Andrea Giachetti, Fabio Bettio

Fondamenti di informatica

Indice

Introduzione

L'utilizzo del calcolatore nelle sue forme moderne, personal computer, portatili, palmari, ecc e l'accesso alle reti con la possibilità di comunicare ed accedere a materiale di ogni tipo sta rivoluzionando diversi aspetti della nostra vita ed il suo effetto non può essere ignorato da chiunque lavori nel mondo della comunicazione, dell'educazione, e di ogni disciplina legata al trattamento dell'informazione.

Eppure, nonostante l'ormai ampia diffusione dei mezzi e la notevole eco che hanno sulla stampa gli argomenti legati al mondo digitale, la cultura di base relativamente all'informatica di base ed al funzionamento e le potenzialità delle tecnologie è decisamente scarsa.

Facciamo un semplice esempio: chi non ha mai sentito parlare di bit e byte, multimedialità, tecnologie client-server? Quasi nessuno. E molti, se venisse loro chiesto se sanno o non sanno di cosa si stia parlando, risponderebbero certamente di sì.
Solo che, posta la domanda ad un campione di studenti universitari e registrate le risposte, si scopre che la stragrande maggioranza di essi ha idee *errate* su cosa questi termini di uso comune significhino.

Le ragioni di questo analfabetismo e della diffusione di informazioni approssimative ed errate possono essere cercate da una parte in una purtroppo dilagante cultura antiscientifica con la pretesa esistenza di una barriera tra cultura umanistica e scienza e tecnlogia, dall'altra al modo decisamente approssimativo e sensazionalistico con cui si parla sui media più diffusi, di ciò che riguarda le tecnologie, privilegiando il fenomeno di costume, la moda (il blog, Second Life) o le paure (virus, pedofilia) all'informazione competente ed obiettiva.

Ma una parte della responsabilità va anche attribuita a noi “addetti ai lavori”.
Un primo, fondamentale passo per la divulgazione di una qualsiasi disciplina consiste infatti nel chiarire il contesto e il significato di ciò di cui stiamo parlando e dare chiare ed esaurienti definizioni dei termini utilizzati. Quando tentiamo di fornire adeguate conoscenze informatiche di base a “non addetti ai lavori” che devono fare ampio uso di calcolatori e programmi per la loro attività, spesso dimentichiamo questa regola di base e diamo per scontato che chi ascolta abbia chiaro di cosa stiamo parlando solo perché magari è un termine di uso comune.

Un altro problema su cui poi, non ci si sofferma abbastanza, è quello di cercare di far capire in che cosa effettivamente consista avere una “cultura” informatica di base, dato che non solo tra la gente comune, ma anche all'interno dei corsi di laurea non specialistici, non si riesce neppure a far capire neppure cosa studino gli informatici e cosa si dovrebbe imparare in un corso base di informatica.

Lo scopo di un buon corso base di informatica non è quello di insegnare a utilizzare in dettaglio determinati programmi, magari commerciali, con vari trucchi e trucchetti (quella che gli anglosassoni chiamano Computer/Information Literacy), ma a capire le tecnologie

ed i meccanismi che stanno alla base dei sistemi di elaborazione in modo da essere in grado di capire le tendenze di mercato, scegliere i prodotti utili, ideare o progettare nuove applicazioni nel proprio campo (Computer/Information Fluency).
Per fare un semplice parallelo, se vogliamo sfruttare al meglio quello che l'automobile può offrirci non è necessario che studiamo nel dettaglio la fisica del motore a scoppio (l'informatica teorica per il calcolatore), ma non è neppure necessario ed è anzi controproducente che impariamo nel dettaglio il manuale di istruzioni della Fiat Punto (l'equivalente di un particolare programma per il calcolatore, per esempio Microsoft Word). Dovremo, invece, avere una sufficiente infarinatura tecnica sul funzionamento dell'automobile e conoscerne le varie tipologie, l'evoluzione del mercato, degli accessori, e tutte quelle nozioni che ci servono a scegliere il modello adeguato ai nostri bisogni ed ad utilizzarlo nel modo più efficiente.

