

Análise dos sistemas de arquivos FAT 16 e FAT 32

Beatriz P. Martis, Eduardo S. Lira, Ewerton A. Assis

28 de Junho de 2013

Resumo

O presente trabalho, que comporá nota parcial na disciplina de Sistemas Operacionais 2, ministrada pelo professor Eduardo S. de Albuquerque, no Instituto de Informática da Universidade Federal de Goiás, tem como objetivo apresentar os sistemas de arquivos FAT 16 e FAT 32; sistemas de arquivo de relevância histórica, inicialmente apresentados no sistema operacional Microsoft MS-DOS. A análise se pautará em suas questões arquiteturais e em alguns detalhes de implementação.

1 Introdução

2 Características do sistema FAT

A tabela de arquivos FAT é situada no início do volume. São armazenadas duas tabelas para o caso de uma ser danificada, além disso as tabelas e o diretório *root* ou raiz devem ser armazenados num local fixo para que os arquivos necessários para a inicialização do sistema sejam localizados corretamente [?].

Originalmente o sistema FAT foi desenvolvido para máquinas com arquitetura IBM PC, por isso os bytes das entradas FAT são armazenados em “little endian”, assim, ao utilizar uma máquina “big endian” é preciso traduzí-los ao serem movidos dados do disco ou para outro disco [?].

Cada volume tem seu próprio FAT, o qual contém informações de alocação para cada arquivo no volume através de listas ligadas de unidades de alocação, isto é, *clusters* formados por setores em potência de 2. Por meio disso, o sistema também pode indicar quais são as unidades livres para serem utilizadas por um arquivo sendo criado ou estendido [?].

Desconsiderando os *clusters* reservados, o FAT16 é limitado por 2^{16} ou 65.536 *clusters*, já o FAT32 por 2^{32} ou mais de 4 bilhões de *clusters* [?] (mais detalhes na Seção 5 - Particionamento de disco). O conceito básico do sistema FAT, como pode ser visto na Figura ??, é que, para cada arquivo ou diretório, é alocado uma estrutura de dados, chamada entrada de diretório, que contém o nome do arquivo, tamanho, endereço de início e outros metadados; seu conteúdo é

guardado em um ou mais *clusters*, que podem ser encontrados através da tabela FAT, a qual identifica o próximo *cluster* e identifica o status de alocação dos *clusters* [?].

O FAT surgiu como uma boa solução para o gerenciamento de disco, principalmente porque a tabela era pequena o suficiente para caber na memória, permitindo um acesso aleatório rápido, mas se tornou inviável com o aumento do tamanho da tabela em discos fixos, o que exigia paginação. Outro problema é que a informação a respeito de espaço livre era espalhada entre os setores do FAT, resultando numa alta fragmentação: para *clusters* grandes, era desperdiçado, em média, metade de um *cluster* em cada arquivo [?].

3 Estrutura do volume e particionamento

Um sistema FAT geralmente é organizado da seguinte forma [?]: um setor de boot, no início do volume (veja a subseção 3.1); as tabelas de alocação FAT1 e FAT2, sendo esta última uma redundância da primeira (veja mais detalhes na seção 4); o diretório *root* ou raiz; e por fim outros diretórios e todos os arquivos representados. As áreas FAT1 e FAT2 geralmente são agrupadas arquiteturalmente sob o título de área FAT e as áreas do diretório *root* e dos demais diretórios e arquivos sob o título de área de dados [?].

Na versão FAT12/16 o diretório *root* ou raiz encontra-se no início da área de dados; enquanto na versão FAT32 o diretório raiz pode estar em qualquer parte da área de dados — tal situação é considerada rara, tendo como padrão o uso da atribuição da versão FAT12/16 [?].

O sistema de arquivos é alocado em *clusters*, unidades de dados do sistema de arquivo [?], que são um grupo consecutivo de setores — o número de setores geralmente é em base 2 (1, 2, 4, 8, 16, 32 ou 64) e o número (“endereço”) do *cluster* deve ser representável em até 16 bits. Curiosamente, os números de *cluster* começam com o número 2: não existem os *clusters* 0 e 1; além destes estarem todos na área de dados. O tamanho de um *cluster* é determinado, geralmente, pelo tamanho do volume de armazenamento. Essas restrições limitam o tamanho máximo de um volume de armazenamento para até 4GB, já que o tamanho mínimo do *cluster* — e, portanto, do setor — é de 512 bytes e o tamanho máximo de um *cluster* é de 64 kilobytes [?] — embora em outra fonte o tamanho máximo de um *cluster* seja 32 kilobytes [?]. Cada *cluster* é identificado como “não-usado”, “em uso”, “*bad cluster*” ou “último *cluster* em um arquivo”; informações que estão disponíveis na tabela de alocação (mais detalhes na seção 4).

