

Universidade Federal do Piauí
Campus Senador Helvídio Nunes de Barros
Curso de Sistemas de Informação

Ana Verônica Ferreira de Carvalho

SSD vs HDD: Um Estudo sobre a viabilidade e os impactos do uso de SSDs como dispositivo de armazenamento secundário em detrimento dos HDDs.

Picos-PI
2014

Ana Verônica Ferreira de Carvalho

SSD vs HDD: Um Estudo sobre a viabilidade e os impactos do uso de SSDs como dispositivo de armazenamento secundário em detrimento dos HDDs.

Trabalho de Conclusão de Curso apresentado ao Curso de Sistemas de Informação do Campus Senador Helvídio Nunes de Barros da Universidade Federal do Piauí como parte dos requisitos para obtenção do Grau de Bacharelado em Sistemas de Informação, sob orientação do Prof. Esp. Ivenilton Alexandre de Souza Moura.

Picos-PI

2014

Eu, **Ana Verônica Ferreira de Carvalho**, abaixo identificado(a) como autor(a), autorizo a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação abaixo discriminada, de minha autoria, em seu site, em formato PDF, para fins de leitura e/ou impressão, a partir da data de hoje.

Picos-PI, 08 de agosto de 2014.

Ana Verônica Ferreira de Carvalho

Assinatura

FICHA CATALOGRÁFICA

Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

C331s Carvalho, Ana Verônica Ferreira de.
SSD vs HDD: um estudo sobre a viabilidade e os impactos do uso de SSDs como dispositivo de armazenamento secundário em detrimento dos HDDs / Ana Verônica Ferreira de Carvalho. - 2014.
CD-ROM : il. ; 4 ¾ pol. (53 p.)

Monografia(Bacharelado em Sistemas de Informação) – Universidade Federal do Piauí. Picos-PI, 2014.
Orientador(A): Prof. Esp. Ivenilton Alexandre de Souza Moura

1. Discos de Estado Sólido. 2. HDD. 3. Implicações ao Sistema.
I. Título.

CDD 004.53

Ana Verônica Ferreira de Carvalho

SSD vs HDD: Um Estudo sobre a viabilidade e os impactos do uso de SSDs como dispositivo de armazenamento secundário em detrimento dos HDDs.

Trabalho de Conclusão de Curso apresentado ao Curso de Sistemas de Informação do Campus Senador Helvídio Nunes de Barros da Universidade Federal do Piauí como parte dos requisitos para obtenção do Grau de Bacharelado em Sistemas de Informação, sob orientação do Prof. Esp. Ivenilton Alexandre de Souza Moura.

Data de Aprovação:

05/08/2014

Ivenilton Alexandre de Souza Moura

UFPI-CSHNB

Leonardo Pereira de Sousa

UFPI-CSHNB

Frank César Lopes Vêras

UFPI-CSHNB

Picos-PI

2014

A Deus que me abençoou e sustentou nos momentos difíceis. A meu pai Valter (in memóriam), sua lembrança me incentivou a continuar sempre, e a minha mãe Silvaneide, sua força e coragem me trouxeram até aqui.

Agradecimentos

Agradeço ao meu Deus que é responsável por tudo de bom que sou e conquistei, a minha Mãe Silvaneide que é exemplo de coragem e persistência, ao meu Pai Valter (in memórium) que não está mais presente mas seus conselhos continuam em minha memória, ao meu tio Luciano que me acolheu em sua casa e me adotou durante esses anos, jamais poderei recompensar ou agradecer a sua ajuda. Agradeço ao meu namorado Jânio Coêlho pelo apoio e incentivo, por suas palavras motivadoras em momentos difíceis, a meus avós Francisca e José Valdir e a minha Irmã Dora, aos meus tios e tias, primos e primas que torceram por mim, aos meus amigos e colegas de Universidade que trilharam essa caminhada comigo sempre me ajudando e torcendo por mim e aos irmãos em Cristo pelas orações. Agradeço ao meu orientador Ivenilton Alexandre que prontamente acolheu minhas ideias, minha banca, a professora Juliana Carvalho e coordenadora Patricia Medynna. Agradeço a todos os funcionários da UFPI pelo seu trabalho que é indispensável para que possamos ter condições de estudar. A todos vocês e aos demais que torceram por mim, obrigada, vocês são responsáveis por essa conquista.

Que os vossos esforços desafiem as impossibilidades, lembrai-vos de que as grandes coisas do homem foram conquistadas do que parecia impossível.

Charles Chaplin

É muito melhor lançar-se em busca de conquistas grandiosas, mesmo expondo-se ao fracasso, do que alinhar-se com os pobres de espírito, que nem gozam muito nem sofrem muito, porque vivem numa penumbra cinzenta, onde não conhecem nem vitória nem derrota.

Theodore Roosevelt

Bem-aventurado o homem que acha a sabedoria, e o homem que adquire o conhecimento, pois a sabedoria é mais proveitosa do que a prata e rende mais do que o ouro. Mais preciosa do que os rubis, e tudo o podes desejar não se pode comparar a ela.

Provérbios: 3 13-15

Resumo

Este trabalho mostrará pesquisas mais relevantes sobre o uso de Dispositivos de Estado Sólido (SSDs), sua adoção pelo mercado em substituição aos HDDs e as implicações que esse uso traz ao sistema em vários aspectos. O objetivo desse trabalho é rever e disponibilizar informação acerca do uso dos SSDs e os impactos da sua adoção como dispositivo de armazenamento secundário. Para isso foram usadas as bases de Pesquisa IEEE (Institute of Electrical and Electronics Engineers) Xplorer, ACM (Association for Computing Machinery), e Periódicos Capes além de outras publicações relevantes na área.

Palavras-chave: Discos de Estado Sólido, HDD, Implicações ao Sistema.

Abstract

This work shows most important researches on the use of Solid-State Disk (SSDs), its adoption by the market as a replacement for HDDs and the implications that this use brings to the system in several aspects. The objective of this study is to review and provide information about the use of SSDs and the impacts of their adoption as a device for secondary storage. For this were used the bases of Research IEEE (Institute of Electrical and Electronics Engineers) Xplore, ACM (Association for Computing Machinery), and Capes Periodicals in addition to other relevant publications in the area.

Keywords: Solid-State Disk, HDD, Systems implications.

Lista de Figuras

Figura 1 -	A evolução do Disco Rígido	18
Figura 2 -	Trilhas e Setores do Disco.	19
Figura 3 -	Estrutura o Disco Rígido	20
Figura 4 -	Gravação longitudinal e perpendicular.	22
Figura 5 -	Diferença física entre SSDs e HDDs.	34
Figura 6 -	Criar um arquivo MP3 ler e reproduzir	35
Figura 7 -	Média de leitura e explosão de leitura	35
Figura 8 -	Latência	43
Figura 9 -	Largura de banda	43
Figura 10 -	Relação Base de dados/ Publicações usadas no estudo.	46
Figura 11 -	Melhor caso e pior caso respectivamente	52

Lista de Tabelas

Tabela 1 -	Variações do padrão ATA	23
Tabela 2 -	Comparação de característica básica de SSDs MLC e SLC com HDD .	34
Tabela 3 -	Características Básicas do SSDs usados na pesquisa	40

Lista de abreviaturas e siglas

ANSI	American National Standards Institute
ATA	Advanced Technology Attachment
CG	Garbage Collection
CRC	Cyclic Redundancy Check
DMA	Direct Access Memory
ECC	Error Correcting Code
EEPROM	Electrically erasable programmable read only memory
FC	Fibre Chanel
FTL	Flash Translation Layer
HDA	Hard Drive Assembly
HDD	Hard Disk Drive
IBM	International Business Machines
IDE	Integrated Development Environment
MLC	Multi Level Cell
NCITS	National Committee for Information Technology Standards
PCM	Phase Change Memory
PCRAM	Phase-change RAM
PIO	Programmed input output
RAID	Redundant array of independent disks
RAM	Randomic Access Memory
RMR	Reliability management on reads
RPM	Rotações por Minuto
SAS	Serial Attached
SASI	Shugart Associates Systems Interface
SATA	Serial-Advanced Technology Attachment
SCSI	Small Computer System Interface
SLC	Single Level Cell
SSD	Sólid State Disk
TLC	Three Level Cell
USB	Universal Serial Bus

Lista de símbolos

μ Nanossegundos

Sumário

1	Introdução	15
1.1	Objetivos	15
1.1.1	Objetivos Gerais	15
1.1.2	Objetivos Específicos	15
1.2	Estrutura da Monografia	16
2	O HDD	17
2.1	A História do HD	17
2.2	Estrutura do HD	19
2.3	Leitura e Gravação dos dados	21
2.4	Interface com o disco	22
2.4.1	ATA	23
2.4.2	SATA	24
2.4.3	SCSI	25
2.4.4	SAS	26
3	SSD baseado em <i>FLASH NAND</i>	27
3.1	Por dentro do SSD	28
3.1.1	Operação de Leitura e Escrita	28
3.1.2	Comando Apagar	29
3.1.3	FTL (<i>Flash Translator Layer</i>)	29
3.1.4	<i>Garbage Collection</i> - Coleta de Lixo	30
3.1.5	Nivelamento de Desgaste - <i>Wear Leveling</i>	31

3.1.6	Fragmentação de Bloco	32
3.1.7	TRIM	33
3.2	SSD vs HDD	33
3.3	Abordagens Híbridas	36
3.3.1	SSD como cache	36
3.3.2	Combinação de SLC, MLC e TLC em SSDs	36
3.3.3	SSD e PCM	37
3.4	Uso de SSDs em <i>Games</i>	37
4	Revisando expectativas acerca dos SSDs	39
4.1	Leitura	40
4.2	Gravação	41
4.3	Comando TRIM	44
5	Resultados	45
5.1	Bases de Pesquisa	45
5.2	Trabalhos Futuros	47
6	Conclusão	48
	Referências	49
	Anexo A – Comparação de desempenho de leitura/escrita sequencial e aleatória	51
	Anexo B – SSD e HDD- Comparação de latência	52

1 Introdução

Nos últimos anos, surgiu no mercado o que se considera o futuro do armazenamento de dados, os Discos de Estado Sólido (SSD). Esse tema vem gerando grande expectativa e atraindo muitos usuários que buscam a rapidez e resistência dos SSDs. Constituído de circuitos eletrônicos os SSDs são feitos a partir de *chips Flash NAND* e não possuem partes móveis.

Por não possuir partes móveis os SSDs são significativamente mais rápidos que os Discos Rígidos (HDs), consomem menos energia e são mais resistentes a impactos. Os SSDs podem ser constituídos de células MLC (*multi-level-cell*) que armazenam 2 bits em cada nível, células SLC (*single-level-cell*) que armazena apenas 1 bit por nível e, recentemente, surgiu a TLC (*three-level-cell*) que armazena três bits por nível. As células SLC são mais caras, mais seguras e duráveis, já as MLC e as TLC são mais baratas, porém mais inseguras e tem menor tempo de vida.

Devido as características particulares dos SSDs foram desenvolvidos recursos que devem ser implementados em uma camada de *firmware* que emula a *interface* de disco rígido. Esses recursos serão explicados nos capítulos seguintes. Essa característica e esses recursos podem ser prejudiciais ao desempenho do sistema e, por causa disso, a adoção dos SSDs como dispositivo de armazenamento secundário ainda é questionada por alguns pesquisadores.

1.1 Objetivos

1.1.1 Objetivos Gerais

O objetivo deste trabalho é esclarecer as diferenças entre HDDs e SSDs e estudar a viabilidade do uso de SSDs para armazenamento secundário e as implicações desse uso.

1.1.2 Objetivos Específicos

Como resultado da pesquisa realizada podemos esclarecer o funcionamento dos SSDs em alguns testes realizados em dispositivos diferentes de vários fabricantes, podemos também compreender o desempenho em situações específicas através de gráficos, sendo possível assim entender as falhas de desempenho em algumas situações. Esse trabalho também fornecerá bases para avaliar a viabilidade de uso dos SSDs.

