

Sistemas Operacionais

Windows 2000



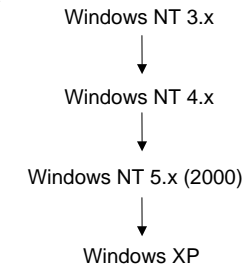
2ª edição

Revisão: Fev/2003

Capítulo 10

Introdução

- Objetivo de desenvolver um sistema operacional multitarefa para ser usado tanto em ambientes monousuário como multiusuário
 - Nome NT vem *New Technology*
 - Aparência (janelas) são oriundas do *Windows 3.x for workgroup*
- Árvore genealógica



Instituto de Informática - UFRGS
Oliveira, Carissimi, Toscani

Sistemas Operacionais

2

Windows2000

- Estrutura básica do sistema operacional Windows NT 4
 - Difere pela inclusão de serviços orientados a sistemas distribuídos e de redes
- Quatro versões diferentes:
 - Professional
 - Server
 - Advanced Server
 - Datacenter Server

Instituto de Informática - UFRGS
Oliveira, Carissimi, Toscani

Sistemas Operacionais

3

Arquitetura do Windows 2000

- Sistema operacional simétrico
- Fortemente inspirada no modelo de microkernel com servidores
 - Processos servidores rodam em modo protegido (supervisor)
- Implementa também uma organização em camadas
 - Cada camada oferece serviços a camada superior
 - Explora orientação a objetos
- Explora também filosofia de máquina virtual
 - Facilidade suportada pelo conceito de subsistemas
 - Processo separado que oferecem um ambiente de execução compatível com um determinado sistema operacional
 - e.g.: subsistema Win32, subsistema POSIX

Instituto de Informática - UFRGS
Oliveira, Carissimi, Toscani

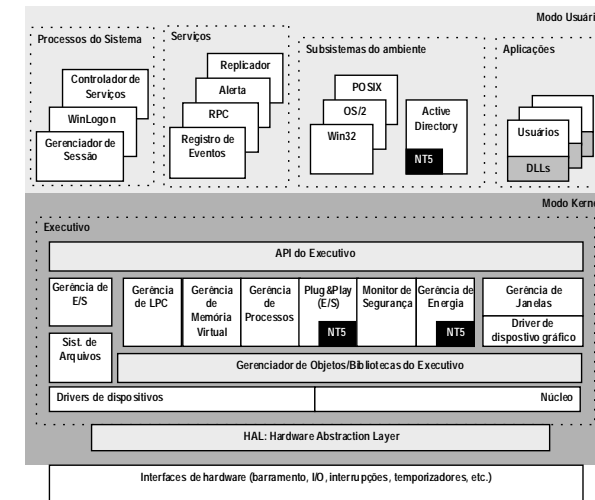
Sistemas Operacionais

4

Estrutura básica do kernel

- Três módulos funcionais:
 - *Hardware Abstraction Layer* (HAL)
 - Drivers de dispositivos
 - Executivo
- O HAL é a camada de portabilidade
 - Implementação específica de serviços de base em função da plataforma
- Drivers de dispositivos
 - Interface padrão entre o subsistema de E/S do executivo e o HAL
 - Portabilidade
- O kernel do windows propriamente dito é o executivo

Componentes da estrutura do kernel



Executivo

- Unidade de execução é a *thread* (objeto *thread*)
 - Podem ser interrompidas em qualquer ponto
 - Em máquinas multiprocessadoras podem ser executadas em qualquer CPU
- O executivo é um programa concorrente
 - Em máquina multiprocessadores utiliza *spin-lock* (*busy waiting*) em monoprocessores não
 - Diferença básica entre a versão monoprocessador e a multiprocessador
- Escalonador apresenta algumas variações para *threads*
 - Tratamento de prioridades
 - Modelo de afinidade (*soft* e *hard*)

Componentes do executivo

- Gerência de objetos
 - Criação, gerenciamento exclusão de objetos do executivo do W2K
 - Objetos são todos recursos do sistema (threads, memória, mutex, etc...)
- Gerência de processos e threads
- Gerência de memória virtual
- Monitor de segurança
 - Implementa políticas de segurança e auditoria no sistema
- Módulo de suporte a *Local Procedure Call* (LPC)
 - Implementa a troca de mensagens entre processos
- Gerência de entrada e saída

Escalonamento

- Unidade de escalonamento é a *thread*
- Escalonador preemptivo com prioridades
- Prioridades organizadas em duas classes:
 - ┆ Tempo real: prioridade estática (níveis 16-31)
 - ┆ Variável: prioridade dinâmica (níveis 0-15)
- Cada classe possui 16 níveis de prioridades
 - ┆ Cada nível é implementado por uma fila em uma política round-robin
 - Múltiplas filas: classe de tempo real
 - Múltiplas filas com realimentação: classe de tempo variável
 - ┆ Threads da classe tempo real tem precedência sobre as da classe variável

