

1.4. Caratteristiche generali dei Sistemi Operativi

1.4.1. Funzione primaria di un S.O.

Il sistema operativo è la parte più importante del software di base ed ha il compito primario di costituire l'interfaccia fra l'uomo e la macchina: esso riceve da un lato istruzioni ad alto livello dall'utente e dall'altro gestisce l'hardware della macchina, inviando attraverso la CPU i comandi elementari alle diverse componenti del computer.

Ad esempio, per copiare un documento da un floppy al disco rigido, l'utente si limita a dare il comando di copia, mentre il sistema operativo si fa carico di inviare al drive del floppy i segnali elettrici di lettura e al drive del disco rigido i segnali di scrittura.

Il sistema operativo è dunque da un lato un'interfaccia tra l'operatore umano e la macchina, dall'altro un gestore della macchina stessa. Una sua parte è sempre nella memoria centrale del computer, pronto a ricevere, interpretare ed eseguire i comandi dell'utente. Una macchina è pura ferraglia senza un sistema operativo installato.

L'utente "vede" la macchina attraverso il suo sistema operativo e la vede come una *macchina virtuale*, con un suo linguaggio (il linguaggio del S.O., vedi in seguito) e le sue funzioni da compiere, quali l'esecuzione dei programmi applicativi richiesti dall'utente, la gestione degli archivi di programmi e dati, la manovra delle unità periferiche, la comunicazione all'utente di messaggi sullo stato del computer e così via.

La politica commerciale del momento è di fornire un sistema operativo preinstallato insieme all'hardware della macchina. Spesso, la fornitura comprende anche il software di utilità ed alcuni semplici programmi applicativi (ad esempio, un editor elementare e un gestore di immagini, giochi vari, etc.).

1.4.2. Alcuni sistemi operativi

Tra i sistemi operativi oggi disponibili possiamo annoverare:

- S.O. indipendenti dalla macchina, specificamente UNIX e suoi derivati (LINUX, XENIX, ULTRIX). UNIX fu sviluppato nei primi anni '70 presso i laboratori di ricerca Bell della AT&T Corporation come sistema operativo orientato allo sviluppo del software e della relativa documentazione; successivamente si è diffuso dapprima in ambito di ricerca e poi in ambito industriale. LINUX, in tempi moderni, è considerato il capostipite del software "Open source", cioè software non protetto da "copyright", che si può usare senza doverlo acquistare;
- S.O. largamente diffusi nei personal computer quali Windows nelle sue varie versioni per i PC "IBM compatibili" oppure MAC-OS per i PC della Apple Computer. Windows deriva dal suo antenato MS-DOS (nato agli inizi degli anni '80 che era un S.O. non "a finestre"); MAC-OS è il primo S.O. che ha usato le "finestre";
- S.O. cosiddetti *proprietary*, cioè tipici di una particolare serie di macchine, di uno specifico costruttore;
- S.O. per applicazioni specifiche, ad esempio controlli industriali.

1.4.3. Struttura e funzioni del S.O.

Il S.O. si compone di più sezioni che possono così schematizzarsi:

- *Interfaccia verso l'utente* detta anche *shell*: è la parte del S.O. che interpreta i comandi dell'utente; di essa si dirà in seguito;
- *Gestione dei lavori*: pianificazione dell' esecuzione dei programmi che l'utente ha chiesto di eseguire, nel senso che definisce l'ordine secondo il quale essi devono essere avviati all'esecuzione, attività indicata con il termine di *schedulazione dei lavori*;
- *Gestione dei dati*: gestione degli archivi di dati registrati nelle memorie di massa (parte essenziale di tale sezione è il *file system* di cui si dirà in seguito);
- *Gestione del sistema* (detta anche *kernel* del S.O.): gestione della CPU, della memoria centrale ed in generale dell'hardware dell'elaboratore.

Mentre si dirà in seguito dello shell e del file system, che interessano direttamente l'utente, delle altre sezioni si forniscono alcuni dettagli nell'approfondimento di questa unità formativa.

1.4.3.1. Gestione dei lavori (approfondimento)

Con la gestione dei lavori, il S.O. effettua la pianificazione dell' esecuzione dei lavori sottomessi dagli utenti, nel senso che definisce l'ordine secondo il quale i lavori devono essere avviati all'esecuzione, attività indicata con il termine di *schedulazione dei lavori a lungo termine*. Questa sezione, inoltre, gestisce gli utenti. Le funzionalità sono pertanto:

- accettazione dell'utente (password, diritti di accesso, etc.);
- attivazione dei lavori di utente;
- simulazione su memoria di massa delle operazioni di ingresso-uscita (*spooling*).