1.2 Estrutura da Monografia

Após a introdução que relatou sobre as motivações e objetivos para o desenvolvimento do presente trabalho, serão apresentados os próximos capítulos que estão organizados da seguinte forma:

- **CAPÍTULO 2-** Traz um apanhado geral sobre os HDDs, a principal mídia de armazenamento secundário utilizado atualmente. Aborda também aspectos históricos, evolução, estrutura, funcionamento e *interfaces* de conexão.
- **CAPÍTULO 3-** Trata dos Discos de Estado Sólido baseado em *flash NAND*. Esse capítulo traz os dados dos SSDs sob a visão de alguns autores, arquitetura, apresenta alguns testes superficiais, uma seção de comparação com os HDDs relata também sobre o uso de discos híbridos e ainda traz um breve relato sobre o uso de SSDs em *games*.
- **CAPÍTULO 4-** Este capítulo aborda uma vertente polêmica: nele é apresentada a visão de alguns autores sobre os SSDs que, baseados em testes mais aprofundados com maior variedade de dispositivos, descreveram o desempenho que os SSDs apresentam diante de algumas situações. Mostram também o impacto e o resultado da execução da Coleta de Lixo e do comando TRIM.
- **CAPÍTULO 5-** Mostra os resultados da pesquisa e as bases de dados que foram consultadas, quantificando as publicações utilizadas.
- **CAPÍTULO 6-** Por fim as conclusões acerca do trabalho.

2 O HDD

Disco rígido, ou HD, é a parte do computador na qual são armazenados dados de forma não-volátil, ou seja, dados que não são perdidos quando o computador é desligado. Os dados armazenados no HD são acessados de forma mecânica: uma cabeça de leitura é movida até a parte do HD da qual se deseja coletar dados, por essa razão, o acesso a esses dados ocorre de forma muito mais lenta se comparado ao acesso a dados armazenados na memória principal (volátil).

O conceito de armazenamento de dados surgiu com Joseph Marie Jacquard, que usou placas perfuradas para tecelagem em meados de 1801. O Recenseamento americano de 1880 levou quase dez anos. Tentando agilizar esse processo o governo americano promoveu um concurso para uma proposta de melhora nesta situação. O estatístico H. Hollerith apresentou uma máquina de tabulação que utilizava cartões perfurados. O censo de 1890 foi computado em alguns meses. Hollerith fundou a *Tabulating Machine Company*, que acabou dando origem à IBM (CARNEIRO, 2006).

Desde o princípio do armazenamento de dados com o cartão perfurado tivemos muitas evoluções e agora dispomos de vários dispositivos para esse fim, o mais presente e mais usado para armazenar dados de forma persistente é o HDD (Drive de Disco Rígido), que será abordado em aspectos gerais nesse capítulo.

2.1 A História do HD

O primeiro uso comercial de discos rígidos, de acordo com Shiratori et al. (2012) é datado de 1956, com o lançamento de um computador chamado IBM 305 RAMAC que possuía uma unidade de armazenamento em discos chamado IBM 350. Esta unidade tinha um conjunto de 50 discos de 24 polegadas (aproximadamente 61 cm) de diâmetro cada, totalizando 5 MB e duas cabeças de leitura/escrita. Por ser muito grande e possuir apenas duas cabeças de leitura/escrita para percorrer todos os discos (todos os discos tinham informações gravadas em ambas as faces), o tempo médio de acesso aos dados armazenados era muito grande (por volta de 600 milissegundos). O aluguel do IBM 305 RAMAC custava US\$3.200 por mês.

A tecnologia aplicada foi de um disco magnético revestido em que os dados pudessem ser escritos, lidos e apagados várias vezes. Tempos de acesso aceitável e taxas de transferência de dados permitidos para desempenho máximo do sistema a um custo acessível por MB

armazenado (GROCHOWSKI; HOYT, 1993).

Com o uso foi-se percebendo que o Disco precisaria de melhoramentos, um desses pontos era o tempo de acesso, percebeu-se que uma cabeça de leitura não era suficiente, pois tornava a leitura e escrita muito lenta. Por isso, como cita Shiratori et al. (2012), o IBM 1301 *Disk Storage Unit*, anunciado em 1961, era idêntico ao sistema de armazenamento do IBM 305 RAMAC, exceto pelo fato de possuir uma cabeça de leitura/escrita para cada superfície dos discos, o que tornava o acesso aos dados muito mais rápido.

Em 1973 a IBM fez um grande avanço e lançou o chamado IBM 3340, também conhecido como *Winchester*. A grande diferença entre o IBM 3340 dos anteriores é que ele possuía uma cabeça de leitura menor e muito mais leve, permitindo que ela flutuasse muito mais próximo à superfície do disco. Na Figura 1 podemos ver a evolução dos discos rígidos.

O IBM 3340 tinha dois discos, um fixo e um removível, ambos com capacidade de 30 MB. Os desenvolvedores do IBM 3340 chamavam-no de “30-30”. O apelido *Winchester* apareceu devido à comparação do “30-30” com um rifle chamado *Winchester 30-30*. Graças ao IBM 3340, por muito tempo os discos rígidos foram chamados de *Winchester*. Embora o produto inicial fosse composto por dois discos de 30 MB cada, foram vendidas apenas versões com discos rígidos compostos por dois discos de 35 MB ou 70 MB cada. Hoje, 35 anos após o *Winchester*, a capacidade dos discos rígidos domésticos já alcançou a marca de Terabytes (aproximadamente 10^{12} bytes), cerca de 200.000 vezes maior do que o *Winchester*. A velocidade de acesso aos dados também foi melhorada: atualmente o tempo médio de acesso é de algumas poucas dezenas de microssegundos (10^{-6} segundos) (SHIRATORI et al., 2012)



Figura 1 – A evolução do Disco Rígido

2.2 Estrutura do HD

Um disco magnético consiste em uma coleção de lâminas, de acordo Oliveira e Correa (2011) elas variam em geral de 1 a 12 unidades, girando em um eixo em 3.600 a 15.000 rotações por minuto (RPM). Essas lâminas, que recebem o nome de pratos, são discos de metal ou vidro recobertos com um material magnético de gravação em ambos os lados, de forma que 10 pratos apresentam 20 superfícies para a gravação. O diâmetros dos discos mais usados varia com pratos de 2,5 ou 3,5 polegadas. Os discos de 2,5 visam o mercado de computadores portáteis enquanto os de 3,5 os computadores de mesa e servidores. Tradicionalmente, os discos com pratos de maior diâmetro têm desempenho mais alto, enquanto os discos menores têm preço mais baixo.

A superfície do prato se divide em círculos concêntricos, denominados trilhas, além das trilhas existem também os cilindros e os setores. Oliveira e Correa (2011) diz que em geral, existem de 5.000 a 30.000 trilhas em cada superfície. Ao conjunto de trilhas sobrepostas em superfícies diferentes se dá o nome de cilindro. Por sua vez, cada trilha é dividida em setores que contêm as informações. Uma trilha pode ter de 100 a 500 setores. Um setor é a menor unidade que pode ser lida ou gravada. Os *mainframes* da IBM, por exemplo, permitem o usuário selecionar o tamanho dos setores, embora a maioria dos sistemas fixe seu tamanho, normalmente em 512 *bytes* de dados. A sequência registrada na mídia magnética é um número de setor, um intervalo, as informações para esse setor incluindo código de correção de erros, um intervalo, o número de setor do próximo setor e assim por diante. A Figura 2 mostra as trilhas e setores do disco.

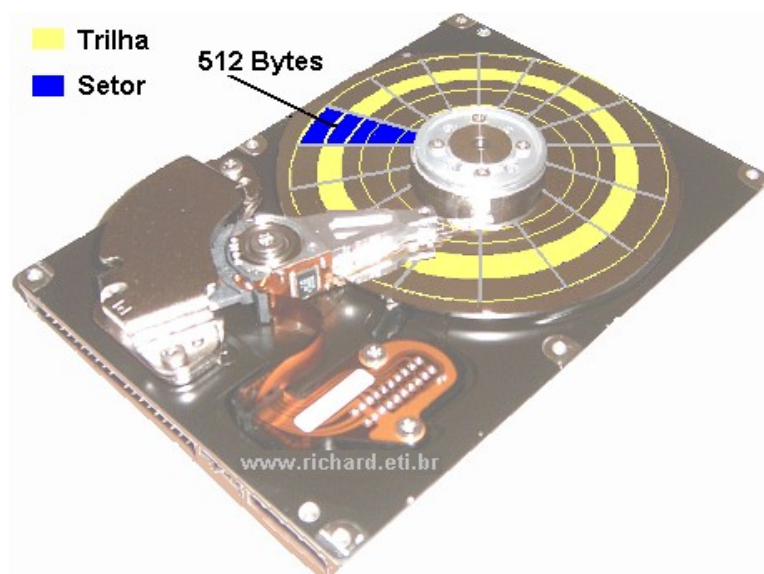


Figura 2 – Trilhas e Setores do Disco.

De acordo com Shiratori et al. (2012) O *Drive* de Disco Rígido é formado basicamente de duas partes: uma ou mais placas lógicas de circuito impresso e um compartimento selado, conhecido como HDA (*Hard Drive Assembly*). As principais partes encontradas na placa lógica são o circuito controlador, o *driver* dos motores e um *chip* de memória RAM (*Random Access Memory*), conhecido como *buffer*. O circuito controlador é responsável por todo o funcionamento da parte mecânica do HD e também pela troca de dados entre o disco e o computador. O *driver* dos motores é um amplificador de corrente, que recebe informações do circuito controlador e repassa essas informações para os motores, mas com uma corrente ampliada. O *buffer* é de grande importância para o desempenho do HD. Quanto maior for a capacidade de armazenamento do *buffer*, maior será a taxa de transferência do mesmo. A Figura 3 mostra a estrutura física dos discos rígidos.



Figura 3 – Estrutura o Disco Rígido

Disponível em: www.blog.techcore.co.mz

O HDA é um compartimento totalmente selado, para evitar entrada de partículas externas, já que, devido à sensibilidade dos discos e à alta velocidade que eles ganham em funcionamento, qualquer partícula pode danificar permanentemente os discos. Dentro do HDA são encontradas as partes físicas do HD, que basicamente são formadas pelos discos magnéticos, pelo motor, pelos braços e pelas cabeças de leitura/escrita, e pelo atuador.

O material não magnético que forma a base do disco recoberto pelas duas finas camadas de material magnético de cada lado do disco, e normalmente é feito de óxidos de ferro ou

cobalto, ficam sobrepostos uns aos outros formando um cilindro presos ao eixo do motor, que é o responsável por fazer os discos girar.

Quanto maior a velocidade do motor, melhor será o desempenho do HD. O dispositivo responsável por gravar os dados nos discos magnéticos é a cabeça de leitura/escrita. Existe uma cabeça para cada uma das superfícies dos discos e elas ficam localizadas na ponta do braço, que é responsável por posicionar as cabeças nas posições corretas. As cabeças são formadas por uma espécie de eletroímã, que é capaz de mudar de polaridade numa velocidade muito rápida. O atuador, também chamado de *voice coil*, tem a função de movimentar o braço entre os discos para que a cabeça possa agir. O atuador dispõe de uma bobina e ímãs, que dão o movimento. Quanto maior for a velocidade do atuador menor será o tempo gasto para posicionar a cabeça na posição correta, portanto melhor será o desempenho do HD (SHIRATORI et al., 2012).

2.3 Leitura e Gravação dos dados

Para ler ou gravar um setor, o controlador de disco envia um comando para mover o braço sobre a trilha apropriada. Essa operação é chamada busca, ou *seek*, e o tempo necessário para mover o braço até a trilha desejada denomina-se tempo de busca. Oliveira e Correa (2011) explica o processo de gravação dos *bits* que ocorre através da magnetização do material ferromagnético depositado na superfície dos discos, discos mais antigos utilizam óxido férrico como superfície de gravação, enquanto discos mais novos utilizam liga à base de cobalto. Mudanças sequenciais na direção da magnetização representam padrões de *bits* que carregam a informação armazenada. A cabeça de leitura e escrita é capaz de detectar esses padrões no caso de operações de leitura e mudar a magnetização no caso de escritas.

A superfície magnética de cada disco é conceitualmente dividida em regiões chamadas de domínios magnéticos que possuem área menor que 1 micrometro quadrado. Nos discos mais antigos as regiões eram magnetizadas de forma horizontal ou paralela à superfície do disco. No início de 2005 os discos começaram a ser gravados com uma orientação perpendicular, que permite domínios magnéticos menores, aumentando assim a densidade de bits por área (OLIVEIRA; CORREA, 2011). A Figura 4 apresenta o conceito.

Devido a natureza policristalina do material magnético, cada região é composta de algumas centenas de grãos magnéticos ou cristalites, que possuem tamanho típico de 10nm. Assim, uma região magnética é composta de grãos que quando alinhados formam um dipolo magnético que carrega a informação armazenada(OLIVEIRA; CORREA, 2011).

