

Sistema de Arquivos e Estrutura de Diretórios

Por Silvio Ferreira

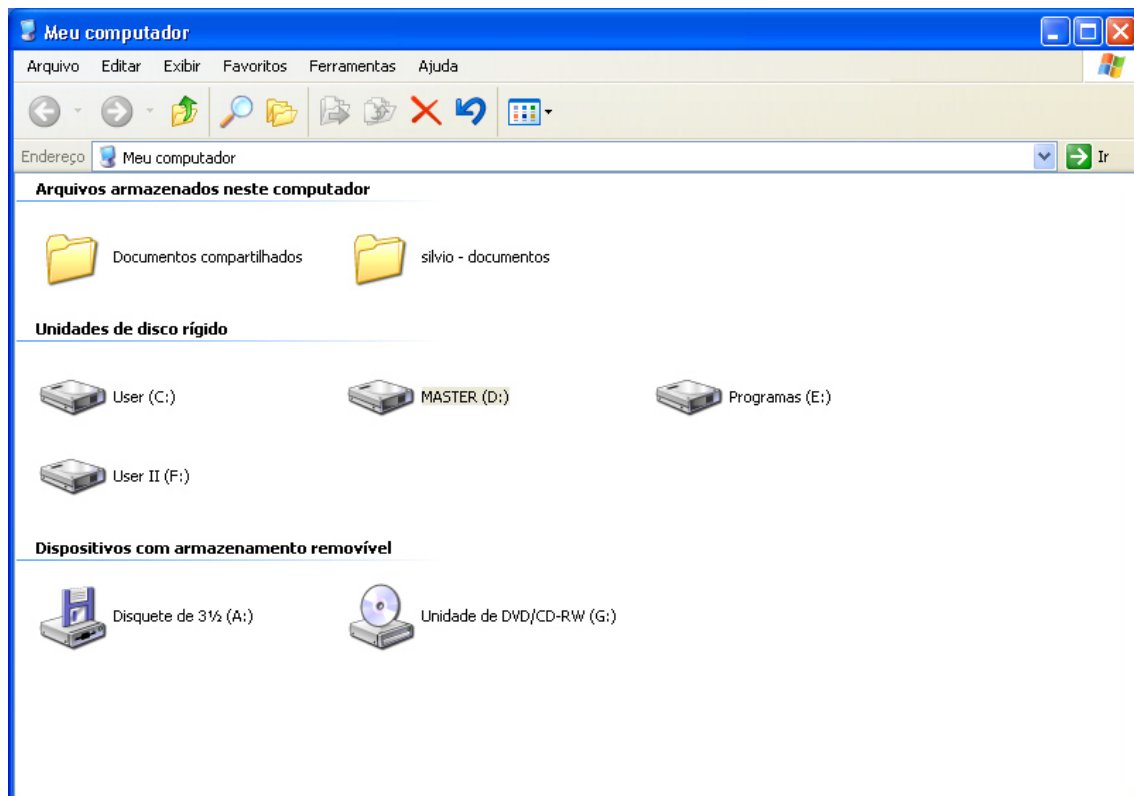


Introdução

Vai formatar um HD e/ou instalar um sistema operacional? Primeiramente, vamos entender sobre sistema de arquivos.

Esquema de diretórios no Windows e no Linux

No Windows particionamos e formatamos o Disco Rígido usando suas ferramentas. Todas as unidades (os Discos Rígidos) e/ou partições receberão nomes de uma letra, como C:, D:, E:. etc.

