FACULDADE ZACARIAS DE GÓES

JUSSARA REIS DA SILVA

SISTEMA DE ARQUIVOS

Valença

2010

JUSSARA REIS DA SILVA

SISTEMAS DE ARQUIVOS

Trabalho apresentado como requisito parcial para AVII da disciplina Sistemas Operacionais do curso de Sistema de Informação da Faculdade Zacarias de Góes, sob orientação do professor Adonai Medrado.

Valença

2010

INDICE

1. INTRODUÇÃO

A era do computador além dos benefícios trouxe também os crimes digitais, para analisar esses crimes, que requer um bom conhecimento e técnicas aprimoradas, a Ciência forense adaptou seus métodos e criou a disciplina Forense Computacional.

Este trabalho tem como objetivo apresentar os conceitos de Sistema de Arquivos, exemplificar alguns sistemas utilizados na plataforma Windows e uma pequena análise sobre Forense computacional e sua aplicação em um ambiente baseado em NTFS.

2. SISTEMA DE ARQUIVOS

Segundo MACHADO (2007); MAIA (2007), os arquivos são gerenciados pelo sistema operacional de maneira a facilitar o acesso ao seu conteúdo, sendo que um arquivo é constituído por informações que podem representar instruções ou dados logicamente relacionados. E a parte do sistema responsável por essa gerência é o sistema de arquivo, que faz manipulações nos arquivos, como por exemplo: identificar, organizar, compartilhar, definir métodos de acesso, proteger e realizar operações de entrada e saída.

A organização do arquivo pode ser definida pelo criador do arquivo quando este criar o arquivo, podendo esta ser do tipo simples - ocorrendo através de uma seqüência não-estruturada de bytes, que é o tipo de organização onde o sistema de arquivo não impõe nenhuma lógica para os dados, deixando o trabalho para a aplicação que definirá toda a organização; existem outros tipos de organizações, por exemplo: a seqüencial, a relativa e a indexada.

Os métodos de acesso dependem de qual forma o arquivo está organizado, estes acessos podem ser *seqüencial, direto e indexado*; através dos métodos de acesso o sistema de arquivos pode recuperar registros de diferentes maneiras, por exemplo, no acesso por chave ou indexado, o arquivo deve possuir uma área de índice onde existam ponteiros para os diversos registros, assim quando a aplicação acessar um registro, será especificada uma chave que o sistema pesquisará na área de índice do ponteiro correspondente, a partir daí liberando o acesso direto ao registro. As rotinas de entrada e saída através de uma interface simples e uniforme facilita a comunicação entre a aplicação e os diversos dispositivos de E/S.

Os atributos de um arquivo são as informações de controle que este possui, tais atributos variam a depender do sistema de arquivo, mas algumas informações como, por exemplo: tamanho do arquivo e identificação do criador, são encontradas na maioria dos sistemas. Alguns atributos podem ser modificados pelo sistema operacional, mas os atributos especificados na criação do arquivo como organização e data/hora da criação não podem ser modificadas.

O sistema de arquivo utiliza os diretórios, que são estruturas de dados que contém informações a referência do arquivo, para organizar os arquivos contidos nos discos. As estruturas de diretório podem ser: *single level directory* (nível único), *user file directory* e *master file directory* são estruturas de dois níveis e *tree-structured directory* (estrutura de diretório em árvore) e utilizado pela maioria dos sistemas.

O controle de quais áreas do disco está livre para uso na criação do arquivo é controlada pelo sistema operacional, que realiza estruturas de dados que armazena

informações em uma lista ou tabela permitindo a identificação de blocos livres para alocação dos arquivos. A gerência de alocação de espaço em disco é tão importante quanto a gerência de espaço livre no disco. As principais técnicas de alocação são:

- Alocação Contígua – permite que um arquivo em blocos seja armazenado e seqüencialmente disposto no disco; Para selecionar qual o segmento na lista de blocos livres deve ser utilizado para alocação, existem estratégias que facilitam a escolha, as principais são: First-fit (o primeiro segmento livre com tamanho suficiente para alocar o arquivo é utilizado), Best-fit (seleciona o menor segmento livre com tamanho suficiente para armazenar o arquivo) e Worst-fit (o maior segmento é alocado).

