Caso di studio: Windows

Ivan Scagnetto

Università di Udine — Facoltà di Scienze MM.FF.NN.

A.A. 2006-2007

Windows NT, 2000, XP

- Storia
- Principi alla base della progettazione
- Componenti di sistema
- Sottosistemi d'ambiente
- File System
- Networking
- Interfaccia di programmazione

Windows NT, 2000, XP

- Sistema operativo a 32-bit con multitasking preemptive per microprocessori moderni.
- Caratteristiche chiave del sistema:
  - portabilità,
  - sicurezza,
  - conformità allo standard POSIX,
  - supporto per sistemi multiprocessore,
  - estendibilità,
  - supporto per l'internazionalizzazione,
  - compatibilità con le applicazioni sviluppate per MS-DOS e MS-Windows 3.x.
- Utilizza un'architettura a micro-kernel.
- Nel 1996, furono vendute più licenze di NT Server che licenze UNIX.

Storia

- Nel 1988, Microsoft decise di sviluppare un sistema operativo portabile basato su una nuova tecnologia (new technology: NT) che supportasse le API sia di OS/2 che dello standard POSIX.
- Nell'ottobre del 1988 Dave Cutler (progettista del sistema operativo DEC VAX/VMS) venne assunto con l'incarico di progettare e guidare lo sviluppo di NT.
- Nonostante in origine NT dovesse supportare l'API di OS/2 come ambiente nativo, durante lo sviluppo si decise di propendere per l'API di Win32, per sfruttare la popolarità di Windows 3.0.
- Le prime versioni del sistema operativo furono Windows NT 3.1 e Windows NT 3.1 Advanced Server.

Alcune differenze fra Windows 95/98 e Windows NT

Item	Windows 95/98	Windows NT
Full 32-bit system?	No	Yes
Security?	No	Yes
Protected file mappings?	No	Yes
Private addr space for each MS-DOS prog?	No	Yes
Unicode?	No	Yes
Runs on	Intel 80x86	80x86, Alpha, MIPS, ...
Multiprocessor support?	No	Yes
Re-entrant code inside OS?	No	Yes
Plug and play?	Yes	No
Power management?	Yes	No
FAT-32 file system?	Yes	Optional
NTFS file system	No	Yes
Win32 API?	Yes	Yes
Run all old MS-DOS programs?	Yes	No
Some critical OS data writable by user?	Yes	No

Confronto fra le dimensioni di alcuni S.O.

Year	AT&T	BSD	MINIX	Linux	Solaris	Win NT
1976	V6 9K					
1979	V7 21K					
1980		4.1 38K				
1982	Sys III 58K					
1984		4.2 98K				
1986		4.3 179K				
1987	SVR3 92K		1.0 13K			
1989	SVR4 280K					
1991				0.01 10K		
1993		Free 1.0 235K			5.3 850K	3.1 6M
1994		4.4 Lite 743K		1.0 165K		3.5 10M
1996				2.0 470K		4.0 16M
1997			2.0 62K		5.6 1.4M	
1999				2.2 1M		
2000		Free 4.0 1.4M			5.8 2.0M	2000 29M

Storia

- Nella versione 4.0 di NT venne adottata l'interfaccia utente di Windows 95, il server web (IIS: Internet Information Server) ed il browser Internet Explorer.
- L'interfaccia utente e le librerie grafiche vennero spostate nel kernel per aumentare le prestazioni, ma resero anche il sistema meno stabile.
- Le prime versioni di NT vennero adattate anche per altre architetture (processori Alpha), ma a partire da Windows 2000 la portabilità venne limitata ai soli processori Intel e compatibili per ragioni di mercato.
- Windows 2000 introdusse notevoli migliorie:
  - Active Directory: servizio di directory basato sullo standard X.500;
  - gestione migliorata dei sistemi di rete;
  - supporto ai dispositivi plug-and-play;
  - file system distribuito;
  - possibilità di aumentare il numero di CPU e la capacità di memoria.

Versioni di Windows 2000

Version	Max RAM	CPUs	Max clients	Cluster size	Optimized for
Professional	4 GB	2	10	0	Response time
Server	4 GB	4	Unlimited	0	Throughput
Advanced server	8 GB	8	Unlimited	2	Throughput
Datacenter server	64 GB	32	Unlimited	4	Throughput

Storia

- Nell'ottobre 2001 l'uscita di Windows XP sostituì Windows 2000 e Windows 95/98, migliorandone le prestazioni, introducendo un supporto multimediale, applicazioni digitali (fotografiche e video) ed un aspetto grafico migliorato.
- Nel 2002 uscì Windows .NET Server (la versione server di XP).
- Windows XP è dotato di un'architettura client-server per implementare diverse personalità (es.: API Win32 e POSIX) tramite processi a livello utente (sottosistemi).
- Windows XP nelle versioni server consente accessi simultanei (tramite servizi distribuiti o istanze multiple dell'interfaccia grafica gestite dal terminal server).
- Nelle versioni desktop invece è possibile avere sessioni virtuali per ogni utente collegato grazie alla tecnica del multiplex applicata a tastiera, mouse e monitor.

Storia

- Windows XP è la prima versione a 64 bit di Windows.
- Il filesystem (NTFS) e molte delle API Win32 usavano già da tempo interi a 64 bit.
- L'estensione di Windows a 64 bit riguarda principalmente il supporto di indirizzi estesi.
- Nelle versioni server per centri di elaborazione dati può gestire
  - fino a 64 GB di memoria e 32 processori (IA32),
  - fino a 128 GB di memoria e 64 processori (IA64).