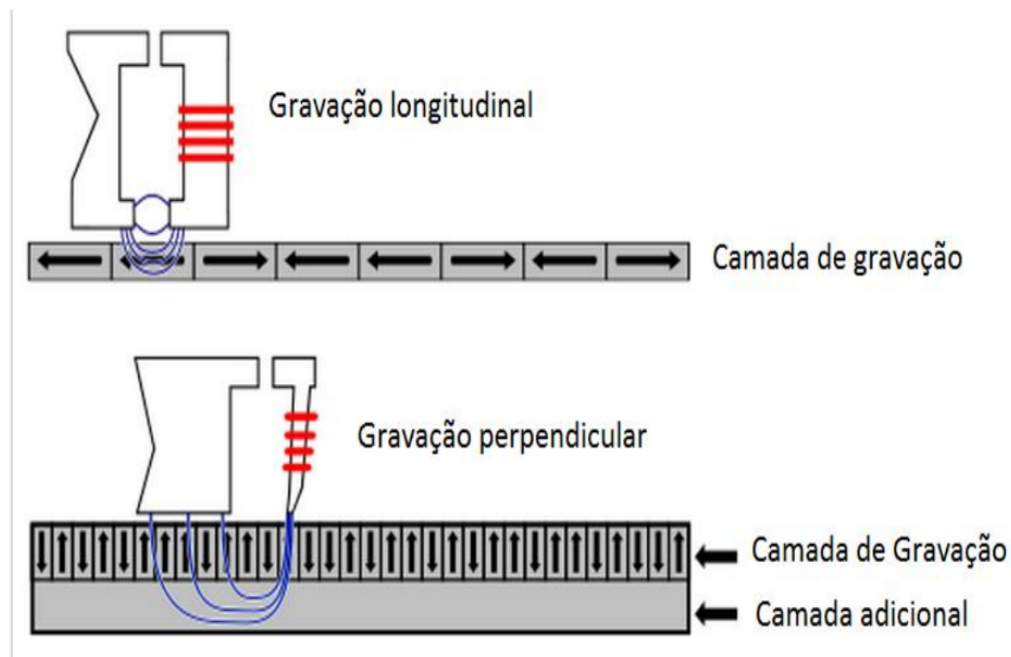


Figura 4 – Gravação longitudinal e perpendicular.

Disponível em: (OLIVEIRA; CORREA, 2011)

O dispositivo de leitura faz o oposto do dispositivo de gravação: quando ele passa por um campo magnético gerado pela superfície do disco, uma pequena corrente elétrica é induzida neste dispositivo, e essa corrente é transmitida até a parte lógica do HD, onde ela é interpretada e um dígito binário é identificado. Esse dispositivo de leitura é protegido por um escudo eletromagnético, para que o campo magnético de outra região do disco não influencie no processo de leitura (SHIRATORI et al., 2012).

2.4 Interface com o disco

Para que os dados sejam trazidos do disco até a memória principal, é necessário um padrão de comunicação eficiente e confiável. Esse papel é desempenhado pela *interface* que é muito importante para desempenho do disco.

Com o aumento da capacidade de armazenamento dos discos magnéticos a *interface* de transferência de dados com o sistema de computação precisa evoluir constantemente para permitir que os dados, cada vez em maior densidade, possam fluir entre o disco e as hierarquias superiores de memória. Nas próximas seções serão mostradas algumas das *interfaces* mais usadas e suas características.

2.4.1 ATA

Em 1986, a *Western Digital, Compaq Computer e Control Data Corporation* fizeram uma parceria para construir o que hoje é conhecido como *interface* ATA. No início, algumas empresas comercializavam esta *interface* com o nome de IDE, permanecendo esta nomenclatura durante alguns anos. Para evitar confusão, o termo ATA/IDE foi adotado como padrão (TAMPELINE, 2011).

A *Advanced Technology Attachment* (ATA) é um padrão para interligar dispositivos de armazenamento, como discos rígidos e drives de CD-ROMs, que opera como um barramento compartilhado entre os dispositivos. A evolução do padrão ao longo do tempo criou diversas variações de *interfaces* como mostra tabela a seguir criada a partir da publicação de (TAMPELINE, 2011)

Tabela 1 – Variações do padrão ATA

Padrão	Características
ATA-1	Este padrão possui modo de endereçamento de 28 <i>bits</i> , sendo possível endereçar até 228 setores de 512 <i>bytes</i> , ou seja, 128 GB. Os modos de transferência são definidos em PIO (<i>Programmed I/O</i>), 0, 1 e 2, que tem taxas de transferência que variam de 3,3 a até 8,3 MB/s. No ATA-1, também foram definidos os modos DMA (<i>Direct Access Memory</i>) <i>single-word</i> 0, 1 e 2 e modo DMA <i>multiword</i> 0.
ATA-2	Também conhecido como EIDE (<i>Enhanced IDE</i>), se caracteriza pela completa implementação do DMA, apresentando taxas de transmissão de 16,67 MB/s. O suporte de gerenciadores de energia, o qual permite a redução das rotações dos discos quando ociosos, foi introduzido nesta especificação.
ATA-3	A especificação ATA-3, publicada no padrão ANSI X3.298-1997, é uma revisão dos padrões ATA a ATA-2. ATA-3 acrescenta ao padrão ATA uma tecnologia de autodetecção de condições adversas conhecida como: <i>Self-Monitoring Analysis and Reporting Technology</i> (SMART).

ATA/ATAPI-4	A extensão ATAPI foi, publicada em 1998 no padrão NCITS 317-1998, ou ATA/ATAPI-4. Além dos modos de transferência especificados pelos padrões ATA anteriores, ATA/ATAPI-4 introduziu os modos de transferência que utilizam UltraDMA 0, 1 e 2 (UDMA). O UDMA otimiza a taxa de transmissão utilizando tanto a subida do <i>clock</i> quanto a descida, aumentando a velocidade de transferência em 100%. No entanto o UDMA 1 e 2 nunca foram implementados pelos fabricantes. Outra melhoria introduzida foi o acréscimo de códigos CRC (<i>Cyclic Redundancy Check</i>) para detecção de erros nas transmissões.
ATA/ATAPI-5	Publicada em 2000, no padrão NCITS 340-2000. Esta atualização proporcionou novos modos de transmissão UDMA 3 e 4 (taxas máximas de 44.4MB/s e 66.7MB/s, respectivamente). Porém, para conseguir maiores velocidades (acima de 33MB/s), foram adicionados mais 40 fios de aterramento (cabo Ribbon).
ATA/ATAPI-6	Publicada em 2002 no padrão NCITS 347-2002. Essa <i>interface</i> suporta o modo UDMA 5 (100MB/s) e disponibiliza recursos de controle automático de emissão de ruídos no disco rígido (<i>Hard Disk Noise Reduction</i>).
ATA/ATAPI-7	Publicada em 2002 no padrão NCITS 361-2002. A transmissão é feita em 16 bits em paralelo, em modo <i>half-duplex</i> , utilizando um cabo Ribbon.
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 de Estado Sólido(SSD), um rascunho do conjunto de comandos foi publicado em 2006. Uma das novas funcionalidades é a tecnologia <i>Hybrid Drive</i> , que utiliza uma memória cache para aumentar a velocidade de acesso para arquivos críticos do SO (Sistema Operacional).

2.4.2 SATA

A interface SATA (Serial ATA), teve como objetivo inicial transformar o IDE/ATA em serial. Seu desenvolvimento foi iniciado pelo SATA *Working Group* em 2002, promovendo resultados promissores com relação a taxa de transmissão. Apesar de possuir todas as funcionalidades de suas antecessoras ATA, esta faz uso de apenas sete fios, sendo 2 para transmissão,

2 para recepção e 3 de aterramento (TAMPELINE, 2011).

O padrão ATA faz transferência de dados de forma paralela, ou seja, transmite vários *bits* por vez, como se estes estivessem lado a lado. No SATA, a transmissão é em série, tal como se cada bit estivesse um atrás do outro. A transmissão paralela de dados (geralmente com 16 *bits* por vez) é mais suscetível ao ruído dos próprios sinais. Tampeline (2011) explica que devido à existência de poucos fios, a SATA pode trabalhar em altas frequências e em baixas tensões, evitando interferências e permitindo taxas de transmissões mais altas. A velocidade inicial do padrão SATA-1 foi 150 MB/s (denominada comercialmente por SATA-150).

Serial ATA, SATA ou S-ATA é o sucessor da tecnologia ATA. Diferentemente dos discos rígidos ATA, que transmitem os dados em um barramento de quarenta ou oitenta fios paralelos, os discos rígidos SATA transferem os dados em série através de uma conexão ponto a ponto. Os cabos Serial ATA são formados por dois pares de fios (um par para transmissão e outro par para recepção) usando transmissão diferencial, e mais três fios para o aterramento, totalizando 7 fios. Isso permite o uso de cabos com menor espessura que facilitam a ventilação do gabinete (OLIVEIRA; CORREA, 2011).

Com o intuito de criar diferenciais competitivos, empresas buscam inserir novas funcionalidades a *interface* ATA. Tampeline (2011) caracteriza algumas delas:

- eSATA (*External SATA*) desenvolvido pela *SATA International Organization* que permite a comunicação por meio da *interface* SATA com dispositivos externos, especificando técnicas para aumentar o tamanho do cabo SATA (de 1 para 8 metros) sem muita perda na taxa de transferência.
- ATA PM (*SATA Port Multipliers*) que permite a inserção de vários em uma única porta;
- *Enterprise* SATA desenvolvida pela *Western Digital* e voltada a requisição de alto desempenho (concorrente do SAS e FC) ;

Devido a grande competição, muitas especificações não são liberadas ao público, dificultando análises mais detalhada de tais *interfaces*. Muitas documentações somente são encontradas quando certo projeto já não é mais classificado como diferencial competitivo

2.4.3 SCSI

De acordo com Tampeline (2011), as raízes do desenvolvimento e padronização do SCSI remontam a 1979, quando o fabricante de discos *Shugart* buscava por uma nova e universal *interface* de discos. Seu principal objetivo era desenvolver uma *interface* que suportasse endereçamento de bloco lógico (ao invés do endereçamento: cabeça / cilindro / sector), transferências paralelas de alto desempenho e comandos de controle genéricos. Esta *interface* foi denominada

de *Shugart Associates Systems Interface (SASI)*. No final de 1981, *Shugart Corporation* juntamente com a *NCR Corporation*, apresentaram o SASI ao comitê ANSI, que o aceitou como projeto e foi denominado de SCSI.

Diversas melhorias foram adicionadas ao SCSI, as principais delas foi a habilidade de competir pelo barramento e liberá-lo temporariamente durante a execução de um comando e a inserção de um adaptador de *host* no barramento SCSI, responsável por intermediar as transações entre os dispositivos e a memória principal

2.4.4 SAS

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

A primeira versão opera a 3 Gbps atingindo uma capacidade real de 300 MBps. Já a segunda versão, lançada no início de 2009 opera a 6 Gbps. Um das características mais atraentes do SAS é a capacidade de operar com discos SATA-II para aplicações de mais baixo custo. Em contrapartida, o barramento SATA-II não tem a capacidade de operar com discos SAS. Para aplicações de mais alto desempenho é possível agregar 4 *interfaces* SAS atingindo assim taxas de transferências de até 2.4 GBps. Outro grande diferencial do SAS é a capacidade de expandir a conectividade para diversos discos em um mesmo barramento. Uma característica que somente surgirá no SATA em sua terceira geração (OLIVEIRA; CORREA, 2011)

3 SSD baseado em *FLASH NAND*

Discos magnéticos têm sido os meios preferenciais para armazenamento de dados por várias décadas. No entanto, a arquitetura de unidades de disco atualmente enfrenta dois desafios segundo (JOSE; PRADEEP.C, 2013).

O primeiro é o desempenho. Os drives de disco são sistemas altamente complexos que consistem em eletrônica e componentes mecânicos, devido a lenta latência mecânica, o subsistema de I / O do disco tem sido repetidamente identificado como um grande gargalo para o desempenho do sistema em muitos sistemas de computação. O segundo é o consumo de energia. *Data center* tradicional que usa unidades de disco como o armazenamento, não está adequado para computação intensiva de dados, devido à sua elevada potência de consumo, a alta latência e baixa largura de banda. A memória *flash* é gradualmente substituição da unidade de disco rígido no centro de dados por causa de seu menor consumo de energia, acesso mais rápido ao acaso, e maior resistência ao choque.

