

# Benchmark do sistema de arquivos.

Lucas Cauã Da Silva De Lima, Leonardo Yokemura Marques

Computação – Centro de Desenvolvimento Tecnológico (CDTec) – Universidade  
Federal de Pelotas (UFPEL)  
CEP 96010-610 – Pelotas – RS – Brazil

{lcdsdlima,lymarques}@inf.ufpel.edu.br

**Abstract.** *This paper will present results and comparisons obtained through a benchmark of the file system will be addressed methodologies used, information about the benchmark, results, conclusions drawn from these results, types of tests among other topics.*

**Resumo.** *Neste artigo serão apresentados resultados e comparações obtidos através de um benchmark do sistema de arquivos, serão abordados metodologias utilizadas, informações sobre o benchmark, resultados obtidos, conclusões obtidas a partir desses resultados , tipos de testes entre outros tópicos.*

## 1. Introdução

Presenciamos a cada dia o desenvolvimento em ritmo cada vez mais acelerado de componentes de hardware, processadores, memórias ram, Gpu's, e consequentemente discos de armazenamento.

Com componentes cada vez mais rápidos é necessário que os discos de armazenamento sejam capazes de suportar toda a demanda de dados gerada por estes, nesse ponto temos o surgimento dos SSD's que aparecem como uma alternativa mais rápida e adaptada as demandas atuais do que os HD's, utilizando memórias flash os SSD's fornecem taxas de leitura e escrita superiores aos dos HD's. No entanto um fator que dificulta a popularização dos SSD's são seu alto custo, por ser produzido com memórias de alta velocidade, ao invés dos discos magnéticos utilizados nos HD's, os SSD's possuem alto custo de produção, o que dificulta a disputa de mercado contra os HD's.

Portanto podemos assumir que os HD's permanecerão por um bom tempo em alta, oferecendo um bom custo benefício na relação valor/quantidade de armazenamento.

Outro ponto importante quando falamos da utilização de discos rígidos é referente ao sistema de arquivos utilizado por estes, já que a forma como os arquivos são armazenados e operados é de extrema importância para determinar um melhor ou pior desempenho. Este é o principal objetivo deste artigo, testar e comparar diferentes tipos de sistemas de arquivos obtendo algumas conclusões sobre o desempenho destes.

## **2. Benchmark desenvolvido**

Para a realização dos testes foi desenvolvido um benchmark para sistemas de arquivos utilizando a linguagem C.

As fases de desenvolvimento, e o código final estão disponíveis no github através do link: <https://github.com/lucaslima94/Benchmark>.

### **2.1. Funcionamento do Benchmark**

O Benchmark desenvolvido possui as opções de tamanho do arquivo e número de repetições, permitindo dessa forma que o usuário realize a execução de forma apropriada ao hardware disponível, a cada execução parcial são geradas as taxas de escrita sequencial, leitura sequencial, escrita aleatória e leitura aleatória em MB/s, além do tempo de execução parcial em segundos.

Ao fim da execução de todas as repetições definidas pelo usuário são geradas as médias das taxas de escrita sequencial, leitura sequencial, escrita aleatória e leitura aleatória em MB/s, o desvio padrão das taxas de escrita sequencial, leitura sequencial, escrita aleatória e leitura aleatória em MB/s e o tempo de execução total em segundos.

A escrita e leitura aleatória são realizadas da seguinte forma: a quantidade de dados a ser lida e dividida em 1000 partes onde cada parte representa um ponto no arquivo, ponto este gerado aleatoriamente, onde serão escritos os dados.

### **2.2. Operações realizadas**

As operações realizadas pelo benchmark são escrita sequencial, escrita aleatória, leitura sequencial e leitura aleatória da seguinte forma:

Inicialmente é criado um arquivo de tamanho definido pelo usuário através da escrita sequencial, logo após é lida uma parte de 1/10 do tamanho do arquivo de forma sequencial, e outro 1/10 do arquivo de forma aleatória, por fim o arquivo original é aberto e são escritos 1/10 do tamanho do arquivo de forma aleatória. Sendo gerados ao fim da execução os dados citados na subseção Funcionamento do Benchmark.

