

## Esame del Sistema Operativo Windows Parte 3 - Indice

1. Architettura del *file system* NT
2. *Master File Table*
  - 2.1 Descrittore base senza estensioni
  - 2.2. Descrizione con estensioni
3. Creazione e localizzazione di *file*
4. Gestione della *cache*

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 237

## Esame del Sistema Operativo Windows Architettura del file system NT - 1

- NT 5.x è capace di supportare l'intera gamma dei FS Windows ed anche **ext2fs** di Linux
  - FAT-16 : ampiezza logica di partizione limitata a  $2^{16}$  blocchi da 32 kB  $\rightarrow$  2 GB
  - FAT-32 : ampiezza fisica di partizione limitata a  $2^{32}$  settori da 512 B  $\rightarrow$  2 TB
  - NTFS : nuova concezione, con indirizzi su disco espressi su 64 *bit*

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 238

## Esame del Sistema Operativo Windows Architettura del file system NT - 2

- Nome di *file* fino a 255 caratteri in codifica Unicode (2 B/c)
  - Se espresso come cammino relativo o assoluto, fino a (32k - 1) caratteri
  - Distinzione tra maiuscolo e minuscolo, ma senza esito per buona parte di **win32 API**
- *File* come aggregato di attributi, ciascuno rappresentato come sequenza (*stream*) di caratteri
  - P.es.: sequenza breve contenente il nome del *file* e l'indirizzo dell'oggetto ad esso associato + sequenza lunga (fino a  $2^{64}$  B !) contenente i dati del *file*
  - Idea copiata da Apple Macintosh © ed introdotta (anche) per compatibilità

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 239

## Esame del Sistema Operativo Windows Architettura del file system NT - 3

- FS ad architettura gerarchica (come Linux)
  - \ invece di / come separatore nei cammini, sia assoluti che relativi
  - Supporto per *directory* corrente (**wd**)
  - Supporto per le entrambe le varietà di **link**
- Servizi di FS resi tramite procedure di libreria **Win32 API**
  - Funzionalmente simili a Linux, ma di concezione assai più bizantina

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 240

## Esame del Sistema Operativo Windows Architettura del file system NT - 4

- NTFS è una collezione di volumi logici
  - Un volume logico può mappare su più di una partizione di disco ed anche su più dischi
  - Il volume è visto come una sequenza lineare di blocchi (*cluster*) di ampiezza fissa
    - Volumi diversi possono avere diversa dimensione di blocco (tra 512 B a 64 kB)
      - Blocco piccolo  $\rightarrow$  bassa frammentazione interna
      - Blocco grande  $\rightarrow$  meno accessi a disco
- Una MFT (**Master File Table**) per volume
  - Fisicamente realizzata come un *file*
    - Perciò può essere salvata ovunque nel volume e non in posizione fissa (che può essere difettosa)
  - Logicamente strutturata come una sequenza lineare di  $\leq 2^{48}$  descrittori di *file* (o *directory*) di ampiezza 1 kB

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 241

## Esame del Sistema Operativo Windows Architettura del file system NT - 5

- Ciascun descrittore (*record*) contiene un numero variabile di coppie di tipo **<descrittore di attributo, valore>**
  - Il 1° campo specifica la struttura e la dimensione dell'attributo
    - 13 attributi sono richiesti dal sistema ed hanno struttura nota
    - Altri attributi, a struttura libera, possono essere aggiunti dall'utente
  - Il 2° campo contiene il valore effettivo dell'attributo (se di piccole dimensioni), altrimenti un puntatore alla sua locazione su disco ( $\rightarrow$  relativo descrittore in MFT)
    - Nell'un caso si parla di attributo residente, nell'altro di attributo non residente
    - Il contenuto effettivo del *file* è visto come il valore dell'attributo "dati"

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 242

### Esame del Sistema Operativo Windows Architettura del file system NT - 6

- I primi 16 descrittori dell'MFT sono riservati per *file* "trascendenti" di sistema (*metafile*), che descrivono la struttura effettiva del contenuto informativo del volume
  - Il 1° descrittore descrive l'MFT stesso
  - Il 2° replica i primi 16 descrittori in modo "non residente", ponendone il contenuto in fondo al volume
    - Per ripristino di FS in caso di corruzione dell'MFT
  - Il 4° descrive il volume (nome, versione di FS, data di creazione, etc.)
  - Il 5° descrive gli attributi usati nel volume
    - Per attributi non residenti si usa un puntatore alla loro descrizione (descrittore secondario in MFT, 48 bit) ed un codice identificatore su 16 bit replicato, per verifica di consistenza, in entrambi i descrittori
  - Seguono: puntatore alla radice del FS; *bitmap* dei blocchi liberi; copia del codice di *boot* di volume o suo puntatore; etc.

