, revision 14e edition, sep 2008.
- [20] F. Gugelmin. Novo MacBook Air deve ter SSD com velocidade de 400 Mbps. <http://www.tecmundo.com.br/apple/11309-novo-macbook-air-deve-ter-ssd-com-velocidade-de-400-mbps.htm>, jun 2011. [Online]. [Accessed Jun. 03, 2012].
- [21] J. Hennessy and D. Patterson. *Computer Architecture: A Quantitative Approach*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition, 2002.
- [22] M. Jackson. *SAS Storage Architecture*. Mindshare Inc, Colorado Springs, CO, revision 1.1 edition, sep 2005.
- [23] M. Kawamoto. HDD Interface Technologies. *Fujitsu scientific and Technical Journal*, 42(1):78–92, jan 2006.
- [24] Z. Kerekes. SSD Buyers Guide. <http://www.storagesearch.com/ssd-buyers-guide.html>, 20-. [Online]. [Accessed Jun. 07, 2012].
- [25] C. Kozierok. Integrated Drive Electronics / AT Attachment (IDE/ATA) Interface. <http://www.pcguide.com/ref/hdd/if/ide/stdATA-c.html>, Mar. 2004. [Online]. [Accessed May. 25, 2012].
- [26] L. Lamers. *Information technology - AT Attachment Interface for Disk Drives*. ANSI, revision 4c edition, 1994.
- [27] B. Martin. Solid State Disk (SSD) Drives, may 2009.
- [28] P. McLean. *Information technology - AT Attachment-3 Interface (ATA-3)*. ANSI, revision 7b edition, 1997.
- [29] P. McLean. *Information technology - AT Attachment with Packet Interface Extension (ATA/ATAPI-4)*. ANSI, revision 18 edition, 1998.
- [30] C. Mellor. SAS or enterprise SATA drives? WD says eSATA, Fujitsu says SAS. <http://features.techworld.com/storage/1973/sas-or-enterprise-sata-drives/>, nov 2005. [Online]. [Accessed Jun. 01, 2012].
- [31] D. Patterson, G. Gibson, and R. Katz. A Case for Redundant Arrays of Inexpensive Disks (RAID). *SIGMOD Rec.*, 17(3):109–116, June 1988.
- [32] E. Pinheiro, W. Weber, and L. Barroso. Failure Trends in a Large Disk Drive Population. In *Proceedings of the 5th USENIX conference on File and Storage Technologies*, FAST '07, pages 2–2, Berkeley, CA, USA, 2007. USENIX Association.
- [33] W. C. Preston. *Using Sans and NAS*. O'Reilly & Associates, Inc., Sebastopol, CA, USA, 1st edition, 2002.
- [34] P. M. Ridge and D. Deming. *The Book of SCSI*. No Starch Press, San Francisco, CA, USA, 1995.
- [35] A. Schwartz. Apple's New MacBook Air: The End of the Conventional Hard Drive? <http://www.fastcompany.com/1696623/apple-announces-new-macbook-air>, oct 2010. [Online]. [Accessed Jun. 03, 2012].
- [36] C. Stevens. *Information technology - AT Attachment 8 - ATA/ATAPI Command Set (ATA8-ACS)*. ANSI, revision 18 edition, dec 2006.