1.4.3.2. Gestione dei dati (approfondimento)

Con la gestione dei dati il S.O. gestisce l'impiego delle memorie di massa e delle periferiche da parte dell'utente; la sezione comprende i seguenti servizi:

- assegnazione degli archivi (file) a supporti fisici: l'utente tratta genericamente con un file di determinato nome, il S.O. assegna il posto fisico ove il file risiede;
- assegnazione all'utente dei file richiesti: (operazione di "apertura", vedi in seguito);
- registrazione delle informazioni sui file;
- condivisione dei file da parte di più utenti;
- protezione dei file.

Le funzionalità offerte dalla gestione dei dati consentono all'utente di fare riferimento ai propri archivi mediante notazioni simboliche, indipendenti dalle unità fisiche sulle quali sono effettivamente registrati. I moduli software per la gestione degli archivi costituiscono il *file system* il quale opera in stretto coordinamento con i sottoprogrammi di ingresso-uscita per l'effettuazione delle operazioni di trasferimento delle informazioni da o per le memorie di massa. In particolare, il file system fornisce le operazioni fondamentali per la gestione dei file a mezzo di opportune istruzioni del linguaggio del sistema operativo.

1.4.3.3. Gestione del sistema (approfondimento)

Con la gestione del sistema, il S.O. realizza la gestione delle risorse hardware (CPU, memoria, ...); in particolare, una parte fondamentale di tale sezione è il nucleo o *kernel* che gestisce la CPU, nel senso che provvede ad attribuirle ai lavori (più precisamente ai processi) che se la contendono, attività detta di *schedulazione a breve termine*, per distinguerla da quella a lungo termine che determina l'avvio del lavoro. Quando infatti esistono vari processi solo uno alla volta

può avere il controllo della CPU ed è proprio il kernel, con la sua politica di gestione, ad alternare i vari processi in gioco.

Oltre alla gestione della CPU, la gestione del sistema comprende anche: la *gestione della memoria centrale* e la *gestione dei dispositivi di I/O*, argomenti per esperti.

1.4.4. Multi-tasking, multi-users e multi-programming

I sistemi operativi moderni più diffusi sono sistemi *multiprocesso* o **multitasking**, cioè supportano l'esecuzione di più processi, nel senso che, in un dato intervallo di tempo, più processi avanzano nel computer. Non è stato sempre così: i sistemi operativi della prima generazione (MS-DOS, ad esempio) erano monoprocesso, cioè occorre che finisse un processo per poterne iniziare un altro.

I sistemi operativi possono anche essere mono o *multi-utente* (*multi-users*), a seconda se sono progettati per servire un solo utente o più utenti: nel secondo caso il S.O. deve gestire la proprietà dei programmi e dei dati, attivando anche sistemi di protezione.

Un sistema in *multi-programmazione* (*multi-programming*) è infine un S.O. nel quale diversi utenti operano *simultaneamente* sul computer mediante unità di I/O distinte: in quest'ultimo caso, i problemi di gestione del S.O. sono molto più complessi.

Ovviamente, un sistema multi-utente (come ad esempio Windows XP su un PC) non è detto che sia un sistema in multiprogrammazione (Windows XP non lo è).

1.4.5. Lavori batch e conversazionali. Il real time

Dal punto di vista delle caratteristiche di un lavoro (o di un programma) sottoposto ad un sistema, si possono distinguere due tipologie di lavori entrambe supportate dal SO:

- lavori *batch*: sono costituiti da programmi in genere complessi (o da programmi di stampa) che vengono eseguiti "in toto", senza soluzioni di continuità;
- lavori *interattivi o conversazionali*: sono suddivisi in tante piccole elaborazioni o *transazioni*; l'utente interviene in forma colloquiale con il computer fornendogli dati e ricevendo elaborazioni in ogni transazione.

Il termine *batch* deriva da una modalità di sottoporre i lavori ad un sistema che oggi si vede poco in giro, ma che era la tecnica fondamentale nel passato (vedi approfondimento). In una visione più moderna, diciamo che un lavoro è in modalità batch se il S.O. riceve dall'utente una lista di istruzioni che esegue interamente, senza soluzione di continuità, e senza ulteriori interventi dell'utente prima che la sequenza non sia terminata.