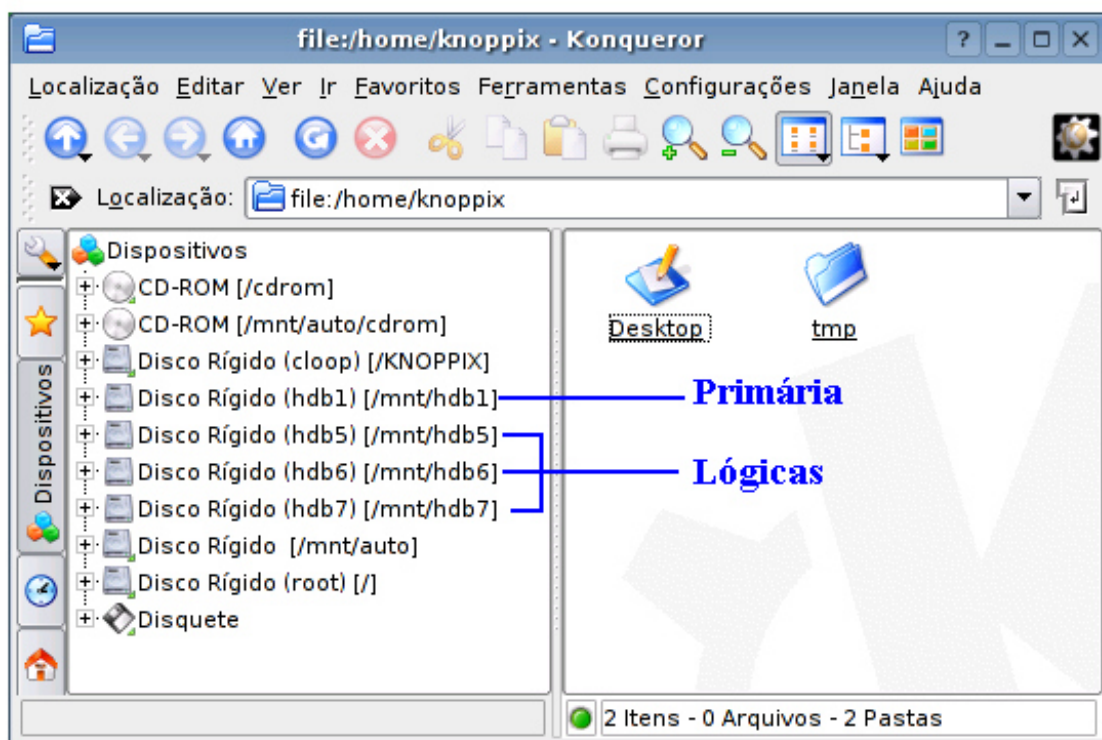


Letras das unidades no Windows

No caso do Linux tanto o particionamento quanto a formatação devem ser realizadas por ferramentas próprias que acompanham, geralmente, cada distribuição. As unidades e/ou partições receberão nomes, geralmente, da seguinte forma: hd+uma letra (a ou b, geralmente)+ um número (exemplo: 1, 2, 5, 5, etc.) como mostrado a seguir:

- /dev/hda1 (Primária)
- /dev/hda2 (Primária)
- /dev/hda5 (Lógica)
- /dev/hda6 (Lógica)
 - /dev/hda7 (Lógica)

Isso pode variar de acordo com o número de partições existente no Disco Rígido e de acordo com o número de Discos Rígido instalados no PC. Na tela a seguir temos o gerenciador de arquivos do Kurumim, onde as unidades estão nomeadas como hdb1, hdb5, hdb6 e hdb7.



Nomes das unidades no Linux

Além dessas diferenças até agora citadas, a forma que encontramos os diretórios no Linux também são diferentes do Windows, bem diferentes para ser mais sensato.

No Windows, tudo é fácil e intuitivo. Já no Linux a estrutura dos diretórios são, geralmente, um pouco mais complexa. Cada diretório do sistema tem uma finalidade específica, o que dá uma maior ordem em tudo. O diretório mais importante em Linux

é o diretório Raiz (marcado por “/”), pois, é abaixo dele que se encontra todos os outros.

No diretório raiz devem estar o conjunto de diretórios mínimo para funcionamento do Linux, como os diretórios /dev, /bin, /proc, entre outros. Voltaremos este assunto no decorrer deste capítulo.

Estrutura de diretórios no Linux

Como acabei de dizer, a estrutura de diretórios no Linux é diferente do que conhecemos no Windows. A seguir temos uma tabela com os principais diretórios bem como uma descrição dos arquivos que estão no diretório.