Princìpi di progettazione

- Estendibilità — architettura a strati.
  - Il modulo esecutivo, che viene eseguito in protected mode, fornisce i servizi di base del sistema.
  - Appoggiandosi sull'esecutivo, diversi sottosistemi operano in modalità utente.
  - Questa struttura modulare consente di aggiungere ulteriori sottosistemi d'ambiente senza influenzare il modulo esecutivo.
- Portabilità — il sistema può essere installato su un'architettura hardware differente con un numero contenuto di modifiche:
  - è scritto in C e C++;
  - il codice dipendente dal tipo di hardware è isolato in una dynamic link library (DLL) denominata Hardware Abstraction Layer (HAL).

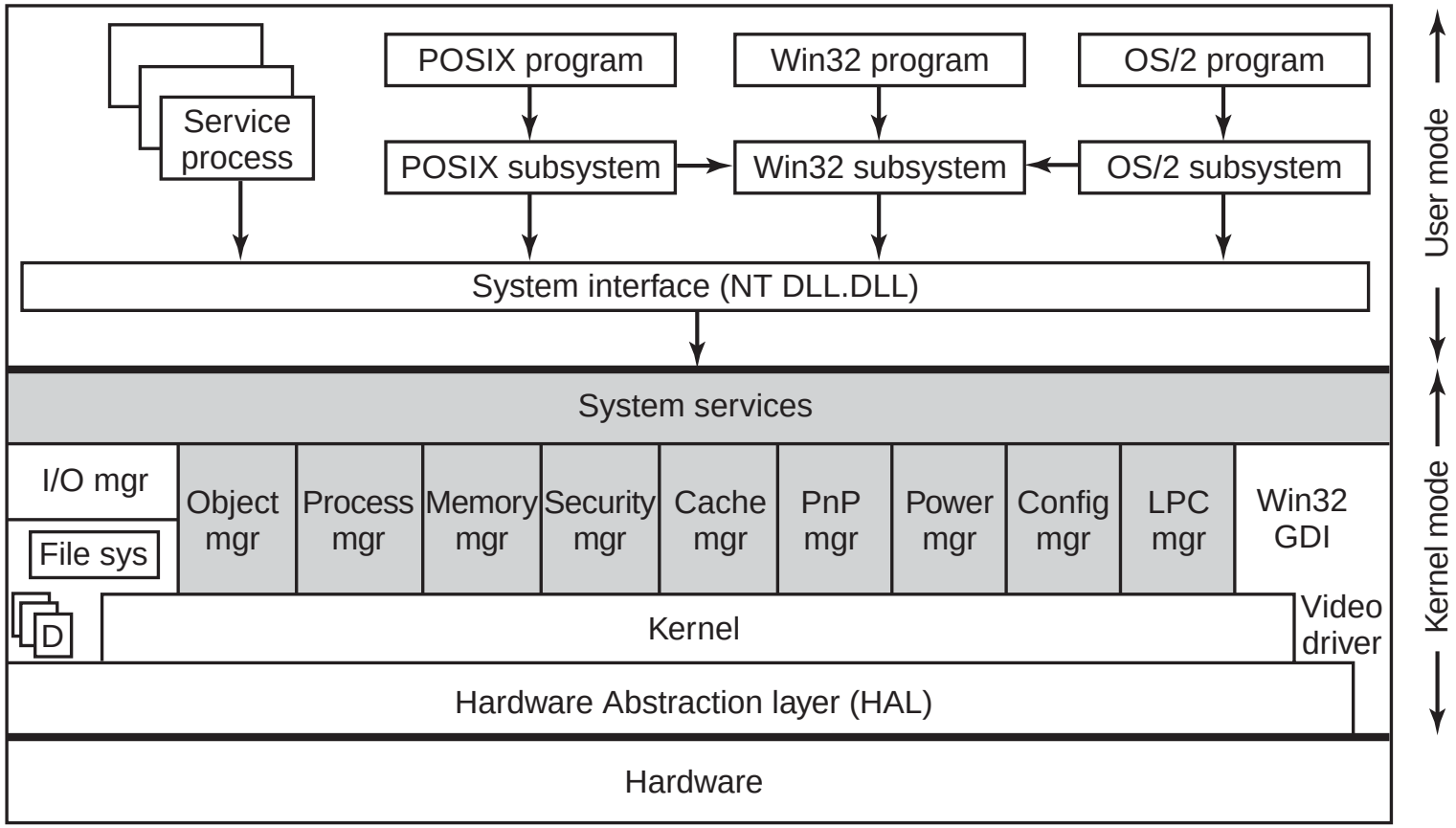
Princìpi di progettazione

- Affidabilità — il sistema utilizza una protezione a livello hardware per la memoria virtuale e dei meccanismi di protezione software per quanto riguarda le risorse del sistema operativo.
- Compatibilità — applicazioni che aderiscono allo standard IEEE 1003.1 (POSIX) possono essere compilate per girare sul sistema senza modifiche al codice sorgente.
- Prestazioni — i diversi sottosistemi possono comunicare fra loro per mezzo di un meccanismo di scambio messaggi molto efficiente.
  - La prelazione di thread a bassa priorità permette al sistema di reagire velocemente agli eventi esterni.
  - Il sistema è progettato per supportare efficacemente multiprocessori simmetrici.
- Supporto per l'internazionalizzazione — l'API national language support (NLS) consente di adeguare il sistema a diverse culture/lingue.

Architettura del sistema

- Consiste in un insieme stratificato di moduli.
- Protected mode — HAL, kernel, modulo esecutivo.
- User mode — collezione di sottosistemi.
  - I sottosistemi d'ambiente simulano diversi sistemi operativi.
  - I sottosistemi di protezione si occupano di garantire la sicurezza.