Memória *Flash* é um tipo particular de EEPROM (siga em inglês para "Memória Somente de Leitura Programável Apagável Eletricamente"). E nada mais que um *chip* de memória que mantém informações armazenadas sem a necessidade de uma fonte de energia. Ela é frequentemente usada em eletrônicos portáteis, tais como dispositivos de música digital (MP3), *smartphones* e câmeras digitais, bem como nos dispositivos de armazenamentos removíveis (os populares *pendrives* e cartões de memória).

Segundo Jose e Pradeep.C (2013), a memória *flash* está em concorrência a nível de custos com o disco magnético. A memória *flash* é amplamente utilizada para armazenamento de código e dados de produtos eletrônicos de consumo, devido à sua característica versátil como a não - volatilidade, de estado sólido, confiabilidade, baixo consumo de energia e resistência a impacto. Seu baixo consumo de energia faz com que seja um meio de armazenamento atraente para *laptops e notebook*.

Deng e Zhou (2010) ressalta que o tradicional termo SSD refere-se a dispositivos semicondutores. Portanto, uma unidade SSD indica a utilização de semicondutores para emular as *interfaces* externas de E/S (*Fibre Channel*, SCSI, SATA e USB, etc.) das unidades de disco magnético convencional. O SSD normalmente consiste de DRAM ou memória volátil, ou memória *flash NAND*, memória não-volátil. A DRAM SSD requer uma bateria interna e cópia em disco rígido para garantir persistência dos dados por isso a maior parte dos SSDs atuais empregam memória *flash* não-volátil como mídia de armazenamento.

Este trabalho focará apenas nos SSDs baseado em memória *flash NAND* por que é mais acessível e utilizada na maioria dos SSDs.

A unidade de estado sólido (SSD), às vezes chamado de disco estado sólido ou disco eletrônico, é um dispositivo de armazenamento de dados que utiliza montagens de circuitos

integrados como memória para armazenamento de dados persistentemente. SSDs não utilizam quaisquer componentes mecânicos móveis, que os distingue dos discos magnéticos tradicionais, tais como unidades de disco rígido ou disquete, que são dispositivos eletromecânicos contendo discos giratórios e cabeça móvel de leitura/gravação.

Os componentes do SSD são basicamente *chips* de memória *flash*, o controlador SSD, um *buffer* DRAM e a *interface* com o *host*. Comparado com discos eletromecânicos os SSDs são tipicamente menos suscetíveis a choque físico, são silenciosos, têm menor tempo de acesso e latência (Período de tempo para que um requisição seja respondida) mas são muito mais jovens em tecnologia e mais caros por unidade de armazenamento. Nas seções seguintes será possível conhecer mais detalhadamente os SSDs.

3.1 Por dentro do SSD

As memórias utilizadas nos SSDs são extremamente simples, sendo cada célula composta de apenas 1 transistor, que nas SLC (*Single-Level-Cell*-Células de memória *Flash* NAND com apenas um nível) guarda apenas 1 *bit*, enquanto as do tipo MLC (*Multi-Level-Cell*-Célula de memória *Flash* NAND com dois níveis) conseguem guardar 2 *bits* em cada célula. Devido à maior simplicidade no funcionamento, as SLC são mais confiáveis: resistem a 100.000 ciclos de escrita, enquanto as MLC resistem a 10.000 ciclos; e mais rápida (principalmente em escrita), porém mais caras, além de oferecerem menor capacidade de armazenamento que as MLC (TAMPELINE, 2011). Essa segmentação define claramente o público alvo de cada tipo de memória, as SLC são voltadas ao mercado corporativo onde se prioriza desempenho e confiabilidade, mesmo que a um custo maior. Enquanto as MLC são voltadas ao consumidor doméstico, para o qual o custo é o fator decisivo.

Uma vez que um pacote de memória *flash* individual só fornece largura de banda limitada (cerca de 40MB/s), memória *flash* SSDs normalmente são baseados e construídas sobre uma matriz de memória *flash*. Como páginas lógicas pode ser distribuídas sobre vários *chips* de memória *flash*, semelhante a um típico armazenamento de RAID-0, alta largura de banda pode ser alcançado por meio do acesso paralelo. Um barramento serial de I/O liga o pacote de memória ao controlador. O controlador recebe e processa pedidos de *host* através de uma *interface* de conexão, como SATA, e emite comandos e transfere dados de/para o módulo de memória *flash* (CHEN et al., 2009).

3.1.1 Operação de Leitura e Escrita

Dirik e Jacob (2009) explica como é feita a operação de leitura em memórias *Flash* NAND; a unidade de acesso menor é uma página, para ler uma página um comando de leitura é

enviado para o registrador de comando e, em seguida, escreve um número de bloco e número da página dentro do bloco para o registro de endereço. Os dados da página completas (2 Kbytes) vão ser acessados em $25\mu\text{s}$ (nanossegundos) e será carregado no registrador de dados. Depois disso, os dados podem ser lidos a partir do registrador de dados por meio do barramento I/O de 8 *bits*. Se páginas sequenciais precisam ser acessadas dentro de um mesmo bloco, o comando de leitura pode ser utilizado no modo de *cache*. Neste modo, quando a primeira página é carregado no registrador de dados, ela vai ser transferida do registrador de dados para o registrador de *cache*. Tipicamente, a cópia dos dados do registrador de dados para o registrador de *cache* leva cerca de $3\mu\text{s}$ (nanossegundos). Enquanto os dados são lidos no registrador de *cache* através do barramento de I / O de 8 *bits*, a página seguinte pode ser lida no registrador de dados.

A operação de leitura é feita de forma semelhante a operação de escrita.

Semelhante a um comando de leitura, a escrita ou comando do programa deve ser emitido no nível da página e as páginas dentro de um bloco são escritas em ordem sequencial. Para programar uma página, é emitido um comando de escrita para o registrador de comando, escreve um número de bloco e número da página para o registro de endereços e carrega dados para o registro de dados. O dado é programado na página em $200\mu\text{s}$. Para programar mais de uma página, o comando de gravação podem ser usado no modo de *cache*, que é semelhante ao modo de cache de leitura descrito anteriormente que permite o acesso simultânea via *cache* e registrador de dados (DIRIK; JACOB, 2009)

3.1.2 Comando Apagar

A única maneira de alterar os valores de *bit* de 0 para 1 é apagando. Ao contrário dos comandos de ler e escrever, o comando de apagamento só pode ser executado a nível de bloco.

Uma vez emitido o comando apagar, todos os valores de *bits* em todas as páginas de um bloco são definidos para 1. Para apagar um bloco, um comando apagar é emitido pelo registrador de comando e é carregado um número de bloco para o registrador de endereço. Em seguida, a memória *flash* irá definir seu status para ocupado por 2 ms, enquanto a operação de apagamento é executado e verificada (DIRIK; JACOB, 2009).

3.1.3 FTL (*Flash Translator Layer*)

Um componente crítico chamado Camada Tradução instantânea (FTL), é implementado no controlador SSD para emular um disco rígido e expõe um conjunto de blocos lógicos para os componentes de nível superior. A FTL desempenha um papel fundamental no SSD e muitos mecanismos sofisticados são adotadas para otimizar o desempenho SSD.

A memória *flash* tem suas próprias peculiaridades, para utilizar a memória *flash* como um dispositivo de armazenamento através de *interfaces* padrões de disco rígido é preciso esconder essas peculiaridades. Para esta finalidade SSDs baseado em memória *flash* utilizam a

Camada de Tradução Instantânea (FTL) entre a *interface* lógica e os meios de armazenamento físico. A *Flash Translation Layer* que emula um disco rígido para sistemas *host*, como é relatado em Dirik e Jacob (2009), além do mapeamento de endereço lógico pra físico a FTL também é responsável pela implementação de todos os recursos de memória *flash*. Nas seções seguintes é relatado sobre a maioria deles.

Mapeamento de Bloco Lógico

A escrita em um SSD não é feita da mesma forma que em um disco, cada página gravada é na verdade, conduzida a uma página física diferente. Dessa forma cabe ao mecanismo de mapeamento usado mapear o endereço lógico para o endereço físico, para isso é feito o mapeamento a nível de página ou de bloco, ou ainda um híbrido.

A granularidade do mapeamento pode ser tão pequena como uma página ou tão grande como um bloco. Ainda que o mapeamento a nível de página seja eficiente e flexível é requerido uma grande quantidade de espaço na memória RAM para armazenar a tabela de mapeamento. Ao contrário o mapeamento a nível de bloco embora o espaço-inteligente eficiente requer extensivas operações ler- modificar- escrever quando escreve somente em uma parte do bloco. Muitos FTLs adotam uma abordagem híbrida usando um mapeamento em nível de bloco para gerenciar a maioria dos blocos, como blocos de dados e usando um nível de página mapeamento para gerenciar um pequeno conjunto de blocos de log, que trabalha como um *buffer* para aceitar solicitações de gravação de entrada de forma eficiente. A tabela de mapeamento é mantida persistente na memória *flash* e reconstruída no *buffer* da RAM volátil no momento da inicialização (CHEN et al., 2009).

Jeon et al. (2007) explicam que, no esquema híbrido de mapeamento, a maior parte do espaço de endereço SSD é mapeado em bloco. Estes blocos são referidos como blocos de dados. Um pequeno número de blocos, chamados blocos de log são mapeados na página. O objetivo principal dos blocos de log é lidar com substituições para os blocos de dados. Já que blocos de log normalmente constituem uma pequena porção do espaço de endereço completo, o tabela de mapeamento de página usada por eles é relativamente pequena. Quando a primeira de substituição acontece de uma página em um bloco de dados, um registro bloco está associado a este bloco de dados.

3.1.4 *Garbage Collection* - Coleta de Lixo

Uma página de memória *flash* pode ser gravável ou não-gravável e qualquer página inicialmente é gravável, as páginas graváveis são chamadas de páginas livres. Uma página gravável se torna não-gravável uma vez que está escrita, por isso as páginas não podem ser reescrita antes de apagar. Quando uma porção de dados em um página for modificada, a nova versão dos dados deve ser escrita em uma outra página disponível em algum lugar e a versão antiga é invalidada.

Se o FTL não tem páginas livres o suficiente para atualizar o bloco, ele tem que executar a coleta de lixo (GC) em uma tentativa de recuperar blocos disponíveis para os quais a solicitação de escrita pode ser encaminhada. Este tipo de GC é referida como recuperando atualização de bloco GC. Do mesmo modo, nos casos em que o FTL tem blocos livres insuficientes, ele deve proteger blocos livres de despejar alguns outros blocos lógicos, chamado bloco livre recuperando GC. Esses processos necessitam migrar todas as páginas válidas a partir da atualização do bloco original para um novo bloco livre (chamada migração de página) e apagar esse bloco (JUNG et al., 2012).

Outra função do coletor de lixo é, com o tempo de uso dos SSD, e como resultado da escrita e atualização de dados nos SSDs por diversas vezes, algumas páginas tornam-se inválidas, gerando uma fragmentação da memória ao longo do tempo por toda a memória. Cabe ao *Garbage Collection* selecionar o bloco apropriado, mover todas as páginas válidas para um bloco novo e apagar o bloco selecionado. Desta forma blocos livres se tornam disponíveis para operação de gravação (GOMES, 2012).

Como mostra Jung et al. (2012), a coleta de lixo tem um impacto considerável no desempenho dos SSDs comerciais mais especificamente na latência (tempo de resposta a uma solicitação) e largura de banda, tendo em vista que a operação de GC requer mais custos do que uma operação normal de I/O.

3.1.5 Nivelamento de Desgaste - *Wear Leveling*

Outra característica importante da memória flash NAND é os ciclos de resistência. Um bloco ultrapassará o limite de desgaste após um determinado número ciclos de programar /apagar, que variam de 10.000 a 100.000. O nivelamento de desgaste é um processo que tenta distribuir uniformemente os dados entre as células de memória para garantir que nenhuma célula será excessivamente sobrecarregada, porque quando alguns blocos de memória *flash* forem desgastados, o *chip* de memória todo irá começar a funcionar mal, um bom sistema nivelamento de desgaste deve ser capaz de manter uma distribuição uniforme da contagem do ciclo de apagamento em todos os blocos.

No que tange a longevidade dos SSD, aqueles que utilizam MLC podem regravar em média apenas 10.000 (dez mil) vezes. Para evitar que algumas áreas, com maior carga de trabalho, falhem prematuramente, inutilizando o disco, os fabricantes utilizam sistemas de distribuição de uso (*Wear Leveling*), que alteram os setores usados cada vez que arquivos são criados ou alterados. Sempre que o conteúdo de um bloco é modificado, o controlador simplesmente lê os dados, faz a alteração e grava o conteúdo em um novo bloco. Isso faz com que o bloco anterior seja novamente utilizado apenas depois que todos os outros blocos do SSD forem utilizados pelo menos uma vez o que evita as penalidades de escrita através de seus algoritmos. Com 10.000 operações de escrita por bloco, um SSD de 80 GB teria uma vida útil estimada em 800 terabytes (TB) de gravações, o que é um número difícil de atingir em situações normais, já que você precisaria gravar 40 GB de dados todos os dias por 60 anos para atingir a marca (GOMES, 2012).