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 243

### Esame del Sistema Operativo Windows Architettura del file system NT - 7

- Il campo <descrittore di attributo> per attributi residenti ha ampiezza 24 B
  - Quello per attributi non residenti è più ampio
- Non tutti i 13 attributi di sistema applicano a tutti i *file* (o *directory*)
  - Gli attributi previsti per i *file* corrispondono a quelli che Linux pone negli *i-node*, con l'aggiunta dell'identificatore del corrispondente oggetto
    - 64 bit per identificatore unico per volume
  - Per *file* di ampiezza < 1 kB il loro contenuto viene memorizzato interamente entro un descrittore, altrimenti il valore dell'attributo diventa la lista dei relativi indirizzi di blocco

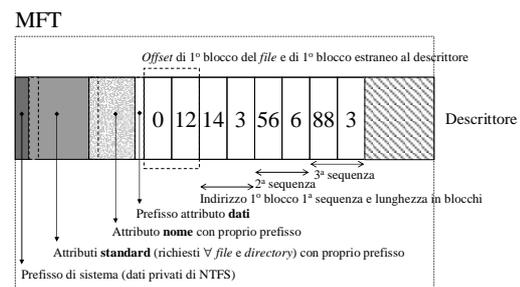
Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 244

### Esame del Sistema Operativo Windows Architettura del file system NT - 8

- Il valore di attributi non residenti viene allocato su sequenze, non necessariamente adiacenti, di blocchi contigui
  - L'allocatore tenta di assegnare sequenze di blocchi contigui piuttosto che singoli blocchi
    - Esattamente come per Linux (cf. pag. 179 delle dispense)
  - Nel caso peggiore si hanno sequenze di 1 solo blocco
- $\forall$  *file* sequenziale in FS  $\square$  1 descrittore base in MFT
  - La struttura interna del descrittore dipende dalla dimensione del *file* e dalla contiguità logica dei suoi blocchi
  - *File* con zone interne non utilizzate (e.g. poste a 0 e riservate per uso futuro) sono chiamati "sparsi" e sono trattati diversamente

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 245

### Esame del Sistema Operativo Windows Descrittore base senza estensioni - 1



Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 246

### Esame del Sistema Operativo Windows Descrittore base senza estensioni - 2

- Nell'esempio in figura un solo descrittore basta per contenere la lista di tutte le sequenze di blocchi di dati del *file*
  - 12 blocchi in totale, suddivisi in 3 sequenze, ciascuna descritta come
    - Indirizzo su disco del 1° blocco della sequenza
    - Ampiezza in blocchi della sequenza
  - Il prefisso dell'attributo dati specifica il numero di sequenze presenti nel descrittore
  - La prima coppia di attributi dati specifica la posizione *relativa* entro il *file* del 1° blocco coperto dal descrittore ed il suo confine (1° blocco non coperto)

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 247

### Esame del Sistema Operativo Windows Descrittore base senza estensioni - 3

- Questa strategia consente di rappresentare *file* di ampiezza illimitata
- Il numero di descrittori necessari per singolo *file* dipende più dalla contiguità dei blocchi che dall'ampiezza assoluta del *file*
  - 1 *file* da 20 GB costituito da 20 sequenze di 1 M blocchi da 1 kB ciascuno richiede 20+1 coppie di valori espressi su 64 bit  $\rightarrow$  ampiamente contenuto in 1 descrittore MFT da 1 kB
  - 1 *file* da 64 kB costituito da 60 sequenze di 1 blocco ciascuna richiede  $(64+1)*(2*8 B) > 1 kB$ , eccedendo la capacità di un singolo descrittore