Architettura del sistema



Componenti di sistema — Kernel

- Il kernel fornisce le basi per il modulo esecutivo ed i vari sottosistemi.
- Non viene mai paginato su disco; la sua esecuzione non è mai soggetta a prelazione.
- Ha quattro compiti principali da svolgere:
  1. scheduling dei thread,
  2. gestione degli interrupt e delle eccezioni,
  3. sincronizzazione a basso livello dei processori,
  4. ripristino del sistema dopo una caduta di tensione.
- Il kernel è object-oriented ed utilizza due insiemi di oggetti:
  1. oggetti dispatcher: controllano consegna (dispatching) e sincronizzazione (eventi, mutanti, mutex, semafori, thread, timer).
  2. oggetti di controllo: (APC: asynchronous procedure call, DPC: deferred procedure call, interrupt, notifiche relative allo stato dell'alimentazione, oggetti relativi a processi e profili).

Kernel — Processi e Thread

- Ogni processo ha uno spazio di indirizzi di memoria virtuale, delle informazioni associate (es.: priorità di base) ed un grado di affinità per uno o più processori.
- I thread rappresentano l'unità di esecuzione gestita dal dispatcher del kernel.
- Ogni thread ha un proprio stato che include una priorità, un grado di affinità del processore e dell'informazione di accounting.
- Un thread può trovarsi in uno fra sei stati: ready, standby, in esecuzione, in attesa, in transizione e terminato.

Kernel — Scheduling

- Il dispatcher utilizza uno schema a 32 livelli di priorità per determinare l'ordine di esecuzione dei thread. Le priorità sono suddivise in due classi.
  - La classe real-time contiene i thread con intervallo di priorità da 16 a 31.
  - La classe a priorità variabile contiene i thread con intervallo di priorità da 0 a 15.
- Caratteristiche dell'algoritmo di scheduling:
  - tende a favorire i thread interattivi (che utilizzano dispositivo di puntamento e finestre) con un tempo di risposta molto buono;
  - permette ai thread I/O-bound di mantenere i dispositivi di I/O occupati.
  - I thread CPU-bound utilizzano in background i cicli di CPU disponibili.

Kernel — Scheduling

- Lo scheduling può avvenire quando un thread si trova negli stati ready o in attesa, quando termina, oppure quando un'applicazione cambia la priorità oppure il grado di affinità rispetto ad un processore di un thread.
- I thread real-time hanno un accesso preferenziale alla CPU, ma il sistema non garantisce che un thread real-time verrà eseguito entro dei limiti temporali precisi.

Kernel — Gestione delle Trap

- Il kernel permette di gestire le trap in occasione delle eccezioni e degli interrupt generati dal software o dall'hardware.
- Le eccezioni che non possono essere gestite dal trap handler sono prese in carico dall'exception dispatcher del kernel.
- L'interrupt dispatcher del kernel gestisce gli interrupt richiamando un'apposita routine di servizio (come avviene per un device driver) oppure una routine interna del kernel stesso.
- Il kernel implementa la mutua esclusione in un sistema multiprocessore utilizzando degli spin lock che risiedono nella memoria globale.

Modulo esecutivo — Object Manager

- Il sistema operativo utilizza gli oggetti per rappresentare tutti i suoi servizi ed entità in gioco; l'object manager supervisiona l'uso degli oggetti:
  - genera un object handle,
  - effettua dei controlli di sicurezza,
  - per ogni oggetto tiene traccia dei processi che lo stanno utilizzando.
- Gli oggetti sono manipolati da un insieme di metodi standard: `create`, `open`, `close`, `delete`, `query_name`, `parse` e `security`.

Modulo esecutivo — Denominazione degli oggetti

- Il modulo esecutivo consente di assegnare un nome (permanente o temporaneo) ad ogni oggetto.
- I nomi degli oggetti sono strutturati in modo analogo ai percorsi nel filesystem di sistemi operativi come MS-DOS o UNIX.
- Viene implementato anche un symbolic link object, simile ai symbolic link di UNIX, che consente di riferirsi allo stesso oggetto tramite nomi (alias) diversi.
- Un processo ottiene una object handle creando un oggetto, aprendone uno esistente, ricevendo una handle duplicata da un altro processo oppure ereditandola da un processo genitore.
- Ogni oggetto del sistema è protetto da una access control list (ACL).

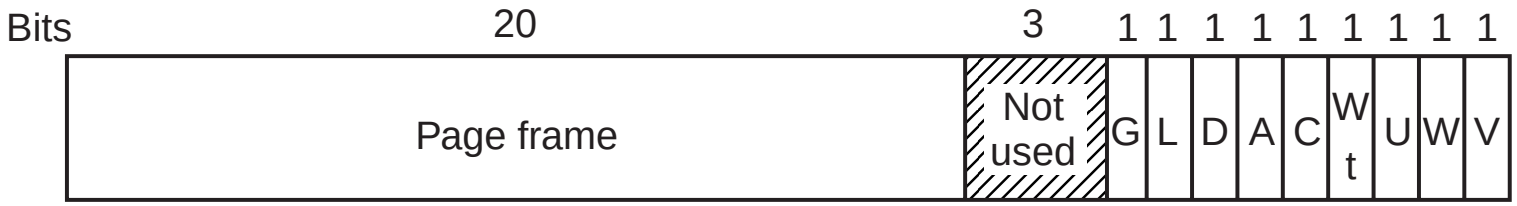
Modulo esecutivo — Manager della memoria virtuale

- Il design del manager della memoria virtuale assume che l'hardware sottostante supporti una traduzione da indirizzo virtuale a fisico, un meccanismo di paginazione, una cache coerente e trasparente sui sistemi multiprocessore, ed aliasing degli indirizzi virtuali.
- Il manager della memoria virtuale usa un meccanismo di paginazione con una dimensione di pagina di 4 KB.
- Per l'allocazione della memoria viene utilizzato un processo a due fasi:
  1. la prima fase riserva una parte dello spazio di indirizzi del processo;
  2. la seconda fase compie effettivamente l'allocazione assegnando dello spazio nel file di paginazione.