Como as áreas FAT e o setor de boot não usam a estrutura de *cluster*, determinar a localização do primeiro *cluster* não é uma tarefa fácil [?]. Na versão FAT12/16 o primeiro *cluster* é o primeiro setor do diretório raiz, que é alocado quando o sistema de arquivo é criado; na versão FAT32, o primeiro *cluster* é o primeiro setor da área de dados.

O diretório raiz contém entradas para cada arquivo e diretório na raiz; o que distingue o diretório raiz dos demais diretórios é sua posição no volume e seu

tamanho fixo: 512 entradas para volumes em disco rígido e um tamanho variável para disquetes, dependendo do tamanho do volume. Como será apresentado na seção 5 (página 4), diretórios são estruturas semelhantes aos arquivos, com estruturas próprias que os distinguem.

3.1 Setor de boot

O setor de boot, como em geral é encontrado em outros sistemas de arquivos, tem como finalidade armazenar código que será executado pela BIOS do sistema computacional durante a inicialização do disco. Localizado no primeiro setor do volume, além de conter código de execução, esta área do sistema de arquivos contém informações sobre o próprio sistema, como o tamanho das tabelas de alocação (áreas FAT) e demais informações sensíveis ao funcionamento do sistema [?]: o número de estruturas FAT e o tamanho destas estruturas; tamanho das estruturas de diretório; tamanho do *cluster*; dentre outras informações. Geralmente a versão é marcada por tag, em formato *string*, com os possíveis valores “FAT12”, “FAT16”, “FAT32” ou apenas “FAT”. Fora esse marcador, não existe, dentre as informações do setor de boot, qualquer dado acurado para a versão do sistema de arquivos FAT; essa informação só pode ser obtida a partir de cálculos feitos nas informações armazenadas [?].

Na versão FAT12/16, o setor de boot ocupa um único setor, enquanto na versão FAT32 muitos outros setores podem ser armazenados. Na versão FAT32 é possível encontrar informações sobre o próximo *cluster* disponível e a quantidade de *clusters* disponíveis (na estrutura de dados FSINFO); o endereço do setor com uma cópia de segurança (*backup*) do setor de boot, geralmente no setor 6, que pode ser usado por ferramentas quando o setor original estiver corrompido; e o endereço do setor de início do diretório raiz.

Outras informações não essenciais estão também disponíveis no setor de boot: o nome OEM (*OEM name* [?]), usado para indicar a ferramenta que criou o sistema de arquivos; um número serial de 4 bytes que determina quando o sistema de arquivos foi criado; e, como já mencionado, uma tag para determinar qual a versão do sistema de arquivos, embora tal informação não seja objetivamente correta. Outra informação importante, a qual é também armazenada no diretório raiz, é o rótulo do volume do sistema de arquivos e um rótulo do sistema de arquivos.

O código de execução, anteriormente mencionado, são instruções de *JUMP* que levam a máquina a executar os códigos de execução, armazenados nos bytes 62 a 509 na versão FAT12/16 e nos bytes 90 a 509 na versão FAT32. Sistemas FAT geralmente apresentam código de execução, mesmo que estes volumes não sejam inicializáveis (*bootable*). A MBR, geralmente no início do disco ou dispositivo e fora da responsabilidade do sistema de arquivos, é responsável por inicializar a execução do código presente no setor de boot.

4 Tabela de alocação

A tabela de alocação no sistema FAT tem uma relevância chave na arquitetura deste sistema de arquivo; como o próprio nome do sistema de arquivos indica (FAT é acrônimo para *File allocation table*, ou tabela de alocação de arquivos, em tradução livre).

Como foi apresentado na seção 3 (página 2), as tabelas de alocação são armazenadas nas áreas de disco FAT1 e FAT2, as quais, como mencionado, não usam a estrutura lógica dos *clusters*. A FAT2 é uma replicação do FAT1 e tem por finalidade atender certo nível de segurança ao indicar possíveis inconsistências que possam surgir no sistema de arquivo. O tamanho da área FAT é geralmente calculado pelo tamanho da estrutura FAT e a quantidade de *clusters* no sistema.

As tabelas de alocação armazenadas na área FAT têm por finalidade indicar o status dos *clusters* da área de dados.

5 Estrutura de diretório

6 Nomes de arquivos

7 Operações com arquivos

8 Aspectos de segurança

9 Conclusão