E' invece frustrante vedere che addirittura a livello universitario si confonda ancora lo studio dell'informatica con l'*uso* di alcune applicazioni per il personal computer e si pensi che un corso di informatica tenuto da informatici dovrebbe insegnare ad usare dei programmi. Nulla di più lontano dalla realtà. L'uso di un determinato programma non richiede competenze informatiche e può benissimo essere lasciato alla lettura del manuale od a tutorial tenuti da utenti dello stesso genere (insegnanti, scrittori, contabili, grafici, a seconda del programma di cui si parla).
L'informatica e gli informatici si occupano, in effetti, di altro. L'"informatica", più che una semplice disciplina, è un insieme di discipline che hanno a che fare con:

- L'informazione e la sua elaborazione
- Il linguaggio e le sue proprietà
- L'architettura e il funzionamento delle macchine calcolatrici
- Le tecniche per elaborare l'informazione con tali macchine
- I sistemi operativi che rendono utilizzabili le macchine
- Le applicazioni pratiche dei calcolatori
- I sistemi che collegano tra loro più calcolatori

Il concetto di "informazione" è qui inteso in senso molto generico: l'informatica si occupa di elaborazione di dati eterogenei, come testi, dati numerici, immagini, suoni.
Anche i concetti di rappresentazione ed elaborazione sono altrettanto ampi: l'informatica si occupa di rappresentazione grafica, meccanica, sonora, trasformazione tra diversi formati, trasmissione di dati a distanza, eccetera. L'unico punto fisso è che tutte le forme di informazione, per essere elaborate dal calcolatore, devono essere "codificate" in segnali binari, sequenze di 0 e 1 o valori alti e bassi di tensione su cui il microprocessore (l'unità di calcolo della nostra macchina) è in grado di operare.
Utilizzando i sistemi a semiconduttore per l'elaborazione di dati binari sono state costruite macchine capaci non solo di fare operazioni aritmetiche ripetitive (il solo compito dei primi calcolatori elettronici), ma anche in grado di affiancarci in numerosissime attività lavorative o ricreative (dalla scrittura, al gioco, dall'educazione alla simulazione).
Tutto questo è stato reso possibile dall'evoluzione di tutte le branche dell'informatica, e si può quindi facilmente comprendere come questa sia diventata una disciplina sempre più varia, complessa ed affascinante.
Certo, non è utile né richiesto (anche se non certo dannoso) che un insegnante, un contabile o un geometra conoscano i problemi dell'algoritmica, della programmazione o dei protocolli di rete, ma la conoscenza dei fondamenti tecnologici degli strumenti che usano sarebbe per loro indubbiamente di grande utilità. Per questo motivo cercheremo di riassumere nelle dispense quel minimo di informazioni sulle tecnologie informatiche che

possono meglio far capire come fanno a funzionare le comuni applicazioni per PC.
Nei capitoli seguenti faremo quindi una panoramica sulle varie problematiche legate all'elaborazione automatica dell'informazione, vedremo la struttura ed il funzionamento di un comune personal computer, come siano fatti i programmi che lo gestiscono e come i calcolatori possano essere connessi in rete e comunicare tra loro.

In dettaglio, nel capitolo 1 vedremo che cos'è esattamente un computer e ne vedremo brevemente l'evoluzione storica, nel capitolo 2 verranno introdotti i concetti di calcolo automatico ed algoritmi, nel capitolo 3 si descrive come l'informazione di nostro interesse sia trasformata in dati elaborabili dal calcolatore (codifica). Nel capitolo 4 passeremo quindi ad analizzare il software di sistema (il "sistema operativo" dei computer) e l'utilizzo moderno delle macchine di calcolo. Nel capitolo 5 si parlerà delle reti di calcolatori. Il capitolo 6 è infine dedicato al mondo di Internet e dei suoi servizi.

1. Che cos'è un “computer”?

Può essere difficile da comprendere per chi al giorno d'oggi usa un PC per scrivere, giocare o guardare un film, a chi aggiorna l'agenda sul cellulare, viaggia col navigatore satellitare o usa l'iPod, ma, in realtà, tutte queste macchine elettroniche discendono direttamente da quei “cervelli elettronici” grossi come armadi, che cominciarono a comparire nel dopoguerra e che elaboravano dati numerici introdotti in genere su schede perforate facendo una serie di calcoli preordinati.
Ma svolgere in sequenza una serie di calcoli aritmetici o logici su dati *digitali* (cioè codificati come numeri binari e fisicamente consistenti in livelli alti o bassi di tensione) è sostanzialmente la stessa cosa che fanno, senza che ce ne rendiamo conto, tutte le apparecchiature basate su microprocessori che utilizziamo quotidianamente per le più disparate attività.
E l'architettura dei moderni computer non si discosta sostanzialmente da quella di quegli armadi che vediamo nei vecchi telefilm di fantascienza.