Este tipo de alocação apresenta um problema que é a fragmentação dos espaços livres, para contornar estes problemas é necessário o uso da rotina de reorganização dos arquivos no disco para que só exista um segmento livre, conhecida como desfragmentação.

- Alocação Encadeada – Um arquivo independente da sua localização física pode ser organizado como um conjunto de blocos ligados pela logicamente no disco. Neste tipo de alocação a fragmentação não ocasiona nenhum tipo de problema, a desvantagem do uso desta técnica é que os blocos de arquivos só podem ser acessados sequencialmente, não podendo ocorrer o acesso direto aos blocos.
- Alocação Indexada – permiti o acesso direto aos blocos do arquivo e não utiliza informações de controle nos blocos de dados, solucionando o problema da alocação encadeada.

Os mecanismos de proteção são muitos importantes para o sistema de arquivo, sendo assim, este utiliza mecanismos de proteção ao acesso das informações gravadas nos discos e possibilitam o compartilhamento de arquivo entre usuários quando requisitado. Esses mecanismos podem ter diferentes níveis de proteção com vantagens e desvantagens diferentes para cada tipo de sistema, os três mecanismos mais utilizados são: senha de acesso, grupos de usuário e lista de controle de acesso.

A implementação do *buffer cache* é feita pelo sistema de arquivo para minimizar a lentidão de acesso ao disco se comparado com o acesso a memória principal. Nesta técnica o sistema operacional reserva uma área da memória para que se tornem disponíveis cachês utilizados em operações de acesso ao disco, e quando uma operação é realizada o sistema verifica se a informação desejada está no buffer cache, se a resposta for positiva, não será necessário o acesso ao disco. Caso a resposta seja negativa, a operação é realizada e o cache é atualizado. Cada sistema adota uma política para substituição de blocos diferentes, por causa do tamanho limite do cache, por exemplo: FIFO e LRU.

Existem vários sistemas de arquivos, mas o Sistema Operacional Windows só reconhece três tipos: FAT16, FAT32 e NTFS, já o Linux e outros sistemas Unix possuem uma variedade grande de sistemas de arquivos, por exemplo: EXT2, EXT3, ReiserFS, JFS entre outros.

3. SISTEMAS DE ARQUIVOS NO WINDOWS

Os sistemas de arquivos reconhecidos pelo sistema operacional Windows que são sistemas baseados em DOS são:

- FAT 12 – pouco usado atualmente, era utilizado nos disquetes e nas primeiras versões do MS-DOS, com endereços de 12 bits para endereçar os clusters e permitia partições de até 16 MB.

- FAT16 – File Allocation Table (tabela de alocação de arquivos). De acordo com MORIMOTO (2007), este sistema de arquivo é compatível com todos os sistemas operacionais e dispositivos como câmeras, mp3, mp4, cartão SD, pendrive de até 2GB e palmtops.

Neste sistema de arquivos, o HD é dividido em clusters, que possui um endereço único permitindo ao sistema localizar os arquivos armazenados.

A maior limitação deste sistema é o uso de endereços de 16 bits para endereçamento dos clusters dentro da partição, permitindo que seja divido em no máximo 65536 clusters, que não podem ser maiores que 32KB, limitando-se a 2GB por partições criadas.

Cada cluster só pode conter um arquivo, caso um o arquivo seja de tamanho menor que o tamanho total do cluster, este espaço restante será desperdiçado, pois não poderá ser ocupado por outro arquivo.

O FAT16 possui uma exceção, sua utilização no sistema operacional Windows NT permitia criar partições de até 4GB, utilizando cluster de 64KB, mas o desperdício de espaço ainda continuava.

- FAT32 – Foi lançada pela Microsoft para resolver o problema do tamanho das partições e resolver o desperdício de espaço causado com o uso da FAT16. Este sistema foi lançado no Windows 95 OSR/2 (revisão do Windows 95 que originalmente somente utilizava sistema FAT16) e continuou sendo usado nas versões posteriores. (Villela, 2007)

No sistema FAT32 é utilizados endereços de 32 bits para endereçamento dos clusters, possibilitando a criação de partições com tamanhos maiores que 2GB, chegando até a 2TB; os clusters diminuem de tamanho, diminuindo o desperdício em discos.