Modulo esecutivo — Manager della memoria virtuale

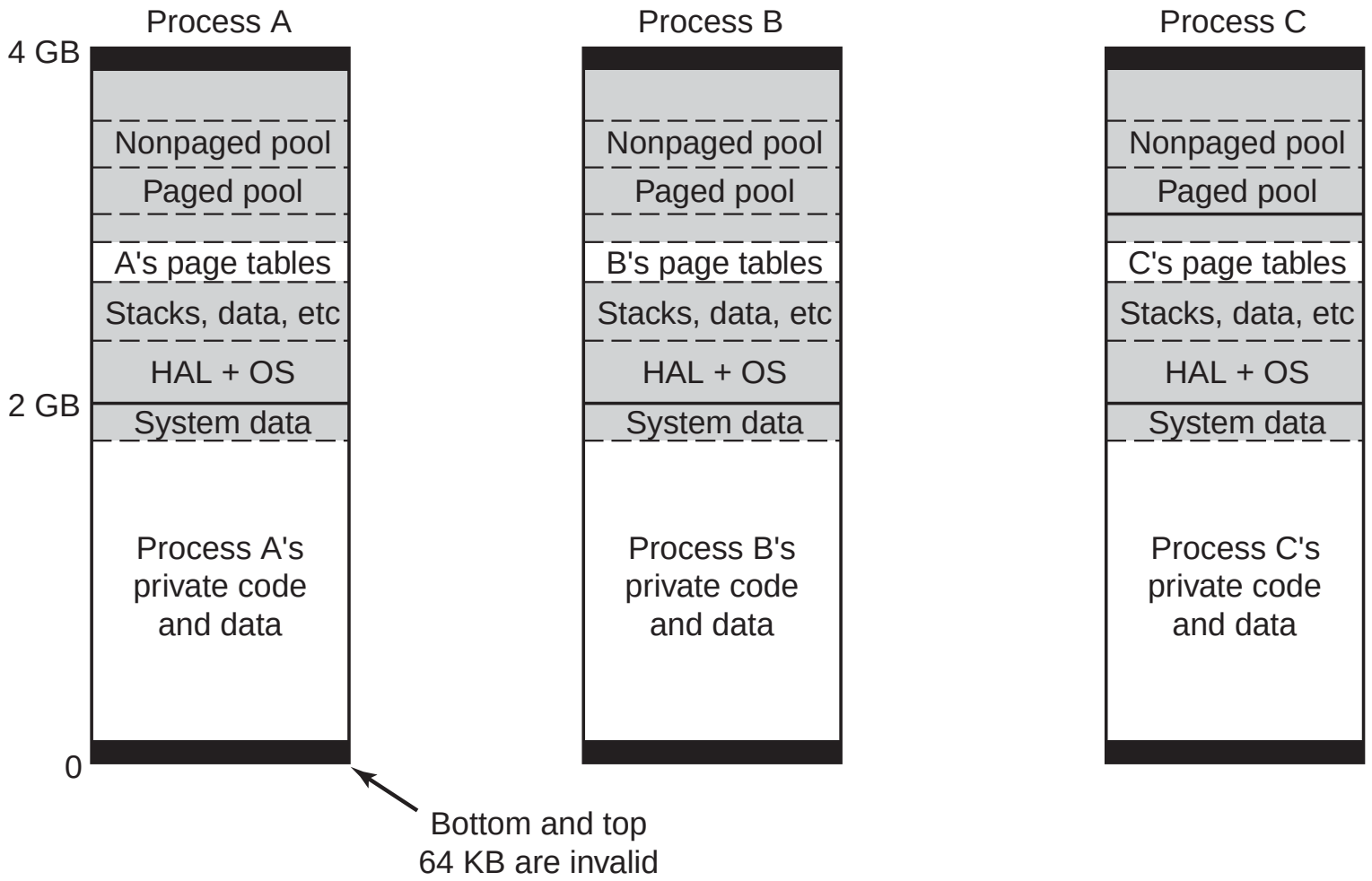
- La traduzione dell'indirizzo virtuale utilizza diverse strutture dati:
  - ogni processo ha una page directory che contiene 1024 page directory entry di dimensione 4 byte.
  - Ogni page directory entry punta ad una page table contenente 1024 page table entry (PTE) di dimensione 4 byte.
  - Ogni PTE punta ad un page frame di 4 KB nella memoria fisica.
- Un intero a 10-bit può rappresentare tutti i valori da 0 a 1023; quindi può essere usato per selezionare un'entry nella page directory o in una page table.
- Questa proprietà è utilizzata nel tradurre un puntatore ad un indirizzo virtuale ad un indirizzo nella memoria fisica.
- Una pagina può ritrovarsi in uno fra sette stati diversi: valida, azzerata, libera, in attesa, modificata, guasta ed in transizione.

La struttura di una PTE (Pentium)



- G: Page is global to all processes
- L: Large (4-MB) page
- D: Page is dirty
- A: Page has been accessed
- C: Caching enabled/disabled
- Wt: Write through (no caching)
- U: Page is accessible in user mode
- W: Writing to the page permitted
- V: Valid page table entry

Esempio: spazi di indirizzamento virtuali di tre processi



Gestione della memoria

Uno spazio di indirizzamento viene descritto per mezzo dei VAD (Virtual Address Descriptor) che tengono traccia di:

- intervallo degli indirizzi mappati,
- il file di backing store utilizzato e l'offset all'interno di quest'ultimo,
- i codici di protezione.

La page table viene creata al primo riferimento ad una pagina del VAD (che memorizza quindi un puntatore alla tabella).
I page fault possono avvenire nelle seguenti occasioni:

- La pagina riferita non è assegnata (committed).
- Si è verificato un errore di violazione della protezione.
- Una pagina condivisa è stata acceduta in scrittura.
- Lo stack ha necessità di espandersi.
- La pagina riferita è assegnata ma non è presente in memoria.