1.1 Architettura del calcolatore

I calcolatori *elettronici* sono macchine che eseguono ripetutamente operazioni *programmate* su dati *digitali*, sfruttando l'elettronica dei transistor.
In tutti i tipi di calcolatore elettronico, abbiamo sempre una (od eventualmente più di una) unità di calcolo (CPU, Central Processing Unit), una memoria centrale che possiamo pensare come un enorme casellario che contiene i numeri binari da elaborare dove la CPU preleva e scrive dati, e dispositivi che permettano all'utente di inserire dati e leggere risultati.
Lo schema logico (Figura 1) prende il nome di architettura di Von Neumann, dal nome dello scienziato ungherese che introdusse, appunto, nei calcolatori del dopoguerra l'idea del computer a “programma memorizzato”. Vedremo tra poco un piccolo riepilogo dell'evoluzione storica dei calcolatori elettronici.

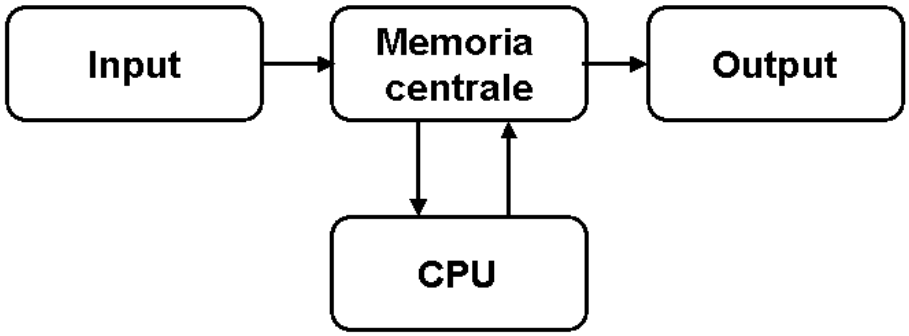

Figura 1: Architettura di Von Neumann

La peculiarità di questo schema sta, come accennato, nell'utilizzo di una sola memoria per contenere sia i dati da elaborare sia i programmi per elaborarli (cioè la lista di operazioni che cogliamo fare eseguire alla CPU in successione). Questo schema venne applicato per

la prima volta nel computer EDVAC del 1950, e successivamente universalmente adottato. Ed è sostanzialmente lo stesso utilizzato nei moderni calcolatori, salvo il fatto che, in genere, la comunicazione tra le parti (CPU, memoria, periferiche di I/O) avviene attraverso dei canali condivisi che prendono il nome di bus di comunicazione (Figura 2).

L'attività del computer consiste quindi, una volta caricato in memoria un programma, nel

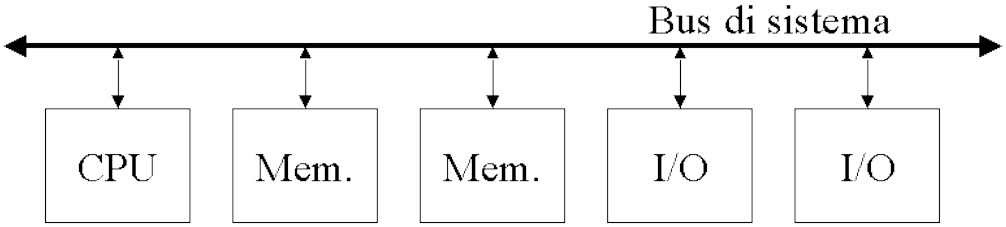

Figura 2: Architettura a bus di comunicazione

trasferimento alla CPU dei suoi singoli passi, nella loro decodifica ed esecuzione e nel passaggio all'istruzione successiva. Tutto qui. L'enorme quantità di cose che possiamo fare oggi con i calcolatori elettronici dipende sostanzialmente dal fatto che:

- Si possano codificare con stringhe binarie, quindi elaborabili dal calcolatore dati di ogni tipo (vedi capitolo 3)
- L'evoluzione delle CPU abbia portato alla possibilità di effettuare quantità enormi di calcoli al secondo anche su macchine dai costi irrisori
- L'evoluzione dei “programmi” di base e dei sistemi operativi (capitolo 4) ha reso possibile facilitare enormemente lo sviluppo di applicazioni nuove per i calcolatori.
- L'evoluzione delle periferiche di Input/output (ingresso/uscita) e delle interfacce utente abbia reso possibile una facile interazione con le macchine anche da parte di utenti non esperti.