A diminuição dos tamanhos dos clusters melhorou o desempenho do uso do espaço livre, mas aumentou a lentidão dos discos rígidos pelo grande número de clusters.

O FAT 32 apresenta alguns outros problemas, como por exemplo: existe uma dificuldade de acesso ao disco com o sistema FAT32 por sistema operacional que não seja o Windows 95 OSR/2, assim como, por alguns utilitários de manutenção de discos rígidos antigos. (TORRES, 1997).

- NTFS – Este sistema começou a ser desenvolvido no inicio da década de 1990, juntamente com o Windows NT e seus conceitos é baseado no sistema de arquivo HPFS (High Performance File System) usado pelo OS/2 da IBM.

O NTFS – New Technology File System - resolveu o problema do desperdício de espaço, pois a menor unidade de alocação possui 512 bytes, e o tamanho do seu cluster poderá ser de 512 bytes independentes do tamanho da partição, mas em contrapartida o problema passou a ser a necessidade de grande processamento para encontrar os dados desejados, devido ao grande número de clusters existentes, diminuindo o desempenho do disco.

O sistema NTFS possui muitas vantagens em relação ao FAT, por exemplos:

- os nomes de arquivos e pastas podem utilizar caracteres em Unicode, facilitando o acesso por usuários que não utilizam o alfabeto ocidental sem o auxilio de um driver ou programa adicional.

- É tolerante a falhas, pois mantém um log de todas as operações realizadas, podendo assim saber em qual ponto uma tarefa parou caso o micro seja desligado bruscamente no meio de uma operação, reduzindo a perda de dados e de tempo.

- O sistema traz recursos de segurança, onde os arquivos podem ser protegidos pelo usuário, podendo este bloquear seus arquivos através do uso de encriptação, como também a compactação do arquivo com senha. (MORIMOTO, 2007)

Existiram diversas versões do NTFS para corrigir falhas e dá suporte a hardware, o NTFS 1.0 ou 3.1 como era conhecido usado no Windows NT 3.1, NTFS 1.1 ou versão 4 usado no Windows NT 4 e no Windows NT 3.51 e o NTFS 5.0 lançada junto o Windows 2000. (ALECRIM, 2004)

4. ANÁLISE FORENSE PARA AMBIENTES NTFS

Segundo THORTON (1997), a ciência forense é exercida em favor da lei para uma resolução correta de um conflito, baseada em procedimentos científicos para obtenção de informações que possam ser utilizadas durante uma disputa judicial.

A ciência forense até pouco tempo atrás era utilizada somente para solucionar mistérios policiais, mas com o advento da era computacional e os problemas envolvendo o meio, fez-se necessário a adaptação da ciência forense para meios computacionais, criando então a forense computacional. (OLIVEIRA,2001)

A forense computacional serve para identificar e entender as relações de causa e efeito de todas as ações ocorridas dentro do sistema, através da manipulação e análise de evidências digitais. Para que se tenha um bom sucesso nas análises é necessário o raciocínio lógico e uma mente aberta para o entendimento das causas e efeito. Estando compreendida a complexidade envolvendo o caso, começa-se a identificar os métodos para manipulação de evidências relacionadas ao sistema de arquivo.

Alguns procedimentos para manipulação de sistemas de arquivos são gerais e não dependem do Sistema Operacional (SO) que está sendo analisado, mas o desenvolvimento de algumas documentações depende de técnicas especificas para cada tipo de SO e situação. Pode-se citar como exemplos:

- **Análise sobre cópias:** As analises devem ser feitas a partir de cópias dos dados originais, de preferência bit a bit, ou seja, por imagem.
- **Assinaturas digitais:** As assinaturas digitais é uma maneira de assegurar a confiabilidade que os dados recolhidos para analise são iguais aos originais.
- **Sem Permissão de escrita e execução:** A desabilitação da função de escrita e execução do sistema na imagem feita dos dados originais evita a alteração das informações a ser analisadas.
- **MACTimes:** É um termo usado para fazer referencia aos três atributos (mtime, atime e ctime) de tempo presente em todos os arquivos e diretórios da maioria dos SO´s. Estas são umas das melhores formas de reconstituir o que aconteceu no sistema de arquivos em um determinado período do passado. No Windows NT esses atributos são chamados de LastWriteTime, LastAcessTime e CreationTime.
- **Arquivos excluídos:** a utilização de um mecanismo de recuperação de arquivos excluídos auxilia na eficiência da analise forense, mas não é sempre que funciona,

pois a exclusão pode ter sido feita por ferramentas de exclusão de forma segura conhecida com *wipping*.