Gestione della memoria

- Il sistema cerca di mantenere un certo numero di pagine libere.
- Ogni processo ha un working set con un numero di pagine regolato da un limite minimo e da un limite massimo.
- Se avviene un page fault e ci si trova al di sotto del limite minimo un nuovo frame viene allocato al processo. Se ci si trova al di sopra del limite massimo viene liberata una pagina per far posto alla nuova.
- Il balance set manager controlla periodicamente se ci sono abbastanza pagine libere.
- Se così non è, interviene il working set manager che provvede ad esaminare i processi in esecuzione alla ricerca di pagine da liberare.

Modulo esecutivo — Manager dei processi

- Mette a disposizione i servizi per la creazione, cancellazione, e l'utilizzo di thread e processi.
- La gestione delle problematiche come le relazioni padre/figlio e le gerarchie di processi è lasciata allo specifico sottosistema a cui appartiene il processo in questione.

Modulo esecutivo — Local Procedure Call (LPC)

- Il meccanismo delle LPC consente di scambiare richieste e risultati fra processi client e server all'interno di un singolo sistema.
- In particolare tale meccanismo viene utilizzato per richiedere servizi ai vari sottosistemi.
- Quando viene creato un canale LPC, possono essere utilizzate tre tecniche principali di scambio di messaggi:
  1. il primo tipo è particolarmente adatto per messaggi di piccole dimensioni, fino a 256 byte; la coda di messaggi della porta viene utilizzata come area di memorizzazione temporanea, mentre i messaggi vengono copiati da un processo all'altro;
  2. il secondo tipo evita di copiare messaggi di grandi dimensioni, puntando ad un oggetto di sessione creato per il canale nella memoria condivisa;
  3. il terzo metodo, denominato quick LPC, viene utilizzato da parti del sottosistema grafico di Win32.

Modulo esecutivo — Manager dell'I/O

- Il gestore dell'I/O è il responsabile dei seguenti aspetti:
  - file system,
  - gestione della cache,
  - device driver,
  - network driver.
- Tiene traccia di quali tra i file system installabili sono caricati, e gestisce i buffer per le richieste di I/O.
- Collabora con il gestore della memoria virtuale per consentire il memory-mapped I/O.
- Controlla il gestore della cache, che gestisce il caching per l'intero sistema di I/O.
- Supporta sia operazioni sincrone che asincrone, stabilendo i time-out per i driver ed i meccanismi per richiamare un driver all'interno di un altro driver.

Modulo esecutivo — Security Reference Manager

- La natura orientata agli oggetti del sistema consente l'utilizzo di un meccanismo uniforme per realizzare il controllo dell'accesso a tempo di esecuzione ed i controlli di audit per ogni entità del sistema.
- Quando un processo ottiene un handle ad un oggetto, il controllore della sicurezza dei riferimenti controlla il token di sicurezza del processo e l'ACL dell'oggetto per verificare che il processo abbia i privilegi necessari.

Sottosistemi d'ambiente

- I processi in modalità utente si appoggiano sui servizi del modulo esecutivo per permettere al sistema di far girare i programmi scritti per altri sistemi operativi.
- Il sistema utilizza il sottosistema Win32 come ambiente fondamentale; Win32 viene utilizzato per mandare in esecuzione tutti i processi. Fornisce inoltre tutti i servizi relativi a tastiera, mouse e capacità grafiche.
- L'ambiente MS-DOS viene implementato come un'applicazione Win32 denominata virtual dos machine (VDM), un processo in modalità utente che viene paginato e gestito dal dispatcher come un qualsiasi altro thread del sistema.

Sottosistemi d'ambiente

- Ambiente Windows a 16-Bit:
  - realizzato tramite una VDM che incorpora Windows su Windows;
  - fornisce le routine del kernel di Windows 3.1 kernel e le stub routine per il window manager e le funzioni GDI.
- Il sottosistema POSIX è stato progettato per mandare in esecuzione applicazioni POSIX conformi allo standard POSIX.1 che si basa sul modello di UNIX.

File System

- La struttura fondamentale del file system (NTFS) è il volume:
  - viene creato tramite il programma di utilità di amministrazione dei dischi;
  - si basa su una partizione logica del disco;
  - può occupare una parte del disco, un disco intero oppure estendersi su più dischi.
- Tutti i metadati, come le informazioni sul volume, vengono memorizzati in un file regolare.
- NTFS utilizza i cluster come unità di allocazione del disco:
  - un cluster è un numero di settori del disco esprimibile come potenza di due;
  - siccome la dimensione del cluster è più piccola della controparte a 16-bit di un file system basato su FAT, l'ammontare della frammentazione interna è ridotto.

File System — Layout interno

- NTFS utilizza numeri di cluster logici (Logical Cluster Number, LCN) come indirizzi di disco.
- Un file in NTFS non è un semplice stream di byte, come in MS-DOS o UNIX; invece si presenta come un oggetto strutturato costituito da attributi.
- Ogni file in NTFS è descritto da uno o più record in un vettore memorizzato in un file speciale denominato Master File Table (MFT).
- Ogni file su un volume NTFS ha un unico ID chiamato riferimento al file:
  - numero a 64-bit che consiste di un numero di file a 16-bit e di un numero di sequenza a 48-bit;
  - può essere utilizzato per effettuare dei controlli di consistenza.
- Lo spazio dei nomi di NTFS è organizzato come una gerarchia di directory; la radice contiene il livello top del B+tree corrispondente.