Diretório	Descrição dos arquivos
/	Esse é o diretório raiz. A partir desse diretório é que se situam todos os outros.
/bin	Arquivos executáveis de comandos essenciais pertencentes ao sistema e que são usados com frequência.
/boot	Arquivos estáticos de boot
/dev	Arquivos de dispositivos do sistema (entrada/saída)
/etc	Arquivos de configuração/ administração do sistema.
/home	Aqui ficam os diretórios locais do usuário
/lib	Aqui ficam de biblioteca essenciais ao sistema
/mnt	Usualmente, é um ponto de montagem de partições
/proc	Aqui ficam as informações do Kernel e dos processadores
/root	Como o nome sugere, este é o diretório local do superusuário
/sbin	Arquivos de sistemas essenciais, mas geralmente, acessíveis somente pelo superusuário
/tmp	Arquivos temporários
/usr	Arquivos dos usuários
/var	Informações variáveis do sistema

Sistemas de arquivos

Quando formatamos um Disco Rígido o que estamos fazendo é preparar o disco para o padrão do sistema operacional, o que damos o nome de *sistemas de arquivos*. O sistema de arquivos definirá como os arquivos serão estruturados, nomeados, acessados, utilizados e implementados pelo sistema operacional.

Todos os aplicativos armazenam e recuperam arquivos no disco, e tudo isso é gerenciado pelo sistema operacional. Se não existissem os sistema de arquivos, nada disso seria possível. O sistema operacional deve organizar os dados no disco de tal maneira que ele “saiba” onde está cada dado. Devido a fatores pertinentes a cada sistema operacional, como desempenho e segurança, existem diversos tipos de sistemas de arquivos e que podem ser incompatíveis entre si. Estaremos vendo a seguir alguns sistemas de arquivos usados pelas várias versões do Windows e pelas distribuições Linux.

Ocorre que o Linux suporta diversos sistemas de arquivos. E isso pode causar confusão em iniciantes que desejam instalar o sistema Linux em algum HD. Muitas distribuições nos permite escolher qual sistema de arquivos. Por isso veremos agora alguns sistemas de arquivos que julgamos ser os “mais conhecidos”.

FAT-16

O sistema de arquivos FAT (*File Allocation Table* - tabela de alocação de arquivos) foi desenvolvido para o DOS 1.0 em meados da década de 80, e foi utilizado também pelo Windows 3X e Windows 95. O seu funcionamento baseia-se em uma espécie de “mapa” para a utilização do disco, que consiste em uma tabela de alocação. Essa tabela indica em qual cluster o arquivo se localiza no disco. O cluster é a menor unidade de alocação de arquivos reconhecida pelo sistema operacional, é formado por vários setores (lembre-se que o setor é a menor divisão física do disco e possui 512 bytes).

O nome FAT 16 é uma referência ao fato que este sistema utiliza 16 bits para o endereçamento de dados, com a máxima de 2^{16} (65.526) posições diferentes. Isso implica no seguinte: os setores possuem 512 bytes e o número máximo de posições permitido é 65.526, dessa forma temos a multiplicação $512 \times 65.526 = 33.554.432$ bytes (32 MB). Mas vamos raciocinar: o sistema FAT 16 permite trabalhar com no máximo 32 MB? Na verdade não. Acontece que o sistema FAT trabalha com clusters, como mencionamos anteriormente, e não com setores individuais, dessa forma, ele agrupa setores vizinhos em uma unidade de alocação (os clusters propriamente ditos) diminuindo assim o número de registros na FAT, tornando possível reconhecer discos de até 2 GB. Para utilizar discos com mais de 2 GB, será necessário particioná-los, ou seja, dividi-los logicamente em outros menores que 2 GB.

O tamanho de cada cluster varia de acordo com o tamanho do disco, veja:

Cluster (em FAT 16)	Capacidade de acesso ao disco
2 KB	128 MB
4 KB	256 MB
8 KB	512 MB
16 KB	1GB
32 KB	2 GB

Um ponto fraco do sistema FAT 16 é o desperdício de espaço. Isso porque cada cluster pode ser ocupado somente por um mesmo arquivo. Em outras palavras, se você armazenar um arquivo de 2 KB em um disco que usa clusters de 32 KB, você estará desperdiçando 30 KB. Esse espaço que sobra simplesmente não é utilizado. O desperdício é proporcional ao tamanho do cluster: quanto maior o cluster, maior o desperdício, que pode chegar até a 25% do tamanho total de um disco. Todo esse desperdício é conhecido como *Slack space* (desperdício em disco).