Escalonamento classe variável

- Dois parâmetros definem a prioridade de uma *thread*:
 - ┆ Valor de prioridade de base do processo
 - ┆ Prioridade inicial que indica sua prioridade relativa dentro do processo
- Prioridade da *thread* varia de acordo com uso do processador
 - ┆ Preemptada por esgotar o *quantum*: prioridade reduzida
 - ┆ Preemptada por operação de E/S: prioridade aumentada
 - ┆ Nunca assume valor inferior a sua prioridade de base, nem superior a 15
- Fator adicional em máquina multiprocessadoras: afinidade!
 - ┆ Tentativa de escalonar uma *thread* no processador que ela executou mais recentemente.
 - Princípio: reaproveitamento de dados na memória cache

Gerência de memória

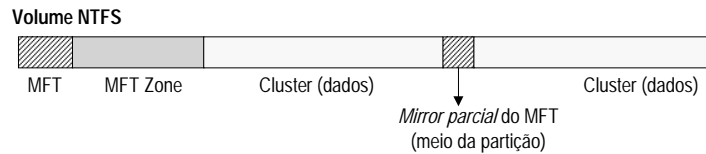
- Implementa memória virtual através de paginação por demanda
 - ┆ Tamanho da página varia entre 4 kbytes e 64 kbytes
- Paginação a dois níveis com espaço lógico de endereçamento de 32 bits
 - ┆ Provê uma extensão para processadores de 128 bits
 - *Very Large Memory* (VLM)
- Política de substituição depende do processador:
 - ┆ FIFO local
 - ┆ LRU implementado através do algoritmo do relógio

Sistema de arquivos: NT File System

- Volume corresponde a uma partição lógica do disco
 - ┆ NTFS gerencia um volume de forma independente
- Volume consiste em:
 - ┆ Arquivos
 - Engloba arquivos de dados, diretórios, *bitmaps*, informações gerenciais
 - ┆ Espaço livre
- Modelo NTFS é "tudo que está no disco é arquivo"

Estrutura de uma partição NTFS

- Espaço do disco é dividido em clusters
 - Tamanho do cluster é definido em função da capacidade do disco
 - Varia de 512 bytes a 64 Kbytes (4K é o valor *default*)
 - Compromisso é desempenho versus desperdício de disco
- Uma partição NTFS é dividida em duas partes:
 - Área MFT (*Master File Table*)
 - Área destinada ao armazenamento de dados



Cluster

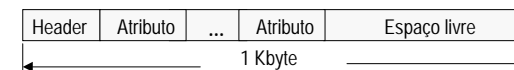
- Unidade de alocação
- Conjunto de setores
 - Fator de cluster
- *Logical cluster number* (LNC)
 - Numeração sequencial de todos os clusters do disco
- *Virtual cluster number* (VCN)
 - Numeração sequencial (a partir do zero) dos clusters alocados a um arquivo
- Um VCN pode ser mapeado para qualquer LCN

MFT - Master File Table

- Arquivo que mantém informações sobre a gerência do sistema de arquivos (*localização, bitmaps, etc*)
- MFT tende a sofrer fragmentação a medida que ele é modificado
 - Introdução da Zona MFT
- Zona MFT
 - Area reservada onde dados não podem ser armazenados, entretanto é contabilizado como espaço livre do disco
 - Em caso de problema de espaço em disco a zona MFT é reduzida
 - Objetivo da zona MFT é deixar o MFT crescer sem provocar fragmentação

Estrutura do MFT

- Arquivo mais importante do NTFS
 - Centraliza informações sobre os demais arquivos do disco
 - Inclusive ele próprio
- MFT é dividido em registros de tamanho fixo de 1 Kbyte
 - Cabeçalho (*header standard*) com informações básicas sobre o arquivo
 - Sequência de um ou mais atributos
- Cada arquivo possui no mínimo um registro



Metafiles

- 16 primeiros registros do MFT são reservados aos *metafiles*
 - O primeiro é o próprio MFT
 - É armazenada uma cópia desses registros no *mirror* MFT
- Metafiles são arquivos do sistema
 - Localizados no diretório *root*
- Responsáveis pelo gerenciamento do sistema de arquivo