Segundo, Deng e Zhou (2010) a coleta de lixo e nivelamento de desgaste têm dois objetivos diferentes que podem entrar em conflito uns com os outros. A coleta de lixo prefere reciclar os blocos que têm um número pequeno de páginas válidas. Pelo contrário, a política de

nivelamento desgaste normalmente recicla os blocos que não são apagadas durante um determinado período de tempo, a fim de eliminar as gravações excessivas para o mesmo local físico da memória. A coleta de lixo pode ser ativada somente quando o número livre de páginas é inferior a um limiar. É sempre um desafio para encontrar um equilíbrio entre a coleta de lixo e nivelamento de desgaste.

3.1.6 Fragmentação de Bloco

Diferente dos HD, o desempenho dos SSD praticamente não é afetado pela fragmentação, já que não faz muita diferença para o controlador se os endereços a serem lidos estão em setores adjacentes ou espalhados pelo disco. Entretanto, particularidades na maneira como os setores são acessados fazem com que em muitas situações o desempenho decaia com o uso, conforme dados são gravados e apagados. À primeira vista, o efeito pode parecer similar à perda de desempenho causada pela fragmentação, mas as razões são completamente diferentes.

Nos SSD, os endereços são organizados em páginas de 4 KB, que são então agrupadas em blocos de 512 KB. As páginas funcionam de maneira muito similar aos *clusters* dos HD magnéticos. Cada página pode armazenar um arquivo ou fragmento de arquivo. Dois arquivos não podem compartilhar a mesma página, o que faz com que arquivos com menos de 4 KB ocupem uma página inteira, desperdiçando espaço. Este não chega a ser um grande problema, já que a maioria dos sistemas de arquivos utilizam *clusters* de 4 KB ou mais de qualquer forma. O grande problema é mesmo a questão dos blocos e das operações de escrita. O grande problema é que as páginas na memória *flash* não podem ser simplesmente regravadas, como os setores em um HD. Sempre que se precisa gravar dados em uma página já ocupada, a controladora primeiro precisa apagar os dados anteriores - levando a célula ao seu estado original - para só então executar a nova operação de escrita. Um fator ainda mais prejudicial é que não é possível apagar as páginas de forma independente. Sempre que se é necessário apagar algo, a controladora precisa limpar todo o bloco (128 páginas) que leva 1 ms ou mais. O processo passa a demorar, então, 1000 vezes mais tempo que a escrita de uma página vazia. Se já houver outros dados válidos dentro do bloco a ser apagado, a operação se torna ainda mais demorada, já que a controladora precisa ler todo o conteúdo do bloco, copiá-lo para o *buffer*, executar o comando de limpeza, fazer a modificação dos dados no *buffer* e em seguida gravar todo o bloco novamente. Essa sequência de passos demora muito tempo, o que explica como o desempenho de muitos SSD pode ser tão baixo em algumas situações (GOMES, 2012).

Para reduzir o número de vezes em que precisa arcar com a penalidade, a controladora procura sempre que possível usar blocos vazios, em vez de regravar o conteúdo de blocos existentes. Nesse caso, um novo bloco é usado para armazenar a versão modificada e a tabela de alocação do sistema de arquivos é atualizada, com a entrada apontando para o novo endereço (GOMES, 2012). No entanto essa não é uma solução definitiva, essa abordagem funciona de forma aceitável no início, mas a partir do momento em que todos os blocos se dão como utilizados pelo menos uma vez, a controladora passa a ser obrigada a voltar e reciclar os blocos anteriormente utilizados, executando todos os passos já citados e isso leva a um fenômeno onde o desempenho do drive novo é consideravelmente melhor que o do drive já utilizado levando a um sintoma similar ao ocorrido em discos rígidos fragmentados, porém com uma causa dife-

rente. A coleta de Lixo e o nivelamento de desgaste citados acima também ajuda a amenizar esse efeito.

3.1.7 TRIM

O comando TRIM é uma nova função que permite ao Sistema Operacional agendar a limpeza das páginas cujo conteúdo foi deletado ou movido em vez de simplesmente marcá-las como vagas. Isso faz com que as páginas livres do SSD realmente estejam prontas para receber novos dados, sem depender unicamente de algum sistema de reciclagem executado pela controladora(GOMES, 2012).

O grande problema com o TRIM é que ele é suportado apenas pelos sistemas mais atuais, o que deixa de fora usuários de sistemas mais antigos. Naturalmente, o TRIM precisa ser suportado também pelo SSD, o que abre outra margem para problemas, já que apenas os *drives* produzidos a partir do final de 2009 oferecem suporte a ele.

3.2 SSD vs HDD

Muito se fala sobre as expectativas a cerca do futuro dos SSDs a sua preferência como sendo o futuro do armazenamento secundário substituindo o HDD. Nessa seção daremos subsídios para que sejam comparados e os posicionamentos que já foram obtidos nas pesquisas a cerca do assunto.

Primeiramente o obvio, os discos rígidos usam pratos para armazenamento dos dados, os SSDs usam *chips* de memórias *Flash NAND*, ou seja, um usa dispositivos eletromecânicos e outro usa só dispositivos elétricos. Como é de esperar partes móveis são mais lentas do que dispositivos semicondutores e também mais sensíveis a falhas e a desgastes. Basicamente essa é a diferença que se mais se destaca entre esses dois dispositivos, mais não é a única.



Figura 5 – Diferença física entre SSDs e HDDs.

Disponível em: www.clubedohardware.com.br

Traduzindo essas informações em números chegamos ao seguinte: Em particular, para o padrão de acesso sequencial os SSDs são até 10 vezes mais rápido e para leituras até cinco vezes mais rápido do que os discos. Para leituras aleatórias, os SSDs oferecem vantagem de desempenho de até 200x. Para gravações aleatórias os SSDs fornecer vantagem de desempenho de até 135x (POLTE et al., 2008).

Tabela 2 – Comparação de característica básica de SSDs MLC e SLC com HDD

	Flash SLC	Flash MLC	HDD
Capacidade	64 GB	256 GB	1T
Leitura	100MB/s (max.)	220MB/s (max.)	8.9 ms (tempo médio de busca) 4.17 ms (latência média)
Escrita	80MB/s (max.)	200MB/s (max.)	175 MB/sec (média de/para o buffer) 300 MB/sec (buffer de/para o host)
Durabilidade	100,000 Ciclos	10,000 Ciclos	N/A
MTBF (Período médio entre falhas)	2,000,000 horas	1,000,000 horas	600,000 POH (porcentagem de horas ligado)

Para entender melhor essas diferenças será apresentado gráficos obtidos através de testes feitos por Rizvi e Chung (2010), ele comparou a velocidade de ler, escrever e copiar

um arquivo no Sistema Operacional *Windows* usando banco de ensaio com um *Flash SSD Samsung 64GB SLC*, *Flash Samsung SSD 256GB MLC* e *HDD Samsung SpinPoint F1 1TB*.

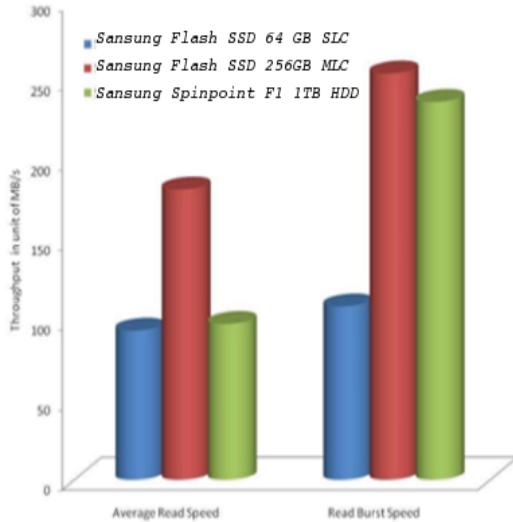


Figura 6 – Criar um arquivo MP3 ler e re-produzir

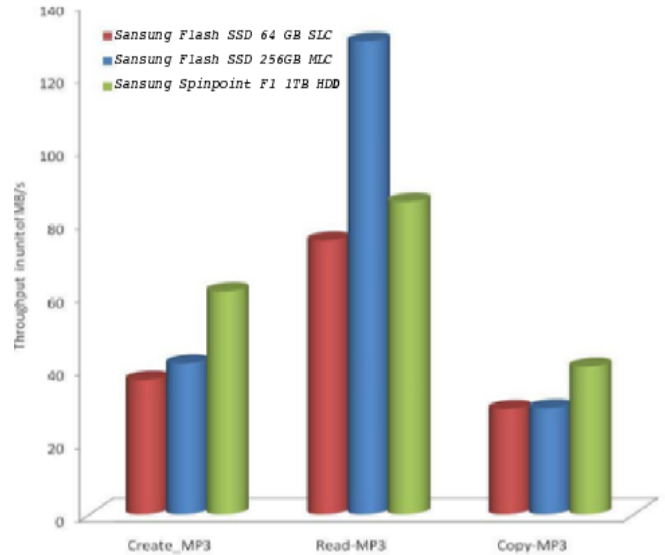


Figura 7 – Média de leitura e exploração de leitura

Durante mais de trinta anos, tem havido um atraso tecnológico na hierarquia de memória entre os tempos de acesso de memórias de acesso aleatório e discos mecânicos (POLTE et al., 2008). O desempenho pequeno de gravação aleatória e o alto consumo de energia é um defeito inerente ao HDD (HUI et al., 2011). Diante disso pesquisadores tem buscado melhorar estes quesitos.

Na área do desempenho, muitos pesquisadores estão interessados nas tecnologias que podem reduzir a distância da cabeça de leitura / gravação, a latência de leitura / escrita e melhorar o desempenho. As tecnologias incluem o seguintes: Projetando algoritmos do elevador para o curto distância de ler ou escrever o movimento da cabeça. A outra é otimizar a operação no nível de sistema de arquivos de acordo com o característico da mídia de armazenamento. A pesquisa métodos utilizados no consumo de energia está a dividir o progresso da operação de HDD em muitos estados diferentes, como ocioso, dormindo, ativo, compensação e assim por diante. A operação de baixo consumo de energia é definida para ser mais longa e a de alto consumo para ser mais curta (HUI et al., 2011).

Memórias *flash* NAND podem fornecer ordens de grandeza e um desempenho mais rápidos do que a mídia tradicional de rotação (HDDs), ainda que ao custo de capacidade reduzida. Empurrando *flash* para densidades mais elevadas, provoca redução significativa em outras métricas - como desempenho, resistência e confiabilidade. Aumentar a capacidade do *Flash*, armazenando um adicional *bit* por célula de memória (1 a 2 bits, ou 2 a 3, por exemplo) reduz a vida útil do *chip*, 5-10%, diminui o rendimento de 22% a 98% (55% em média) e aumenta a latência 1,3 x a 4,0 x (2,3 x em média). O aumento da densidade via escalonamento leva a menor, mas ainda declínios significativos.

3.3 Abordagens Híbridas

Diante da inviabilidade de substituir os HDDs por SSDs, graças ao custo elevado dos SSDs, e baixa densidade e problemas de confiabilidade, encontrou-se um novo mercado, uma maneira de aproveitar os benefícios dos discos de estado sólido sem abrir mão da alta densidade, do baixo custo e da confiabilidade do disco rígido, fazendo uma espécie de arranjo entre os dispositivos. Existem várias propostas de arranjos híbridos com SSDs, nas seções seguintes serão mostradas algumas abordagens já usadas e algumas propostas de uso de SSDs em abordagens que buscam otimizar o funcionamento de sistemas fazendo uso dos benefícios que essa mídia traz.

3.3.1 SSD como cache

A forma mais comum e mais eficiente de combinar SSDs e HDDs é usando SSD como *cache* para HDD essa abordagem não só reduz consumo total de energia, mas também aumenta a capacidade de resposta de um sistema.