VFAT

É basicamente o mesmo funcionamento do sistema FAT 16 com o acréscimo de um recurso que permite arquivos com nomes longos. No sistema FAT 16 é permitido apenas nomes no formato 8:3 (oito caracteres no nome + três caracteres na extensão). Esse sistema armazena o nome do arquivo no formato 8.3, e, o nome longo fica oculto em entradas “fantasmas” do diretório. Esse sistema é utilizado pelo Windows 9X/ME.

FAT-32

Com o desenvolvimento dos sistemas operacionais, ficou claro que o sistema de arquivo FAT 16 era totalmente defasado, principalmente pelo fato de não reconhecer uma unidade única que tenha mais de 2 GB e pelo sério problema de desperdício de espaço. A solução para isso é óbvia: diminuir o tamanho dos clusters. E isso foi feito já a partir da última revisão do Windows 95 (Windows 95 OSR2). Esse sistema reconhece Discos Rígidos de até 2 terabytes (2.048 GB).

Cluster	Capacidade de acesso ao disco
512 bytes	256 MB
	8 GB
4 KB	
8 KB	16 GB
16 KB	32 GB
32 KB	2 TB

Sistema NTFS

Mesmo com a diminuição do tamanho dos clusters feitos no sistema FAT 32 o problema de desperdício de espaço não foi resolvido. A grande verdade é que para resolver esse problema de vez o sistema FAT deveria ser trocado, não adiantava mais simplesmente implantar algumas melhorias, era necessário um novo sistema de arquivos. Mas esse sistema já existia a muito tempo, e já era utilizado desde a década de 80 pelo Windows NT: o sistema NTFS. NTFS significa **NT File System** (sistema de arquivos do NT). Existem diferenças imensas desse sistema para o FAT 16 e FAT 32. A começar pelos próprios sistemas operacionais: o Windows 95, 98 e ME (estou citando somente estes apenas por motivos de linha de tempo. Quando o NTFS foi criado não existia o Windows 10 por exemplo. Mas, o Windows 10 é instalado em NTFS também.) foram desenvolvidos para uso doméstico, onde a segurança e instabilidade sempre deixaram a desejar. Já no caso do Windows NT é diferente, pois foi desenvolvido desde o início para ser usado em máquinas que exigem o máximo de estabilidade e segurança. Resumindo: o Windows NT foi construído com objetivos diferentes do Windows 9X.

Quanto ao desperdício de espaço provocado pelos tamanhos dos clusters, não acontece com o NTFS, simplesmente porque não há cluster. A menor unidade de alocação é o próprio setor, ou seja, 512 bytes. Outro problema que ocorre no sistema FAT é quando o PC trava ou é desligado derrepente, o que faz com que seja gerado agrupamentos perdidos (entre outros problemas). No caso do NTFS, é mantido um Log com todas as operações realizadas, e, se o PC travar ou desligar derrepente, ao ser ligado novamente ele examina esse Log, identifica onde foi interrompido e consegue retomar as operações, podendo dessa forma corrigir automaticamente os problemas. As informações dos arquivos como nome, atributos de segurança, localização entre outros, são armazenados no MFT (Master File Table).

O sistema NTFS é usado pelo Windows NT, 2000 e XP, Windows 7 e 8 e o Windows 10.

NTFS5

O NTFS5 é um sistema de arquivo utilizado pelo Windows 2000. Dispõem de um novo recurso que consiste em criptografar os dados gravados no disco rígido, fazendo assim que somente o usuário tenha acesso a ele, enquanto este estiver rodando o Windows.

HPFS

É um sistema de arquivo com basicamente os mesmos recursos do NTFS, porém é utilizado pelo OS/2.

EXT /EXT2

O sistema de arquivos EXT (Extended File System) surgiu de uma evolução no desenvolvimento do Linux que no início utilizava o sistema Minix FileSystem. O EXT2 obviamente foi concebido para atender ao desenvolvimento crescente dos Disco Rígido. Enquanto o EXT reconhece partições de até 2 GB, o EXT2 consegue reconhecer até 4 TB.