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 248

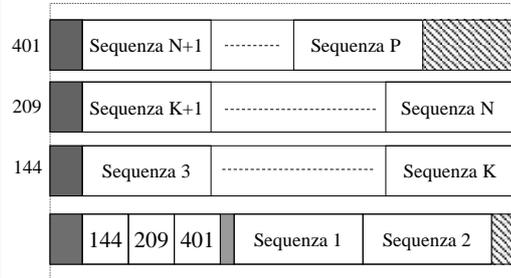
### Esame del Sistema Operativo Windows Descrizione con estensioni - 1

- La rappresentazione di alcuni *file* può talora richiedere più di un descrittore
- In tal caso, NTFS usa il meccanismo delle "continuazioni" (già visto per gli *i-node* di UNIX e Linux)
  - Il descrittore base contiene un puntatore ad uno o più descrittori secondari in MFT, dove si trova la rappresentazione dei blocchi del *file*
  - Se rimane spazio, il descrittore base può includere la rappresentazione delle prime sequenze di blocchi
- Se non vi fosse abbastanza spazio in MFT, l'intera lista di descrittori secondari del *file* verrebbe trattata come un attributo non residente e posta in un "*file*" dedicato, con un corrispondente descrittore in MFT

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 249

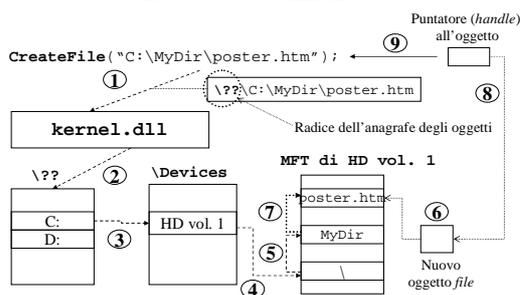
### Esame del Sistema Operativo Windows Descrizione con estensioni - 2

MFT



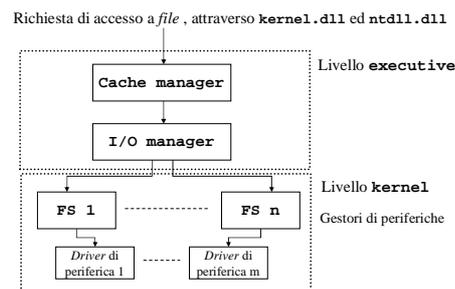
Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 250

### Esame del Sistema Operativo Windows Creazione e localizzazione di file



Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 251

### Esame del Sistema Operativo Windows Gestione della cache - 1



Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 252

### Esame del Sistema Operativo Windows Gestione della cache - 2

- Ogni accesso a *file* viene indirizzato in prima istanza al **cache manager**
  - **Indipendentemente** dal tipo di FS
  - Tratta blocchi **virtuali**, perché non conosce la struttura fisica dello specifico FS
    - Blocco virtuale = (*stream, offset*)
    - Blocco fisico = (partizione, indice di blocco)
- Ciascuno specifico FS è visto come un gestore di periferica logica controllato dall'**I/O manager**

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 253

### Esame del Sistema Operativo Windows Gestione della cache - 3

- Per ogni *file* in uso, il **cache manager** alloca 256 kB di spazio di indirizzamento virtuale di **kernel**
  - Indipendentemente dalla dimensione del *file*
  - Lo spazio complessivo a sua disposizione è parametro di configurazione
    - Se necessario, rialloca spazio assegnata a *file* vecchi
- Le richieste di accesso a *file* vengono soddisfatte attingendo allo spazio di **kernel**
  - Per ogni dato non disponibile, l'**I/O manager** tratta l'errore (*page fault*), **trasparentemente** al **cache manager**, caricando il blocco richiesto

Esame del Sistema Operativo Windows Architettura degli elaboratori 2 - T. Vardanega Pagina 254

## Esame del Sistema Operativo Windows *Gestione della cache - 4*

- Ogni accesso simultaneo da parte di più *thread* distinte ad uno stesso *file* mappato in memoria riferisce la stessa area assegnata dal **cache manager** al *file*
  - Il *file* è mappato una sola volta nello spazio del kernel, indipendentemente dal numero di utenti
  - Le scritture avvengono nella zona di nucleo, le letture copiano dati nell'area del richiedente
- Questo meccanismo garantisce la coerenza dei valori trattati