Há muitos benefícios em usar SSD como *cache* de um disco para HDD. Com SSD ligados ao sistema, não apenas o tempo de vida do SSD se estende através da diferenciação da direção da leitura e as operações de gravação, ao nível do sistema de desempenho aumenta em executar sequências de inicialização de aplicativos e dados a partir do SSD que tem um alto desempenho de leitura. Em termos de poder de consumo de um sistema, é preferível operar em SSD em vez do disco rígido, porque quando as leituras são direcionadas para o SSD, ele permite ao HDD ficar em estado de *spin-down* mais tempo e economizar cerca de 500mW (LEE et al., 2011)

Essa abordagem de uso de SSDs como *cache* segundo Li et al. (2012) é viável também para motores de busca, sua proposta é usar SSDs como uma *cache* secundária para a memória, o principal objetivo é melhorar o desempenho dos motores de busca, reduzindo o custo de operação dentro de SSD. Os resultados experimentais demonstram a proposta de arquitetura melhora a taxa de acertos em 13,31%, o desempenho em 41,05%, o tempo médio de acesso dentro de SSD em 43,83% e reduz operações de apagamento por bloco em 71,52%.

3.3.2 Combinação de SLC, MLC e TLC em SSDs

Com o objetivo de aumentar a densidade dos SSDs sem aumentar drasticamente os custos, os *chips* TLC, que armazenam 3 bits por célula, estão sendo considerado uma alternativa. No entanto a vida útil e o desempenho dos *chips* TLC são relativamente limitadas inviabilizando um dispositivo feito puramente de módulo TLC.

Diante disso Oh et al. (2013) propõe uma abordagem onde módulos SCL, que são mais confiáveis e duráveis, são usados em conjunto com módulos TLC que tem maior capacidade de armazenamento, baseado em um modelo de análise que propõe um esquema de distribuição

de dados ideal que distribui dados entre os *chips* SLC e TLC para uma determinada carga de trabalho, de tal forma que o desempenho ou a vida pode ser otimizado.

Existe também a possibilidade de combinar células MLC e SLC que é feita da mesma forma do arranjo citado acima, mesclando em um dispositivo *chips* SLC e MLC, aproveitando assim a maior capacidade de armazenamento e menor custo do MLC com a maior durabilidade do SLC (CHANG, 2008).

3.3.3 SSD e PCM

Durante o desligamento, a maioria dos sistemas de acolhimento, inicia um comando (o *STANDBY* Comando imediato) para um SSD para dar ao SSD tempo suficiente para preparar para o desligamento. Isso permite que o SSD salve os dados atualmente em transição (em *buffers* temporários) para a mídia NAND não volátil. No entanto, durante um desligamento de energia insegura, o SSD perde abruptamente alimentação antes de o sistema possa iniciar a *STANDBY* comando imediato. Isto evita que os dados no *buffers* temporários sejam salvos na memória não volátil. Essa abordagem apresenta uma maneira de evitar a perda dos dados caso a alimentação seja interrompida abruptamente.

A proposta de Liu et al. (2011) apresenta uma maneira de preservar os dados de um SSD caso a alimentação seja interrompida. Ele propõe arquiteturas híbridas SSD que são compostas de PCM(*Phase-Change Memory*) e *Flash* NAND. Nestas arquiteturas, perda de dados durante desligamento de energia não planejada pode ser evitado. *Phase-Change Memory* (memória de troca de fase- também conhecido como PCM, PRAM, e PCRAM) é um tipo de memória de computador não-volátil. PCM utiliza o comportamento exclusivo de vidro calcogeneto, que pode ser "ligado"entre dois estados, cristalina e amorfa, com a aplicação de calor. PCM é uma tecnologia de memória emergente com muitas características atraentes: é não-volátil, *byte* endereçável, 2-4X mais denso do que o DRAM, e ordens de magnitude melhor do que NAND *Flash* na latência de leitura, latência de gravação e resistência de gravação.

3.4 Uso de SSDs em Games

Jogos muitas vezes exigem recursos significativos de seus computadores, exigindo processadores rápidos, placas gráficas e *drives* de disco a fim de proporcionar uma experiência de jogo imersiva. Antes de jogar, no entanto, um jogo precisa ser inicializado e iniciado, o que pode levar uma quantidade significativa de tempo para entrar no jogo. Qualquer redução neste momento é cobiçado, especialmente para jogadores que mudam com frequência entre os jogos ou para o jogador casual que pode ter menos interesse em jogar, e portanto, menos tolerância

para atrasos na obtenção para o jogo.

Claypool et al. (2011) publicaram uma experiência do uso de SSDs em jogos e quais vantagens teriam em relação ao HDD. Os resultados foram os seguintes: A análise dos resultados considera SSDs cerca de duas vezes mais rápido que HDDs para leitura, e 15% mais rápido para a escrita. SSDs fornecer Tempos de inicialização do jogo 25% mais rápido do que HDDs, início do jogo 20% mais rápido, mas nenhuma aceleração significativa para jogo poupa tempo. Parente velocidades de acesso de base (leitura e escrita) em um *laptop* em comparação a um *desktop* nem sempre são refletidas no arranque do jogo, inicializar e salvar desempenho com ou sem um SSD.

4 Revisando expectativas acerca dos SSDs

Muito tem se falado sobre os Dispositivos de Estado Sólido e suas muitas vantagens principalmente em relação ao HDD, os SSDs tem sido usados em sistemas de computação e altamente difundidos como uma revolução no sistema de armazenamento. No entanto essa tecnologia ainda é relativamente jovem e maioria das especificações e informações de desempenho são as divulgadas pelos fabricantes e os comportamentos internos dos SSDs têm sido ignorados e podem causar grande impacto na performance do sistema.

Alguns pesquisadores tem refutado a fama dos SSDs e afirmando que as expectativas sobre os SSDs e seu fantástico desempenho são altamente questionáveis. Nesse capítulo serão apresentados os resultados de algumas pesquisas publicadas que testam diversas condições de uso, o impacto delas e a resposta que ela provocam.

A pesquisa de Jung e Kandemir (2013) descobriu por exemplo, que os SSDs não são inclinados para cargas de trabalho de leitura intensiva em termos de desempenho e confiabilidade, especificamente desempenho de leitura aleatória do SSDs é pior do que a seu desempenho de gravação sequencial e aleatória em 40% e 39% em média, e mais importante ainda, o desempenho de leituras sequenciais fica significativamente pior ao longo do tempo. Além disso, descobriu que leituras pode encurtar a vida do SSD mais do que gravações, o que é muito lamentável, dado o fato de que muitos sistemas / plataformas existentes já empregam SSDs como *cache* de leitura ou em aplicações que leituras são altamente intensivas.

Embora muitas dessas aplicações SSD e cenários de uso são propostos e desenvolvidos com base em expectativas comuns de SSDs, SSDs modernos e sistemas de *flash* NAND tem sofrido mudanças tecnológicas e mudanças arquitetônicas graves no último par de anos. Especificamente, as células de memória *flash* NAND encolheram a partir de 5x nm para 2x nm nos últimos quatro anos, e agora menos elétrons são armazenados por porta flutuante. Estas características em nível de célula faz dispositivos *flash* menos confiáveis e introduz operações adicionais para concluir com êxito as solicitações de I/O, que por sua vez impõe latências mais longas. O aumento da capacidade de *chips* de memória *flash* NAND, o tamanho de uma página SSD tem aumentado significativamente ao longo dos últimos anos, por exemplo, a partir de 512B na década de 1990 para 16KB em 2010. Hu et al. (2011) afirma que, quando uma página é atualizada parcialmente (por exemplo, com solicitações de pequenas e aleatórias de I / O em tais cargas de trabalho como MSN), alguns dados devem ser lidos a partir da página antiga e

escritos para uma nova página após serem mesclados com os novos dados, páginas maiores são mais propensas esse problema que leva a degradação do desempenho. Na verdade, ampliando o tamanho da página do *flash* conduz a que o tempo de resposta médio é aumentado para 1,8 vezes, em comparação com a do melhor desempenho *flash* tamanho da página, sob a carga de trabalho do MSN.

4.1 Leitura

Em sua pesquisa Jung e Kandemir (2013) testou diversas vertentes sobre leitura em SSDs, algumas delas são mostradas abaixo: Na pesquisa não foi divulgado o fabricante dos SSDs, pois a finalidade do estudo não era evidenciar ou desfavorecer marca alguma, apenas mostrar o estado da arte dos SSDs. Todos os modelos testados foram fabricados apartir do ano de 2012 com a ultima versão do firmware disponíveis. As características dos dispositivos de SSD usados na pesquisa estão na tabela abaixo.

Tabela 3 – Características Básicas do SSDs usados na pesquisa

	SSD1	SSD2	SSD3	SSD4	SSD5	SSD6
Capacidade(GB)	120	256	256	256	240	240
DRAM(MB)	128	256	0	256	0	0
Número de <i>chips</i>	16	16	8	8	16	16
Tecnologia(nm)	34	25	25-32	32	25	25
Número de núcleos)	2	3	1	3	1	1

1- Leitura aleatória - O desempenho com acessos de leitura aleatórios são piores que outros tipos de padrões de acesso e operações, incluindo mesmo acessos de gravação aleatória, o que está em contraste direto com a expectativa amplamente divulgada sobre desempenho de leitura de SSDs na literatura. Especificamente, os valores de banda com leitura aleatória dos 3 modelos testados são piores do que o valores correspondentes com gravações aleatórias por 59,7%, 39,4% e 23,7%. Características latência de leitura não são muito diferente da largura de banda; os valores de latência observados com leitura aleatória são piores do que as latências observadas com gravações sequenciais, gravações aleatórias, e leituras sequenciais 41,3%, 35,2% e 35,9%, em média.

2- Desempenho de leitura com acessos sequenciais -Este teste foi realizado com SSDs novos e usados, o que se percebeu foi que a maior parte das solicitações de leitura em todos os SSDs novos são atendidas dentro de 300 a 400 ms. Os SSDs usados levam mais de 600 ms para atender os pedidos de I / O, o que é duas a três vezes pior do que os SSDs

novos. Pode-se concluir, a partir desta análise que, características de desempenho de leitura sequencial pioram com o envelhecimento.

3 -Funcionamento do ciclo programar/apagar (PE) e o impacto sobre a leitura - Nos dispositivos testados Pode-se ver a partir desta experiência que os ciclos PE aumentam em cada rodada de avaliação, de forma direta contrasta com o que a literatura atual sobre sistemas que exploram SSDs levaria a crer. Em leituras sequenciais, o ciclos máximas PE atingem a metade dos ciclos de PE nas gravações. Ironicamente, os ciclos máximas de PE com a carga de somente leitura aleatória são maiores do que a de escrita em cerca de 12x. Além disso, a disparidade entre o máximo e a média de ciclos de PE nos conta outra história; estratégias de nivelamento de desgaste do *firmware* de *flash* atual centram-se principalmente nas gravações, não em leituras. Enquanto o *firmware* mantém a redução das disparidades entre a média máxima de ciclos de PE em gravações, os ciclos máximos PE sobre leituras são 247 vezes maiores do que os ciclos médios de PE em cada ciclo, o que faz com que certos blocos desgastem mais rapidamente e piores características de resistência do SSD.

4 -Impacto do gerenciamento de confiabilidade na performance de leituras - Todas as atividades para lidar com distúrbios de gestão de leitura, má gestão tempo de execução do bloco, e ECC, a que se refere coletivamente como gerenciamento de confiabilidade nas leituras (RMR) exigem operações adicionais de E / S e ciclos de computação. Estes custos não são revelados aos usuários, mas pode contribuir para longa latências em operações normais. Para nossa avaliação, nós executamos uma carga de trabalho aleatória de somente leitura, em três dispositivos de SSD. A latência de leitura com RMR é pelo menos cinco vezes maior que a latência de leituras sem RMR, o que seria inaceitável para aplicações SSD sensíveis à latência. Além disso, os custos gerais RMR são mais consideráveis com padrões de acesso aleatório a pequenos volumes (1 até 64 setores), que é o tamanho dominante pedido em muitos sistemas de arquivos. Considerando 8 setores (4KB) como exemplo, enquanto que as latências de leitura sem RMR sobre os dispositivos testados são de 75, 60 e 47 microssegundos, respectivamente, o aumento da latência com RMR são 685, 1787, e 4944 microssegundos, na mesma ordem.