EXT3

Na verdade o EXT3 é o EXT2 com o acréscimo de um recurso chamado *journaling*, que consiste no armazenamento de cada uma das operações realizadas sobre os seus registros. É como se fosse uma espécie de “agenda”. Assim, antes dos dados serem escritos, eles são “agendados”. Dessa forma, em casos de falhas (travamentos, desligamento inesperado, etc), será possível retornar para o último estado consistente, em outras palavras, se o sistema falhar o problema poderá ser corrigido automaticamente.

O Ext3 suporta 16TB de tamanho máximo do volume e 2TB de tamanho máximo de um arquivo.

EXT4

Posso te dizer que esse é a melhor escolha de sistema de arquivos ao instalar o Linux. Inclusive ele já é o padrão de muitas distribuições atuais.

Obviamente ele é uma evolução do EXT3 e possui muitas novas funcionalidades, tais como:

- Verificação de integridade do journal: é feito um exame da área de metadados do filesystem para, permitindo dessa forma, que sejam detectados e reparados problemas com integridade estrutural antes de reiniciar;
- Suporte para tamanhos maiores de volumes e arquivos: tamanho máximo para arquivos é de 16 TB para um sistema com blocos de 4k. O tamanho máximo do volume é de 1 EB;
- Numero ilimitado de subdiretórios: dentro de um mesmo diretório pode haver ilimitados subdiretórios;
- Alocação tardia: conhecida Também por atribuição de atraso. Reduz a fragmentação e melhora o desempenho.
- Compatibilidade com versões anteriores: totalmente compatível com ext3 e ext2.

ReiserFS

Sistema desenvolvido por Hans Reiser especialmente para os sistemas Linux. Da mesma forma que o sistema EXT3, o ReiserFS é um sistema de arquivos com suporte a *journaling*. Se faz presente no Linux a partir da versão 2.4 do Kernel. Quanto a utilização do espaço em disco, esse sistema tem se mostrado muito eficiente em comparação aos outros. Ao invés dele usar clusters de tamanhos pre-definidos, ele trabalha com um método de ajuste do tamanho do cluster de acordo com o arquivo que será gravado, o que podemos chamar de “clusters dinâmicos”. Isso quer dizer que ele aloca o tamanho exato que o arquivo necessita. Sem dúvida alguma um sistema muito mais eficiente em termos de utilização de espaço. Além disso, o ReiserFS consegue ser de 10 a 15 vezes mais rápidos que ext2 e ext3 em processo de leitura de arquivos pequenos. Isso porque quando não temos clusters de tamanhos definidos, arquivos pequenos também terão clusters pequenos. Isso faz com que eles fiquem muito próximos uns dos outros, o que torna a leitura mais rápida.

Dessa forma, as principais características desse sistemas são:

- suporte a *journaling*;
- Utilização de espaço muito eficiente;
- Mais rápido nos processos de leitura de arquivos pequenos;

Esse sistema é o mais indicado para utilização em sistemas Linux atualmente. Com a introdução do ReiserFS4, o ReiserFS recebe as vezes o nome de “ReiserFS3”.

Qual sistema de arquivos usar?

Depende do sistema operacional em questão. Em geral você usará:

- **MS-DOS e Windows 95:** FAT 16. Portanto o FAT16 é para sistemas muito antigos, aqueles “pré-históricos”;
- **Windows 95 OSR/2, Windows 98, Windows 2000, ME e XP:** FAT 32;
- **Windows NT, Windows XP, Windows 2000, Windows 7, Windows 8 e Windows 10:** NTFS;
- **Linux:** ReiserFS, EXT, EXT2, EXT3 ou EXT4. Ao instalar Linux sugiro sempre ReiserFS ou EXT4 (escolhe esse se tiver essa opção. É muito melhor).

Resumidamente é isso. Se você é um iniciante, use essa “receita de bolo”. Você vai usar outros sistemas de arquivos somente se for um “expert” e em casos específicos. Se você for um especialista na área, pode usar até FAT se desejar (ou precisar), tudo vai depender do que pretende fazer, estudar ou testar.