La modalità conversazionale è, invece, quella tipicamente usata nei PC e che si concretizza operativamente con le finestre di dialogo per agevolare l'interazione utente/sistema e sistema/utente mediante lo scambio di informazioni: l'utente dà un'unica istruzione, relativamente semplice, al SO che la esegue e resta poi in attesa di un'ulteriore istruzione, e così via fino al termine del lavoro.

Fra le tipologie dei lavori al computer, va anche annoverato il lavoro in *tempo reale*: si intende un lavoro che deve dare i risultati in tempo utile per l'utente. È, in generale, un lavoro interattivo con un vincolo assegnato sul *tempo di risposta* (tempo intercorrente dal momento in cui l'utente chiede una risposta al sistema e la risposta stessa), vincoli che sono rispettati a seguito di attenta progettazione dell'hardware, del software e del sistema operativo.

1.4.5.1. Sistemi batch (approfondimento)

Secondo la definizione "classica", dicesi in *batch (a lotti)* la modalità operativa secondo la quale i lavori da eseguire sono raggruppati in lotti ed immessi in tale forma nel sistema. Una definizione più moderna considera la modalità batch in antitesi a quella conversazionale: è in batch un lavoro di cui si chiede l'esecuzione "in toto", senza soluzioni di continuità.

Il lavoro in batch monoprogrammato si sviluppa in forma rigidamente sequenziale (un programma dopo l'altro) mentre nel batch multiprogrammato più lavori del lotto si sviluppano in concorrenza secondo la strategia adottata dal sistema operativo. L'obiettivo che persegue un sistema batch multiprogrammato è lo sfruttamento ottimale delle risorse hardware in modo da aumentare la *capacità di attraversamento (throughput)* del sistema, ossia il numero di lavori eseguiti per unità di tempo.

Nella visione classica del batch, ciò si otteneva mediante un oculato assortimento dei lavori di un lotto; oggi questo problema non si pone.

In un sistema batch multiprogrammato "classico", per motivi di ottimizzazione della velocità operativa, le operazioni di stampa erano simulate su memoria di massa e successivamente trasferite da questa alla stampante mediante appositi processi di sistema (detti di *spool-out*), che venivano trattati come lavori del lotto, in parallelo all'elaborazione vera e propria (il termine "spool" è l'acronimo della frase "Peripheral Symultaneous Operation On Line").

La tecnica batch classica oggi si vede raramente realizzata in un PC, ma la tecnica dello spooling è ancora usata per i processi di stampa.

1.4.5.2. Sistemi interattivi (approfondimento)

In un sistema *interattivo o conversazionale* esiste un colloquio continuo tra l'utente e l'elaboratore; sulla base dei risultati parziali ricevuti, l'utente interagisce col sistema, fornendo nuovi dati, avviando nuovi programmi, richiedendo stampe e così via.

La modalità interattiva è quella oggi prevista dalla maggioranza dei sistemi operativi, su PC o su sistemi di grande dimensione. Nel primo caso si tratta di applicazioni gestite in monoprogrammazione, nel secondo di applicazioni multiprogrammate che operano su grandi archivi condivisi da più utenti (ad esempio, i grandi sistemi di prenotazioni aerei o ferroviari).

In generale, all'attività interattiva è affiancata un'attività di tipo batch: tipicamente nei PC mentre si svolge il lavoro interattivo, può operare un programma che si dice in *background* e che sfrutta gli istanti di inattività della CPU, ad esempio un programma complesso (non conversazionale) oppure un programma di stampa.

1.4.5.3. Sistemi real time (approfondimento)

I sistemi in tempo reale sono quelli per i quali è fondamentale il *tempo di risposta*. In realtà è opportuno distinguere due classi fondamentali di applicazione:

- applicazioni per controllo di processo, in cui è richiesto che il tempo di risposta non superi un valore limite assegnato, necessario ad eventuali interventi correttivi del sistema controllato (si pensi ad esempio a sistemi nei quali è affidato al computer un controllo critico a valle del quale esso debba intervenire per evitare una catastrofe): il sistema deve essere opportunamente dimensionato;
- applicazioni di tipo gestionale, nelle quali è prefissato un valore medio del tempo di risposta, ma è tollerabile che il tempo effettivo superi quello medio: si tratta in genere di un sistema interattivo a transazioni che resta valido solo se l'utente attende un tempo

ragionevole, ma che ogni tanto può anche superare tale limite. Anche qui si richiede una attenta progettazione, ma i limiti di ritardo non sono rigidi.