4.2 Gravação

Além dos testes relacionados a leitura Jung e Kandemir (2013) fez testes de gravação, alguns resultados estão descritos abaixo:

1-Impacto da latência do pior caso em SSDs - Comparamos todos os SSDs testados e dois

tipos de HDDs (7K RPM e 10K RPM) tanto em latência média quanto em latência do pior caso. Para quantificar o latência média, usamos dispositivos como novo, e para o pior caso a avaliação de latência, que empregam os dispositivos totalmente utilizado para SSD e HDD. A latência média de todos os SSDs são melhores do que os HDDs, independente do tipo de carga de trabalho utilizada. No entanto, o pior caso em latências de SSDs totalmente utilizados são muito piores do que a dos discos rígidos, que é um problema para muitas aplicações SSD de intensa gravação. Especificamente, as latências de pior caso de todos os SSDs testados são, em média, 12 a 17 vezes pior do que a dos HDD testados.

2-Buffer de DRAM e redução dos custos da Coleta de lixo(GC) - Uma vez que as gravações podem se armazenadas nos buffers (Memória de armazenamento temporário onde os dados ficam antes de serem gravados definitivamente na memória) de DRAM, em SSDs modernos é esperado que se esconda os custos com Coleta de Lixo. Experimentos feitos com cache de DRAM ativa e desativada em dois modelos SSDs diferentes deram os seguintes resultados: O pior caso de latência em um dos dispositivos testados é escondido pelo buffer DRAM antes do precipício gravação começar. No entanto, uma vez que GCs começam a ser chamadas em série, a latência do segundo modelo de SSD com *cache* ativa torna-se duas vezes pior do que com *cache* inativa no outro dispositivo. No primeiro dispositivo testado, a disparidade de desempenho entre *cache* ativa e *cache* inativa são mais acentuados.

A coleta de lixo (GC) - Recurso que visa recuperar blocos invalidados (gravados) para torna-los novamente graváveis- introduz latências significativas e grandes variações de desempenho, o que torna a adoção generalizada de SSDs difícil. Jung et al. (2012) testou as latências acrescentadas pela GC, podemos ver os resultados nos gráficos a seguir:

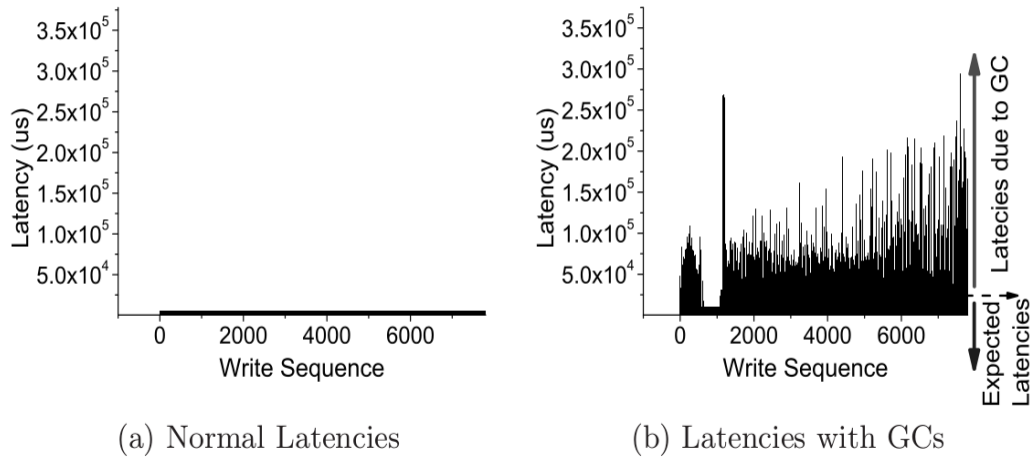


Figura 8 – Latência

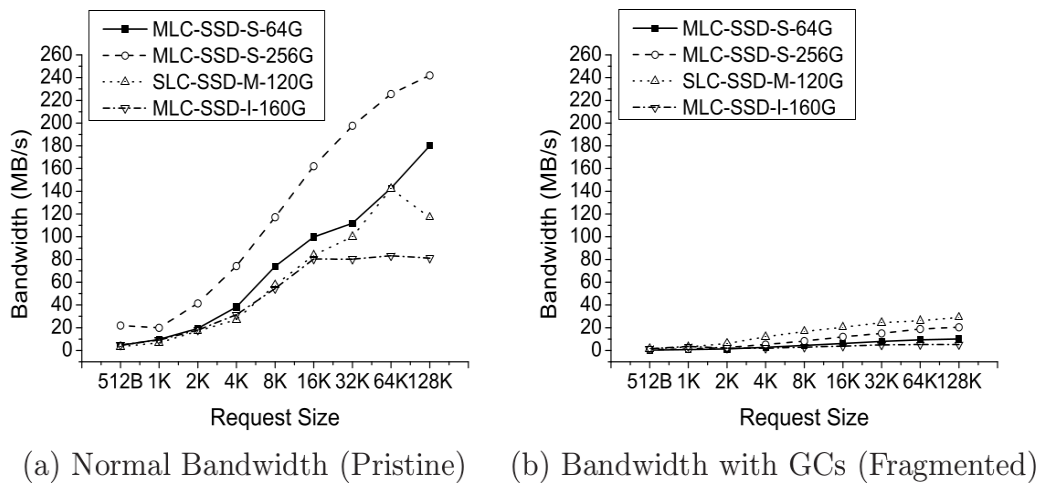


Figura 9 – Largura de banda

Na figura 8 podemos ver a comparação de latência para um padrão de acesso de gravação aleatória com um pedido com tamanho de 1MB usando um SSD real baseada no MLC. Com

isso podemos verificar a latência adicionada pela coleta de lixo.

A figura 9 mostra o rendimento do sistema para o estado da arte de quatro SSDs (diferentes fornecedores e tipos de NAND). Note que todos os SSDs testados sofrem degradação significativa no desempenho, uma vez iniciada as coletas de lixo.

4.3 Comando TRIM

De acordo com os testes, a magnitude dos ganhos de desempenho com os comandos TRIM (O comando TRIM é uma função que permite ao sistema operacional agendar a limpeza das páginas cujo conteúdo foi deletado ou movido em vez de simplesmente marcá-las como vagas) varia significativamente, dependendo do modelo de solicitação TRIM. Enquanto SEQ-TRIM (sequencial) pode efetivamente eliminar custos gerais de Coleta de Lixo, RND-TRIM (aleatório) não tem impacto positivo sobre a performance. Além disso, os SSDs exigem tempos de execução muito longo para processar o TRIM, que pode levar a inesperada degradação de performance (JUNG; KANDEMIR, 2013).

Diante dos dados obtidos através dos testes citados acima, podemos concluir que a leitura aleatória em SSDs em termos de largura de banda e de latência, ao contrário do que se pensava, é pior que qualquer outra operação, inclusive a escrita. Percebeu-se também que a leitura sequencial se degrada com o tempo. No que se refere a escrita o que se concluiu é que o rendimento de gravações cai significativamente e o pior caso latência aumenta acentuadamente quando o penhasco gravação é atingido. Além disso, o pior caso de latência de SSD é muito mais elevada do que HDDs, o que implica que ele precisa ter muito mais atenção especialmente no contexto de aplicações sensíveis à latência.

5 Resultados

Este capítulo apresenta os resultados obtidos nessa pesquisa, as fontes de materiais usados para se chegar aos resultados e algumas considerações referentes aos SSDs.

5.1 Bases de Pesquisa

Para realização dessa pesquisa foram consultadas as bases de dados ACM, IEEE Xplore e Periódicos Capes, além de outras publicações relacionadas. Nas buscas realizadas nos três repositórios de dados pelo termo genérico *Solid State Disk* retornaram 4.200, 1002 e 3.311 resultados na ACM, IEE Xplore e Periódicos Capes respectivamente. Nos resultados retornados pela busca no repositório da ACM, cerca de 30% deles tinham sido publicados pela IEEE, e conseqüentemente foram retornados na busca da IEEE Xplore também. Os Periódicos Capes, disponibilizam publicações, tanto da ACM quanto da IEEE, o que fez com que uma média de 25% dos resultados retornados na busca fossem retornados também no seu repositório de origem, fazendo com que aproximadamente 55% dos resultados encontrados fossem repetidos. Das 3830 publicações retornadas, muitas eram bastante antigas e, portanto, fora de contexto, pois as buscas foram de publicações feitas desde 1961, outras não tinham relevância para o foco desse trabalho e ainda outras tratavam de assuntos iguais ou semelhantes. Os resultados apresentados acima foram retornados na busca por um termo genérico, na busca por termos específicos como por exemplo, Coleta de Lixo, Nivelamento de Desgaste, Comando TRIM, SSDs e comandos avançados e entre outros, o número de resultados retornados foi em média 90% menor. Depois de filtrar os resultados mais relevantes, as publicações utilizadas nessa pesquisa foram 94.

No gráfico abaixo vemos a relação publicação/base de dados do material que foi utilizado na pesquisa.

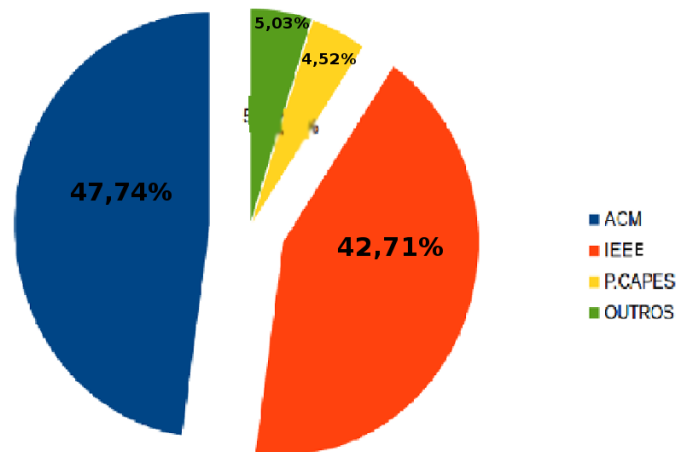


Figura 10 – Relação Base de dados/ Publicações usadas no estudo.

Das publicações a acima citadas 23 foram citadas diretamente neste trabalho, as demais deram suporte indireto a pesquisa.

Analisando publicações feitas baseadas em testes pode-se perceber que, conceitualmente os SSDs se mostram muito atraentes e parecem o meio perfeito de se conseguir alta performance e desempenho com menor consumo de energia e mais resistência já que há muito tempo os HDDs vem assumindo o posto de memória secundária e as evoluções significativas foram poucas ao longo dos últimos anos. No entanto esses resultados fantásticos são obtidos em testes genéricos, com dispositivos novos e condições ideais. A realidade é que muitos fatores degradam o desempenho e comprometem a vida útil dos SSDs, o que nos faz questionar o seu uso generalizado.

A maioria dos autores publicam estudos baseados em dados fornecidos pelos fabricantes ou em testes comparativos com dispositivos de Estado Sólido novos. O que resulta em dados de desempenho perfeito. E para os transtornos trazidos pelas características peculiares dos SSDs de gravar e apagar apresentam a coleta de lixo e o comando TRIM. Mais o fato é que esses comandos também degradam o desempenho do disco, a coleta de lixo adiciona latências muito grandes as operações e, com isso compromete significativamente a largura de banda, o que na prática, reduz muito as vantagens dos SSDs sobre os HDDs.

Outro aspecto verificado foi o tempo de vida do SSD que, principalmente em aplicações que precisam de muito escrita, pois os blocos precisarão serem limpos com muita frequência isso degrada o desempenho e reduz o tempo de vida do SSD.

Apesar das controvérsias, as memórias *Flash NAND* e os SSDs são uma grande pro-

messa que aos poucos se torna realidade. No que se refere a dispositivos de armazenamento portáteis esse tipo memória domina, nossos *Pen Drives* por exemplo, são memórias *Flash NAND* assim como cartões de memórias, muitos computadores adotam discos híbridos e SSDs para inicializar rapidamente o sistema. Pude ver na construção desse trabalho algumas propostas de arquitetura e protótipos de SSDs com ideias de melhoramentos, o que me faz acreditar no futuro dos Discos de Estados Sólidos e nas vantagens que ele pode oferecer.

5.2 Trabalhos Futuros

Como pode-se perceber, não há muitas publicações com testes mais profundos, em que são testados SSDs sob perspectivas mais reais, as especificações divulgadas por fabricantes são incompletas pois são consideradas diferenciais competitivos e, portanto, são preservadas. Cabe as pesquisas revelar características e comportamentos mais aprofundados sobre os SSDs.

Fica como sugestão para trabalhos futuros a realização de testes exaustivos tanto em SSDs como em HDDs avaliando por exemplo, latência para leitura e escrita, funcionamento dos dispositivo com uso contínuo sob diversas cargas de trabalho, confiabilidade e falhas.