File System — Ripristino

- Tutti gli aggiornamenti delle strutture dati del file system vengono effettuate utilizzando delle transazioni:
  - prima di alterare la struttura dati, la transizione scrive un record di log contenente le informazioni per il ripristino e l'annullamento dell'operazione;
  - dopo la modifica alla struttura dati, un record di commit viene aggiunto al log per denotare il successo della transazione;
  - in seguito ad un crash, le strutture del file system possono essere ripristinate ad uno stato consistente analizzando i record del log.

File System — Ripristino

- Lo schema analizzato non garantisce che tutti i dati nei file degli utenti potranno essere ripristinati dopo un crash. Soltanto le strutture dati del file system (i metadati) risulteranno integri e rifletteranno uno stato consistente prima del crash.
- Il log viene memorizzato nel terzo file di metadati all'inizio del volume.
- La funzionalità di logging è fornita dal servizio del file di log.

File System — Sicurezza

- La sicurezza di un volume NTFS deriva direttamente dal modello ad oggetti del sistema.
- Ogni oggetto file ha un attributo descrittore della sicurezza memorizzato nel suo record MFT.
- Tale attributo contiene il token d'accesso del proprietatio del file ed un'ACL che stabilisce i privilegi d'accesso che vengono concessi agli utenti che possono accedere al file.

Gestione dei volumi e tolleranza agli errori

- `FtDisk`, il driver dei dischi fault tolerant del sistema, fornisce diversi modi per combinare più dischi SCSI in un unico volume logico.
- Un volume set è una concatenazione logica di più dischi per formare un volume logico di dimensioni maggiori dei singoli componenti.
- È anche possibile combinare più partizioni fisiche in modalità round-robin per formare un insieme di stripe (ovvero, un RAID di livello 0).
  - Variante: insieme di stripe con parità, ovvero, RAID di livello 5.
- Il sistema supporta anche il disk mirroring o RAID di livello 1: è costituito da due partizioni di uguale dimensione che si trovano su due dischi diversi e che contengono gli stessi dati.
- `FtDisk`, per gestire i settori guasti, utilizza una tecnica hardware denominata sector sparing, mentre NTFS utilizza una tecnica software denominata cluster remapping.

File System — Compressione

- Per comprimere un file, NTFS suddivide i dati del file in unità di compressione, che consistono in blocchi di 16 cluster contigui.
- Per file radi, NTFS utilizza un'altra tecnica per risparmiare spazio:
  - i cluster che contengono solo zeri non vengono allocati/salvati su disco;
  - invece i buchi vengono preservati nella sequenza di numeri di cluster virtuali memorizzata nell'entry MFT del file;
  - al momento della lettura di un file, se viene individuato un buco fra i numeri di cluster virtuali, NTFS riempie di zeri la porzione corrispondente del buffer del chiamante.

Networking

- Il sistema supporta il networking sia peer-to-peer che client/server; fornisce anche degli strumenti per la gestione della rete.
- Per descrivere il networking del sistema, ci riferiamo a due delle interfacce interne di rete:
  - NDIS (Network Device Interface Specification) — Separa gli adattatori di rete dai protocolli di trasporto in modo che i cambiamenti apportati sugli uni non influenzino gli altri.
  - TDI (Transport Driver Interface) — Abilita qualunque componente dello strato di sessione all'utilizzo di un meccanismo di trasporto disponibile.
- Il sistema implementa i protocolli di trasporto come driver che possono essere caricati e scaricati dinamicamente.

- Il protocollo server message block (SMB) viene utilizzato per inviare richieste di I/O sulla rete. Prevede quattro tipi di messaggi:
  - `controllo di sessione`,
  - `file`,
  - `stampante`,
  - `messaggio`.
- Il Network Basic Input/Output System (NetBIOS) è un'interfaccia di rete che astrae dall'hardware. Viene utilizzato per:
  - stabilisce nomi logici per la rete,
  - stabilisce connessioni logiche o sessioni fra due nomi logici in rete,
  - supporta trasferimenti di dati affidabili per una sessione basata su richieste NetBIOS o SMB.

Networking — Protocolli

- NetBEUI (NetBIOS Extended User Interface): protocollo di default per il peer networking su Windows 95 e Windows for Workgroups; viene utilizzato quando NT, 2000 o XP vuole condividere delle risorse con queste reti.
- I protocolli TCP/IP vengono utilizzati normalmente per connettersi ad una vasta gamma di sistemi operativi e piattaforme hardware.
- PPTP (Point-to-Point Tunneling Protocol) viene utilizzato per consentire la comunicazione fra moduli server per l'accesso remoto su macchine connesse tramite Internet.
- Il protocollo NWLink connette il NetBIOS alle reti Novell NetWare.

Networking — Protocolli

- Il protocollo Data Link Control (DLC) è utilizzato per accedere ai mainframe IBM ed alle stampanti HP che sono connesse direttamente alla rete.
- La comunicazione con calcolatori Macintosh tramite il protocollo Appletalk è possibile per mezzo del package Services for Macintosh.

Networking — Meccanismi per il calcolo distribuito

- La applicazioni distribuite sono supportate tramite named NetBIOS, named pipes e mailslot, Windows Socket, Remote Procedure Call (RPC) e Network Dynamic Data Exchange (NetDDE).
- Le applicazioni basate su NetBIOS possono comunicare in rete utilizzando NetBEUI, NWLink o TCP/IP.
- Le named pipe sono un meccanismo di scambio messaggi orientato alla connessione in cui i nomi vengono assegnati tramite la uniform naming convention (UNC).
- Le mailslot sono un meccanismo di scambio messaggi non orientato alla connessione sono utilizzate per applicazioni con comunicazioni di tipo broadcast, come per esempio per ricercare dei componenti sulla rete.
- Winsock, l'API delle windows socket, è un'interfaccia dello strato di sessione che fornisce un'interfaccia standard a molti protocolli di trasporto che possono avere degli schemi di indirizzamento differenti.