Arquivos metafiles do NTFS

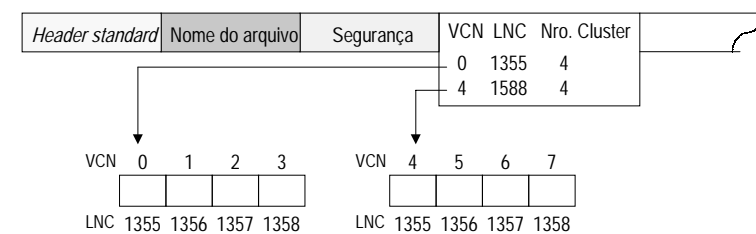
- \$MFT : *Master File Table*
- \$MFTMIRR: *Mirror* do MFT
- \$LOGFILE : Arquivo de log da journalização
- \$VOLUME : Informações sobre o volume (número, data criação, bit *sujo*, etc)
- \$ATTRDEF: Definições de atributos
- \$. : Diretório raiz
- \$BITMAP : Informações de clusters livres/ocupados
- \$BOOT : Setor de *boot*
- \$BADCLUS: Lista dos blocos danificados do disco
- \$QUOTA : Informação sobre quotas de espaço em disco (usuários)
- \$UPCASE : Mapeamento de caracteres minúsculos para maiúsculos

Atributos

- Arquivos são mantidos na forma de pares de atributo/valor
 - Um dos atributos são os dados
 - Demais atributos são nome, *timestamps*, segurança, etc
- Atributos residentes
 - Valor do atributo é armazenado no próprio MFT
- Atributos não residentes
 - Valor do atributo é armazenado em uma área a parte (*run* ou *extent*)
 - Atributos podem necessitar mais de um MFT
 - Atributo "lista de atributos" indica o próximo MFT

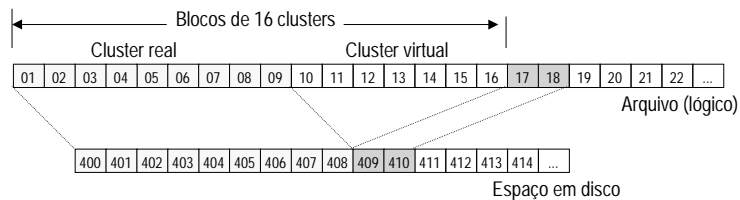
Exemplo de atributo

Registro MFT:



Compressão

- É um atributo do arquivo
- Qualquer arquivo (inclusive diretório) pode ser armazenado de forma compactada
 - Transparente para a aplicação
- Algoritmo de compactação utiliza como base blocos de 16 *clusters*
 - Cluster virtual



Conceito de arquivos e *streams*

- *Stream* é um arquivo cujo dados são armazenados no próprio registro MFT na área de "espaço livre"
 - Destinado a arquivos sem dados ou com poucos dados
 - Não ocupa um cluster (prevenção contra a fragmentação interna)
- Atributos são um tipo de stream
 - Pode haver mais de um stream associado a um arquivo

Diretório

- Arquivo especial que mantém registros sobre arquivos e diretórios
- Um registro é composto por:
 - Nome do arquivo
 - Atributos de base
 - Referência para um elemento MFT
- Internamente é organizado como uma árvore binária
 - Ordenação alfabética
- Diretório raiz possui uma referência no MFT

Segurança e proteção

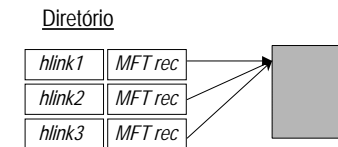
- Cada arquivo possui um atributo de segurança no MFT
- Atributo de segurança fornece:
 - Direitos de modos de acesso
 - Lista de controle de acesso para usuários/grupo não proprietários do arquivo
- Através de uma add-on permite que arquivo e diretório sejam encriptados
 - *Encrypting File System* - EFS

Jornalização no NTFS

- Transações são mantidas em um metafile (\$LogFile)
- Armazena apenas metadados
 - Garante a consistência do sistema de arquivos não de dados de usuário

Hard link

- Mecanismo de alias
 - Um arquivo com mais de um nome
- Cada nome corresponde a uma entrada em diretório todas apontando para o mesmo registro MFT
 - O arquivo só é completamente apagado quando o último nome é removido



Link simbólico

- Possibilita a criação de diretórios virtuais
- Não é reconhecido pelo sistema operacional
 - Aplicativos próprios para seu gerenciamento
 - **Atenção:** remover um link simbólico através de um gerenciador de arquivos do Windows (explorer, exemplo) leva a remoção do objeto apontado

Gerência de entrada e saída

- Baseada em trocas de mensagens entre o driver de dispositivo e o subsistema de entrada e saída
 - IRP (*I/O request Packet*)
- Diferencial é a melhoria ao suporte plug-and-play
 - Adoção da especificação ACPI (*Advanced Configuration and Power Interface*)
 - Definição do WDM (*Win32 Driver Model*)
 - Objetivo é portabilidade de drivers

Leituras complementares

- R. Oliveira, A. Carissimi, S. Toscani; Sistemas Operacionais. Editora Sagra-Luzzato, 2001.
 - Capítulo 10
- <http://www.winternals.com>
 - Site que fornece vários artigos e tutorias sobre a tecnologia *Windows*