## **3. Estado da arte para benchmark em sistemas de arquivo**

O evolução cada vez mais acelerada de memórias e processadores em relação a pequena evolução dos HD's faz com que cada vez mais o acesso ao disco secundário se torne o principal gargalo em um sistema. Embora varias técnicas sejam utilizadas a fim de diminuir o numero de acessos ao HD, como caches, aumento da capacidade das memórias ram, uso de swap, ainda não foi possível fazer com que os acessos ao disco rígido deixem de ser um gargalo.

Com o intuito de melhorar a performance durante o acesso aos discos secundários o desenvolvimento de sistemas de arquivos mais eficientes se tornou uma área de pesquisa promissora, consequentemente os benchmarks de sistema de arquivos ganharam um foco especial, a fim de definir a real eficiência de um sistema de arquivos os benchmarks tendem a ser cada vez mais complexos e eficientes.

Outro ponto que tem alavancado as pesquisas na área de sistemas de arquivos

diz respeito aos sistemas de arquivos distribuídos, com dados cada vez mais complexos e que necessitam de muito processamento se populariza o uso de computação distribuída, e consequentemente a criação e aperfeiçoamento de sistemas de arquivos específicos para esta área, ocasionando o aumento de pesquisas de benchmarks para estas áreas.

#### **4. Metodologias de Avaliação**

Para realização dos testes utilizou-se uma máquina com os seguintes componentes de hardware: Processador Intel Pentium 4 (3.0Ghz), Memória ram 1.5 Gb DDR2 533 mhz, Disco Rígido Maxtor 80GB 7200 Rpm, Unidade de armazenamento flash USB Sandisk cruzer blade 8.0 GB conectado a uma porta USB 2.0;

Os testes foram realizados com as unidades de armazenamento da seguinte forma: Disco rígido Maxtor 80GB 7200 Rpm : uma partição em formato ext4 e uma partição em formato NTFS, Unidade de armazenamento Flash USB Sandisk cruzer blade 8GB, uma partição em formato ext4.

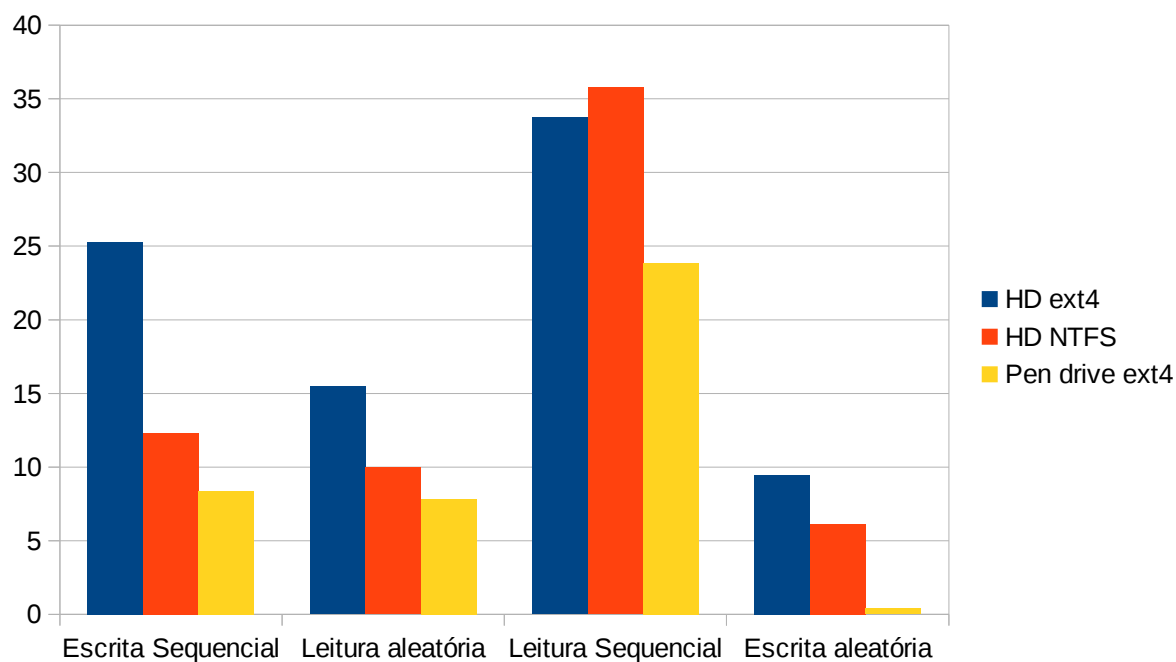
Os testes foram realizados utilizando arquivos de tamanho 2.0 GB, a fim de evitar armazenamento do arquivo em cache da memória ram. Foram realizadas 10 repetições do teste em cada Unidade para garantir um resultado mais fiel a realidade.