A análise forense no sistema operacional Windows NT, pode-se dizer que é bastante complexo, pois se trata da utilização de um sistema operacional com pouca documentação e sistema fechado. O ambiente favorável para análise forense computacional é um ambiente que seja controlado e com ferramentas confiáveis, por exemplo, outra plataforma de sistema operacional. Porém o sistema de arquivos NTFS possui características próprias não suportadas por drivers de outros sistemas, por exemplo, o Linux.

Como citado no item 3, o sistema de arquivo NTFS é nativo do Windows NT, que foi desenvolvido para suprir as necessidades do mercado não satisfeita com o FAT. O processo de formatação de uma partição com o NTFS cria um MFT (Master File Table) e diversos arquivos de sistema. AMFT contém informações sobre todos os arquivos e diretórios do volume.

Muitos problemas afetam a maioria dos SO´s em relação ao MACTimes, um desses problemas é a falta de registro histórico com todas as modificações dos arquivos, a solução para este poderia ser a habilitação de um log para registrar o acesso e fornecer os dados com o que foi feito pelo usuário aos arquivos do sistema.

Além do problema citado o ambiente Windows possui diversos problemas com a análise do *MACTimes*, por exemplo, quando se copia um arquivo para um outro de nome diferente, data da modificação continua a do original, só mudando a data do último acesso e criação, ficando o arquivo com a data da modificação antes da data da criação.

Alternate Streams é um mecanismo utilizado para embutir um arquivo dentro do outro, sem que seu conteúdo seja alterado. No sistema NTFS este mecanismo é um dos pontos importantes na análise forense, pois pode-se dizer que todos os arquivos NTFS possui um outro arquivo sem nome embutido.

Existem diversas ferramentas para efetuar uma análise, por exemplo: o TASK, é um conjunto de ferramentas de código aberto para analise forense com suporte ao sistema de arquivo NTFS, desenvolvida por Brian Carrier.

5. CONSIDERAÇÕES FINAIS

O conhecimento do funcionamento detalhado dos sistemas de arquivos possibilita um melhor conhecimento de quais técnicas utilizar para um melhor desempenho do sistema. Como também, a noção da complexidade da análise forense de um software proprietário, por não disponibilizar de documentação adequada e possui características não suportadas por outros sistemas operacionais.

6. REFERÊNCIAS BIBLIOGRÁFICAS

MACHADO, Francis B.; MAIA, Luiz P.. **Arquitetura de Sistemas Operacionais**. 4ª Ed. Rio de Janeiro: LTC, 2007. Pg 213-229.

VILLELA, Anderson. **Sistema de Arquivos do Windows XP.** Disponível em: http://www.OFICINADANET.com.br/artigo/298/sistema_de_arquivos_do_windows_xp. Acessado em: 19 nov. 2010.

MORIMOTO, Carlos E. **Hardware, o Guia Definitivo.** Disponível em: http://www.gdhpress.com.br/hardware Acessado em: 18 nov 2010.

TORRES, Gabriel. **Sistema de arquivos.** Disponível em: http://www.clubedohardware.com.br/artigos/313 Acessado em: 19 nov 2010.

ALECRIM, Emerson. **Sistemas de Arquivos NTFS**. Disponível em: http://www.infowester.com/ntfs.php Acessado em: 19 nov 2010.

THORTON, J.; **The general assumptions and rationale of forensic identification**. Modern Scientific Evidence: The Law and Science of Expert Testimony; West Publishing Co.; Volume 2; 1997;

OLIVEIRA, Flávio. **Metodologias de análise forense para ambientes baseados em NTFS.** Disponível em: http://www.las.ic.unicamp.br/paulo/papers/2001-SSI-flavio.oliveira-forense.ntfs.pdf. Acessado em: 25 nov. de 2010.

HOLPERIN, Marco; LEOBONS, Rodrigo. **Análise Forense.** Disponível em: http://www.gta.ufrj.br/grad/07_1/forense/task.html Acessado em: 26 nov 2010.