6 Conclusão

Esse trabalho procura disponibilizar, fazendo um apanhado geral do cenário de dispositivos de armazenamento secundário, bases para conhecer melhor os SSDs. A partir desse trabalho pode-se conhecer melhor a respeito dos SSDs e seu funcionamento em comparação com os HDDs, e compreender através de testes o desempenho dos SSDs e as implicações no desempenho.

Enfim, os Dispositivos de Estados Sólido são uma tecnologia promissora, os pesquisadores acreditam no futuro desse dispositivo como dominante no cenário de armazenamento secundário e demais aplicações que possa exercer. Porém é uma tecnologia jovem que precisa ser amadurecida antes de ocupar o lugar dos tradicionais HDDs, pois em um mundo cada vez mais tecnológico em que a automatização de transações, das mais triviais às mais excepcionais, estão cada vez mais comuns, e que o equilíbrio do mundo real é refém do virtual, é preciso que os passos dados sejam bem calculados.

A partir desse trabalho teremos subsídios para avaliar e adequar as escolhas às necessidades, pois os SSDs tem outros usos além de armazenamento secundário, já vivemos há anos usando memórias *Flash* (Base da composição de SSDs) em nossos *Pen Drives* e cartões de memória, e boa parte dos computadores portáteis atualmente saem de fábrica com uma memória SSD para *Boot*. Muito provavelmente os SSDs não sairão de cena, como também os HDDs, pelo menos pelos próximos anos. Por isso a função principal desse trabalho é acima de tudo fornecer bases para que se tenha a consciência do real cenário que envolve SSDs, HDDs e questões referentes a substituição.

Referências

- CARNEIRO, José Luís. **Histórico e Hardware**. Salvador, Brasil, 2006. Disponível em: <<http://www.jlcarneiro.com/wp-content/uploads/File/inf/hardware.pdf>>. Acesso em: 20 jun. 2014.
- CHANG, Li-Pin. **Hybrid Solid-State Disk: Combining heterogeneous NAND Flash in Large SSDs**. *IEEE*, Hsin-Chu, Taiwan, jul. 2008.
- CHEN, Feng; KOUFATY, David A.; ZHANG, Xiaodong. **Understanding Intrinsic Characteristics and System Implications of Flash Memory based Solid State Drives**. *ACM*, Columbus, OH 43210, jun. 2009.
- CLAYPOOL, Mark et al. **On the Performance of Games using Solid State Drives**. *IEEE*, Worcester, MA 01609, USA, 2011.
- DENG, Yuhui; ZHOU, Jipeng. **Architectures and optimization methods of flash memory based storage systems**. *Journal of Systems Architecture*, Elsevier B.V., Guangzhou 510632, PR China, dez. 2010.
- DIRIK, Cagdas; JACOB, Bruce. **The Performance of PC Solid-State Disks (SSDs) as a Function of Bandwidth, Concurrency, Device Architecture, and System Organization**. In: . Austin, Texas, USA: [s.n.], 2009. Acesso em: 20 Nov. 2013.
- GOMES, Jeremias Moreira. **A forense computacional e os discos de estado sólido**. *ICoFCS*, Brasília, Brasil, 2012.
- HU, Yang; JIANG, Hong; FENG, Dan. **Performance Impact and Interplay of SSD Parallelism through Advanced Commands, Allocation Strategy and Data Granularity**. Wuhan, China, 430074, jun. 2011. Acesso em: 11 Jul. 2013.
- HUI, Sun; RUI, Zeng; JIN, Cheng. **Analysis of file system and block I/O scheduler for SSD in performance and energy consumption**. *IEEE Asia -Pacific Services Computing Conference*, 1037 Luoyu Road, Wuhan, China, 2011.
- JEON, Hyeran; MAGHRAOUI, Kaoutar El; KANDIRAJU, Gokul B. **Investigating Hybrid SSD FTL Schemes for Hadoop Workloads**. Los Angeles, USA, 2007.
- JOSE, Soya Treasa; PRADEEP.C, Prof. **Design of a Multichannel NAND Flash Memory Controller for Efficient Utilization of Bandwidth in SSDs**. In: . Kerala, India: [s.n.], 2013.
- JUNG, Myoungsoo; KANDEMIR, Mahmut. **Revisiting Widely Held SSD Expectations and Rethinking System-Level Implications**. *ACM*, University Park, PA 16802, USA, 2013.
- JUNG, Myoungsoo; PRABHAKAR, Ramya; KANDEMIR, Mahmut Taylan. **Taking Garbage Collection Overheads Off the Critical Path in SSDs**. *International Federation for Information Processing*, Middleware, 2012.

LEE, Seongjin et al. **Performance Analysis of SSD HDD Hybrid Storage Manager**. Seoul, Korea, 2011.

LI, Ruixuan et al. **An Efficient SSD-based Hybrid Storage Architecture for Large-scale Search Engines**. *International Conference on Parallel Processing*, IEEE, Wuhan, P.R.China, 2012.

LIU, Yining; ZHOU, Chuangshi; CHENG, Xiaohua. **Hybrid SSD with PCM**. *IEEE*, Shenzhen, Guangdong 518057, P. R. China, 2011.

OH, Yongseok et al. **Hybrid Solid State Drives for Improved Performance and Enhanced Lifetime**. Korea, jul. 2013.

OLIVEIRA, Vinícius Garcia de; CORREA, Ricardo. **Tecnologias de Discos Magnéticos**. Compinas. São Paulo, brasil, 2011.

POLTE, Milo; SIMSA, Jiri; GIBSON, Garth. **Comparing Performance of Solid State Devices and Mechanical Disks**. *IEEE TRANSACTIONS ON MAGNETICS*, Pittsburgh, PA, 15213, 2008.

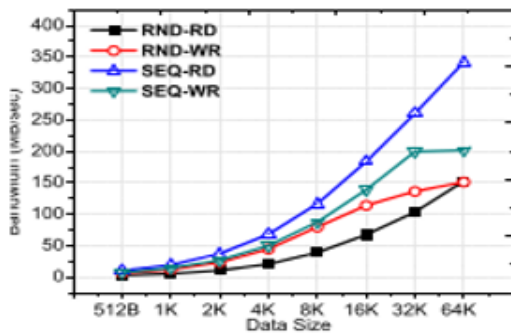
RIZVI, Sanam Shahla; CHUNG, Tae-Sun. **Flash SSD vs HDD: High Performance Oriented Modern Embedded and Multimedia Storage Systems**. *IEEE*, Korea, 2010.

SHIRATORI, Felipe H.; YAMASHITA, Hugo H.; POZZATO, Guilherme H. G. **Por dentro do HD**. Compinas. São Paulo, brasil, 2012.

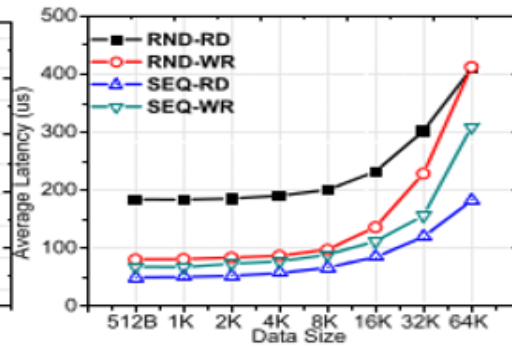
TAMPELINE, Leonardo Garcia. **Tecnologias de Discos: ATA, SATA, SCSI, SAS e Fusion-io**. Compinas. São Paulo, brasil, 2011.

GROCHOWSKI, Edward G.; HOYT, Roger F. **Magnetic Hard Disk Drive Form Factor Evolution**. *IEEE TRANSACTIONS ON MAGNETICS*, 650 Harry Road, San Jose, CA, v. 29, n. 6, nov. 1993.

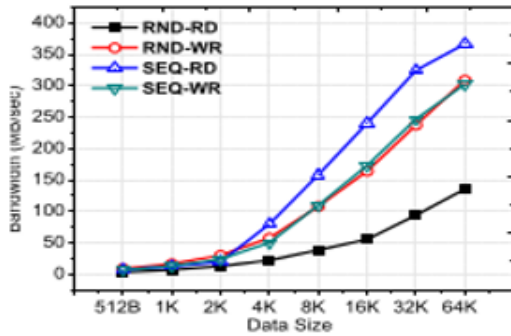
ANEXO A – Comparação de desempenho de leitura/escrita sequencial e aleatória



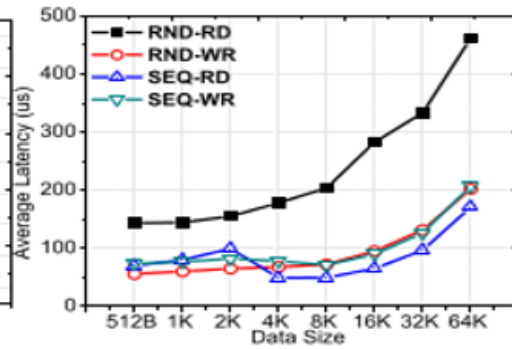
(a) SSD-L Bandwidth.



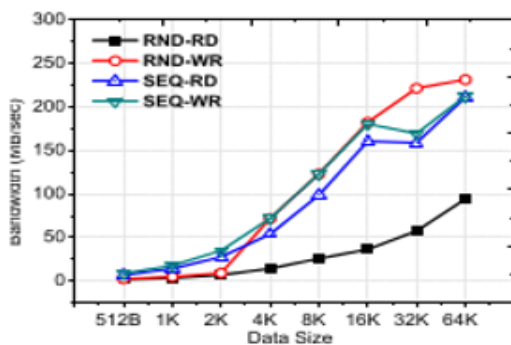
(b) SSD-L Latency.



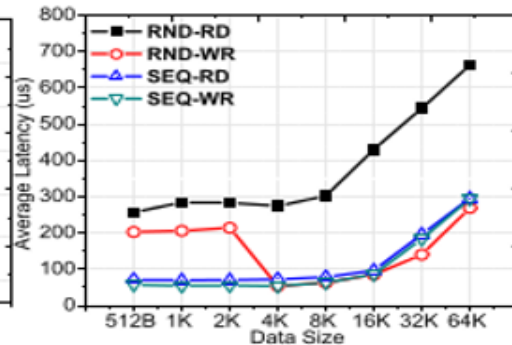
(c) SSD-C Bandwidth.



(d) SSD-C Latency.



(e) SSD-Z Bandwidth.



(f) SSD-Z Latency.

ANEXO B – SSD e HDD- Comparação de latência

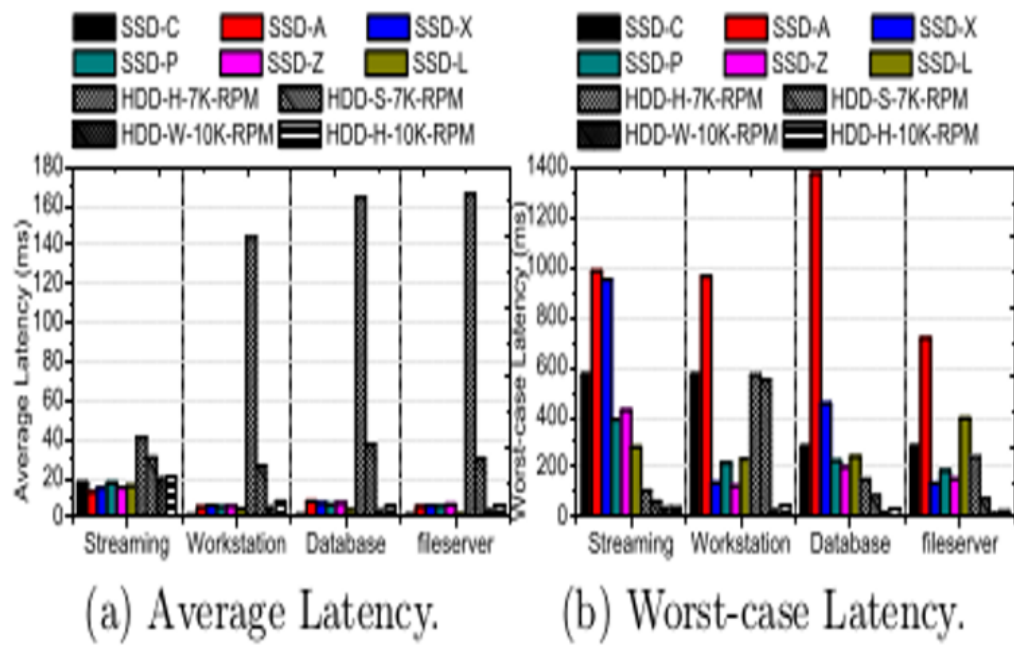


Figura 11 – Melhor caso e pior caso respectivamente