Meccanismi per il calcolo distribuito

- Il meccanismo RPC segue lo standard denominato Distributed Computing Environment per i messaggi RPC in modo che i programmi del sistema risultino portabili.
  - I messaggi RPC vengono inviati utilizzando NetBIOS o Winsock su reti TCP/IP oppure named pipe su reti Lan Manager.
  - Il Microsoft Interface Definition Language permette di descrivere i nomi delle procedure remote, gli argomenti ed i risultati.

Networking — Redirector e Server

- Un'applicazione può utilizzare l'API per l'I/O per accedere a file su un computer remoto, come se fossero locali, a patto che sul sistema remoto stia girando un server MS-NET.
- Un redirector è l'oggetto lato client che inoltra le richieste di I/O relative ai file remoti al server che è in grado di soddisfarle.
- Per ragioni di performance e sicurezza, i redirector ed i server vengono eseguiti in kernel mode.

Accesso ad un file remoto

- L'applicazione richiama il gestore dell'I/O per richiedere l'apertura di un file (assumendo che il nome del file segua il formato standard UNC).
- Il gestore dell'I/O forma un pacchetto di richiesta di I/O.
- Il gestore dell'I/O identifica che l'accesso è relativo ad un file remoto e richiama un driver denominato Multiple Universal Naming Convention Provider (MUP).
- Il driver MUP invia il pacchetto di richiesta di I/O in modo asincrono a tutti i redirector registrati.
- Un redirector in grado di soddisfare la richiesta risponde al MUP.
  - Per evitare di ripetere la stessa richiesta a tutti i redirector in futuro, i MUP utilizzano una cache per memorizzare quale fra i redirector può soddisfare la richiesta.

Accesso ad un file remoto

- Il redirector invia la richiesta di rete al sistema remoto.
- I driver di rete del sistema remoto ricevono la richiesta e la passano al driver del server.
- Il driver del server inoltra la richiesta al driver di sistema appropriato per il file system locale.
- Il device driver appropriato viene richiamato per recuperare i dati.
- I risultati vengono restituiti al driver del server, che li invia al redirector che aveva effettuato la richiesta.

Networking — Domìni

- Il sistema utilizza il concetto di dominio per gestire i diritti d'accesso globali all'interno dei gruppi.
- Un dominio è un gruppo di macchine su cui gira un sistema NT, 2000, XP di tipo Server che condividono una politica di sicurezza ed un database di utenti.
- Il sistema fornisce quattro modelli di dominio per gestire domìni multipli all'interno di una singola organizzazione:
  - single domain model: i domìni sono isolati;
  - master domain model: uno dei domìni viene eletto come master domain;
  - multiple master domain model: ci sono più domìni master che si considerano fidati (trusted) l'uno con l'altro;
  - multiple trust model: non c'è un master domain; ogni dominio gestisce i suoi utenti, ma i vari domìni si considerano fidati l'uno con l'altro.

Risoluzione dei nomi in reti TCP/IP

- Su una rete IP, la risoluzione dei nomi rappresenta il processo di conversione di un nome simbolico in un indirizzo IP, ad esempio, `www.bell-labs.com` viene risolto in `135.104.1.14`.
- Vengono forniti diversi metodi per la risoluzione dei nomi:
  - Windows Internet Name Service (WINS)
  - broadcast name resolution
  - domain name system (DNS)
  - hosts file
  - LMHOSTS file

Risoluzione dei nomi

- WINS consiste in uno o più server WINS che mantengono un database dinamico di associazioni fra nomi e indirizzi IP e del software client per interrogare i server.
- WINS utilizza il Dynamic Host Configuration Protocol (DHCP), che aggiorna automaticamente le configurazioni di indirizzi nel database WINS, senza l'intervento esplicito dell'utente o dell'amministratore.

Interfaccia di programmazione — Accesso agli oggetti del kernel

- Un processo ottiene l'accesso ad un oggetto del kernel denominato `XXX` richiamando la funzione `CreateXXX` per ottenere un handle relativa a `XXX`; tale handle è unica all'interno del processo.
- Un handle può essere chiusa richiamando la funzione `CloseHandle`; il sistema può rimuovere l'oggetto se il conteggio dei processi che lo utilizzano raggiunge il valore 0.
- Vi sono tre modi per condividere degli oggetti fra processi:
  - un processo figlio eredita un handle relativa all'oggetto;
  - un processo assegna un nome ad un oggetto al momento della sua creazione ed un secondo progetto apre l'oggetto con quel nome;
  - utilizzare la funzione `DuplicateHandle`:
    - assegnato un handle ad un processo ed il valore dell'handle, un secondo processo può recuperare un handle allo stesso oggetto, condividendolo.

Interfaccia di programmazione — Gestione dei processi

- Un processo viene lanciato tramite la funzione `CreateProcess` che carica tutte le DLL usate dal processo stesso e crea un thread primario.
- Thread addizionali possono essere creati per mezzo della funzione `CreateThread`.
- Ogni DLL o file eseguibile che viene caricato nello spazio di indirizzi di un processo viene identificato da un instance handle.

Gestione dei processi

- Lo scheduling in Win32 utilizza quattro classi di priorità:
  - `IDLE_PRIORITY_CLASS` (livello di priorità 4)
  - `NORMAL_PRIORITY_CLASS` (livello 8) – tipico per la maggior parte dei processi
  - `HIGH_PRIORITY_CLASS` (livello 13)
  - `REALTIME_PRIORITY_CLASS` (livello 24)
- Per garantire prestazioni adeguate per i programmi interattivi, per i processi nella classe `NORMAL_PRIORITY_CLASS` viene utilizzata una particolare regola:
  - si distingue fra il processo in primo piano che viene selezionato sullo schermo ed i processi in secondo piano che non sono selezionati al momento;
  - quando un processo passa in primo piano, viene incrementato il quanto per lo scheduling di un certo fattore (tipicamente 3).