Os resultados podem ser vistos na seção resultados.

#### **5. Resultados**

Veremos agora os resultados obtidos após 10 repetições do benchmark, utilizando tamanho de arquivos de 2 GB, logo após serão feitas comparações entre os diferentes discos e diferentes partições.

Inicialmente vemos os resultados gerais obtidos na forma de gráfico (Gráfico 1) e duas tabelas, que mostram as médias das taxas de transferência de dados obtidas durante a execução (Tabela 1) e os desvios padrão da execução (Tabela 2).



**Gráfico 1** – Média em MB/s de transferência para as operações. .

	HD ext4	HD NTFS	Pen Drive ext4
Escrita sequencial	25,30 MB/s	12,33 MB/s	8,36 MB/s
Leitura aleatória	15,46 MB/s	9,98 MB/s	7,83 MB/s
Leitura sequencial	33,73 MB/s	35,77 MB/s	23,85 MB/s
Escrita aleatória	9,48 MB/s	6,14 MB/s	0,40 MB/s

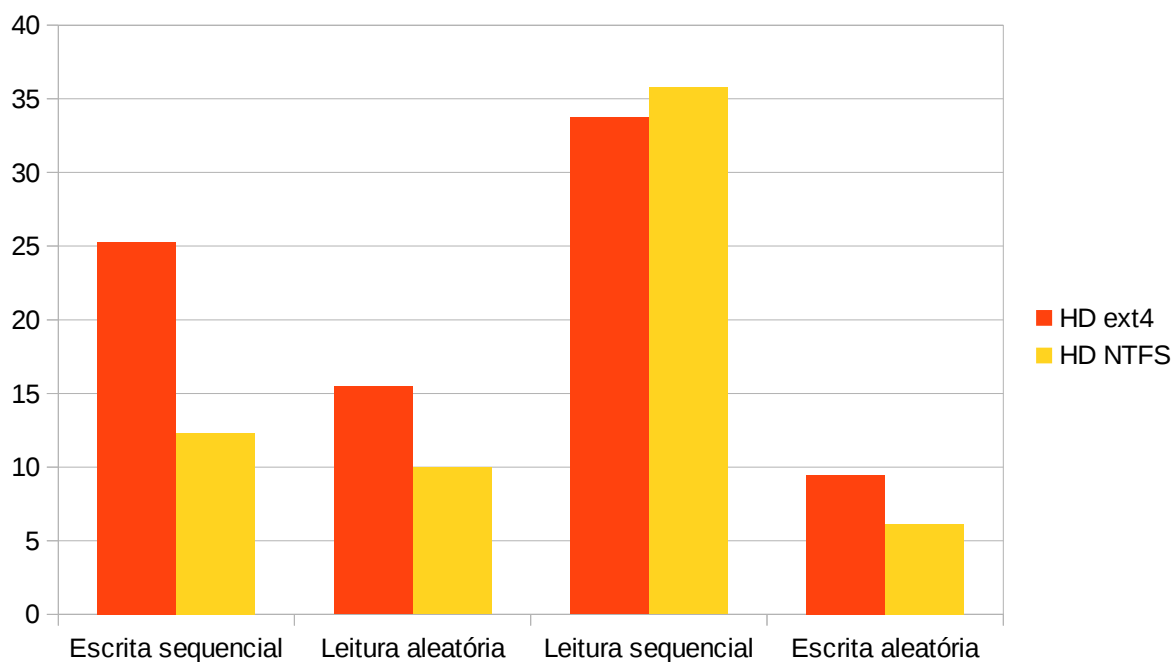
**Tabela 1** – Média em MB/s de transferência para as operações.