Gestione dei processi

- Il kernel aggiorna dinamicamente la priorità di un thread in base al fatto che sia I/O-bound o CPU-bound.
- Per sincronizzare l'accesso concorrente agli oggetti condivisi dai thread, il kernel fornisce degli oggetti di sincronizzazione, come semafori e mutex.
  - Inoltre i thread possono sincronizzarsi utilizzando le funzioni `WaitForSingleObject` o `WaitForMultipleObjects`.
  - Un altro metodo di sincronizzazione nell'API di Win32 è la sezione critica.

Gestione dei processi

- Una fibra è del codice in modalità utente che viene schedulato in base ad un algoritmo definito dall'utente:
  - soltanto una fibra in ogni dato istante può essere in esecuzione, anche su hardware multiprocessore;
  - le fibre sono utilizzare per facilitare il porting di applicazioni UNIX di tipo legacy che sono state scritte secondo il modello di esecuzione con fibre.

Interfaccia di programmazione — Interprocess Communication

- Le applicazioni Win32 possono comunicare fra loro condividendo degli oggetti del kernel.
- Un altro meccanismo consiste nello scambio di messaggi, che è particolarmente utilizzato nella applicazioni per la GUI:
  - un thread invia un messaggio ad un altro thread oppure ad una finestra;
  - un thread può anche inviare dei dati nel messaggio.
- Ogni thread Win32 ha la sua coda di input in cui può ricevere i messaggi.
- Questa soluzione è sensibilmente migliore della coda condivisa della versione a 16-bit di Windows, dato che in questo caso un'applicazione in crash può bloccare l'input anche alle altre applicazioni.

Interfaccia di programmazione — Gestione della memoria

- Memoria virtuale:
  - `VirtualAlloc` riserva o blocca della memoria virtuale.
  - `VirtualFree` sblocca o rilascia la memoria.
  - Queste funzioni permettono all'applicazione di determinare l'indirizzo virtuale in corrispondenza del quale viene allocata la memoria.
- Un'applicazione può utilizzare la memoria eseguendo il memory mapping di un file nel proprio spazio di indirizzi.
  - Multistage process.
  - Due processi possono condividere la memoria mappando lo stesso file nelle rispettive memorie virtuali.

Gestione della memoria

- Uno heap nell'ambiente Win32 è una regione di uno spazio di indirizzi riservati.
  - Un processo Win32 viene creato con 1 MB di default heap.
  - L'accesso è sincronizzato per proteggere le strutture dei dati per l'allocazione dello spazio da danneggiamenti dovuti ad aggiornamenti concorrenti da parte di thread multipli.
- Siccome le funzioni che si basano su dati globali o statici tipicamente non riescono a girare correttamente in un ambiente multithread, il meccanismo di gestione della memoria locale alloca spazi globali per i singoli thread.
  - Il meccanismo fornisce metodi sia statici che dinamici per creare delle aree di memoria locali ad un thread.

Il registro di Windows

- Windows tiene traccia di un gran numero di informazioni cruciali sull'hardware, sul software e sugli utenti del sistema in un database centrale: il registro.
- Il registro è costituito da una collezione di directory contenenti sottodirectory ed entry, ovvero, valori (è simile ad un file system).
- Il registro può essere esplorato e manipolato tramite programmi appositi come `regedit`/`regedt32`.
- Il registro è manipolabile anche da programma grazie ad opportune funzioni dell'API Win32.
- Al momento dello spegnimento/riavvio del sistema, la maggior parte delle informazioni contenute nel registro viene salvata in file chiamati hive (tipicamente in `\windows\system32\config`).

Il registro di Windows

Key	Description
HKEY_LOCAL_MACHINE	Properties of the hardware and software
HARDWARE	Hardware description and mapping of hardware to drivers
SAM	Security and account information for users
SECURITY	System-wide security policies
SOFTWARE	Generic information about installed application programs
SYSTEM	Information for booting the system
HKEY_USERS	Information about the users; one subkey per user
USER-AST-ID	User AST's profile
AppEvents	Which sound to make when (incoming email/fax, error, etc.)
Console	Command prompt settings (colors, fonts, history, etc.)
Control Panel	Desktop appearance, screensaver, mouse sensitivity, etc.
Environment	Environment variables
Keyboard Layout	Which keyboard: 102-key US, AZERTY, Dvorak, etc.
Printers	Information about installed printers
Software	User preferences for Microsoft and third party software
HKEY_PERFORMANCE_DATA	Hundreds of counters monitoring system performance
HKEY_CLASSES_ROOT	Link to HKEY_LOCAL_MACHINE\SOFTWARE\CLASSES
HKEY_CURRENT_CONFIG	Link to the current hardware profile
HKEY_CURRENT_USER	Link to the current user profile

Avvio del sistema

- Il BIOS contenuto nella ROM viene eseguito determinando il dispositivo di sistema da avviare ed eseguendo il caricatore d'avvio.
- Il caricatore avvia il programma NTLDR (NT file system LoaDeR).
- NTLDR carica HAL, il kernel ed il contenuto del file hive relativo al sistema (in modo da avviare i driver necessari); alla fine viene eseguito il codice del kernel.
- Il kernel avvia due processi:
  - il processo di sistema contenente tutti i thread di lavoro interni;
  - SMSS (session manager subsystem): il primo processo utente creato (analogo a `init` per UNIX).
- SMSS lancia in esecuzione WINLOGON e CSRSS (client/server runtime server subsystem), ovvero, il sottosistema delle API Win32.
- WINLOGON installa il resto del sistema fra cui il sottosistema per la sicurezza LSASS (local security authority subsystem).