	HD ext4	HD NTFS	Pen Drive ext4
Escrita sequencial	0,49 MB/s	0,02 MB/s	0,30 MB/s
Leitura aleatória	1,02 MB/s	0,79 MB/s	0,33 MB/s
Leitura sequencial	4,29 MB/s	2,48 MB/s	0,29 MB/s
Escrita aleatória	0,45 MB/s	0,28 MB/s	0,004 MB/s

**Tabela 2** – Desvio padrão em MB/s das operações.

### 5.1. Comparação de resultados em diferentes formatos de partição do HD.

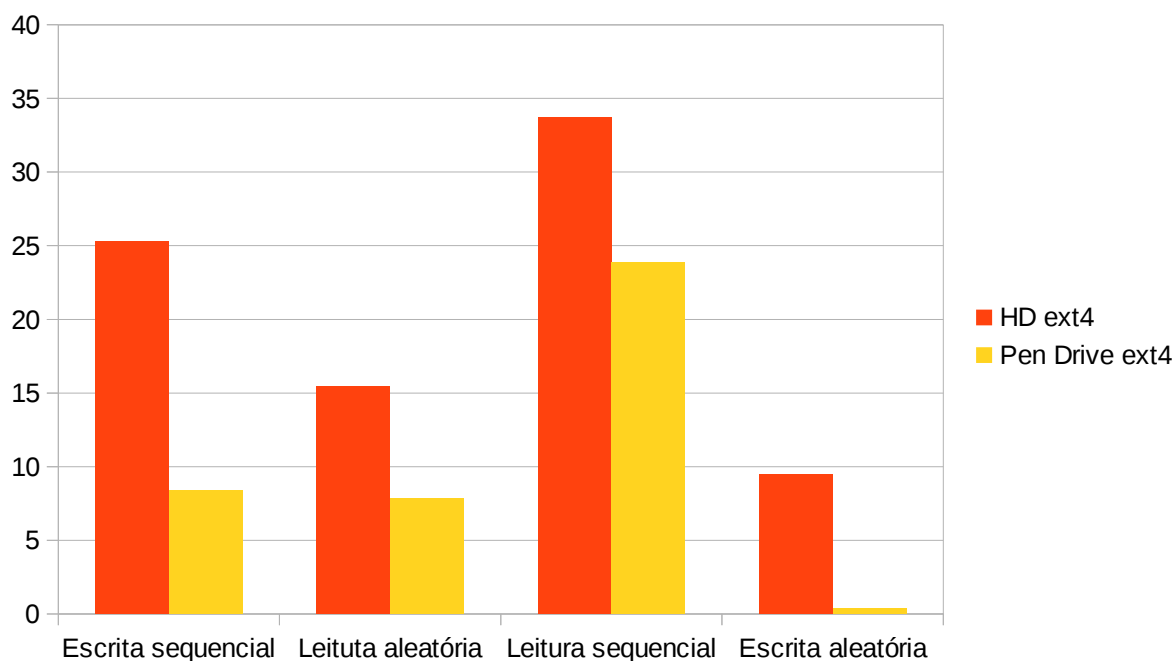
No gráfico 2 é apresentada uma comparação entre as partições em formato NTFS e ext4 em um mesmo HD.



**Gráfico 2** – Comparação entre partições NTFS e ext4.

## 5.2. Comparação de resultados de um mesmo formato de partição em diferentes discos.

No gráfico 3 é apresentada uma comparação entre diferentes discos utilizando o formato ext4.



**Gráfico 3** – Comparação de diferentes disco em formato ext4.

## **6. Conclusões**

Comparando os resultados apresentados é possível notar que quando comparamos um mesmo HD com partições ext4 e NTFS, temos melhores valores para a partição ext4 em 3 quesitos: escrita sequencial, leitura aleatória e escrita sequencial, já na leitura sequencial a vantagem é para a partição NTFS.

Quando comparamos dois discos, Disco rígido e flash drive USB, com o mesmo formato de partição ext4, notamos que o disco rígido leva vantagem em todos os quesitos, no entanto a medida de um flash drive USB possui outras variáveis a serem consideradas, principalmente o limite de transmissão de uma porta USB 2.0.

## **Referencias**

Isaila, Florin. File Systems: State of the Art. Disponível em [http://www.di.unipi.it/~coppola/dags/florin\\_abstract/abstract.html](http://www.di.unipi.it/~coppola/dags/florin_abstract/abstract.html) . Acessado em 20-jun-2015.

IOZone File System Benchmark. Disponível em : <<http://www.iozone.org/>>. Acessado em 20-Jun-2015.

Gabriel Torres, Sistemas de arquivos. Disponível em <<http://www.clubedohardware.com.br/artigos/sistema-de-arquivos/313>>. Acessado em 20-Jun-2015.