Vediamo qualche dettaglio su questa architettura di base dei calcolatori, andando ad analizzare, per prima cosa, il suo “cuore”, cioè la CPU

1.2 La CPU

La CPU (Central Processing Unit) è appunto il cuore del nostro sistema. Essa contiene gli elementi circuitali necessari al funzionamento dell’elaboratore, quei circuiti integrati che operano effettivamente sulle sequenze di valori binari.
La CPU esegue i programmi che risiedono nella memoria centrale, e coordina il trasferimento dei dati tra le varie unità. Essa si compone di:

- una unità di controllo (CU)
- una unità aritmetico-logica (ALU)
- alcuni dispositivi di memoria detti registri

L’Unità aritmetico-logica è costituita da dispositivi circuitali che consentono di eseguire le operazioni aritmetiche somma, sottrazione, prodotto, divisione (ADD, SUB, MUL, DIV) o logiche (AND, OR, NOT) sugli operandi memorizzati nei registri interni all'ALU, come il registro accumulatore (A), dove vengono inseriti risultati parziali, il registro operando (OP) il registro di stato (PSW) ove vengono inserite informazioni sullo stato di avanzamento dei comandi (ad esempio riporti risultati di confronti, ecc.). Una caratteristica fondamentale della CPU è quindi la dimensione dei registri che corrisponde al numero di bit che possono essere elaborati contemporaneamente. I processori attuali arrivano a 64 bit; ovviamente alla dimensione dei registri dovrà corrispondere un analoga o proporzionale ampiezza del

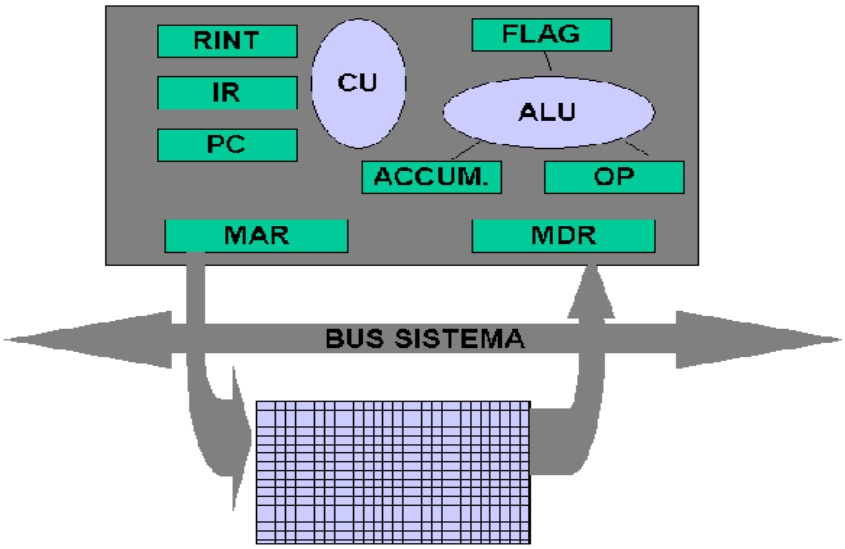

Figura 3: Schema della CPU

bus di comunicazione e delle caselle di memoria. L’unità di controllo si occupa della sequenzializzazione delle operazioni cioè della corretta esecuzione in successione dei comandi. Manda i segnali di controllo e gestisce lo svolgimento del ciclo fondamentale della macchina, cioè la ripetizione continua delle fasi di lettura delle istruzioni di programma, dei dati, decodifica ed esecuzione dell’istruzione e passaggio all’istruzione successiva.

Anch’essa conterrà dei registri per memorizzare lo stato delle operazioni in corso e per gestire operazioni di servizio.
Registri fondamentali della CPU sono:

- il registro degli indirizzi di memoria (MAR Memory Address Register): in
- esso viene memorizzato l’indirizzo di memoria dove estrarre o scrivere il dato ed è collegato direttamente al bus indirizzi
- il registro dei dati di memoria (MDR Memory Data Register), che contiene una copia del dato letto o da scrivere ed è quindi collegato al bus dati
- il contatore di programma (PC), che in realtà non conta nulla, ma memorizza l’indirizzo di memoria dove si trova la prossima istruzione da eseguire
- il registro della istruzione corrente (IR), che contiene l’istruzione corrente che deve essere decodificata ed eseguita
- il registro delle interruzioni

Il microprocessore, che integra le componenti della CPU, è realizzato unendo minuscoli transistor, cioè dispositivi a semiconduttore in grado di assumere stati diversi in funzione delle tensioni applicate (acceso/spento). La capacità di elaborazione del processore dipende ovviamente dal numero di transistor “integrati” in esso.
La stupefacente crescita dipende anche dalla miniaturizzazione dei contatti dei circuiti integrati, che sono ormai giunti alla tecnologia 0.13-0.15 micron (un capello è spesso circa 100 micron) ed è sempre stata approssimata dalla legge empirica nota come “legge di Moore” (da una previsione fatta nel 1965 da Gordon Moore) secondo cui il numero di transistor raddoppierebbe ogni due anni (corretti poi in 18 mesi). Al di là del fatto che non si possano usare leggi di questo tipo per fare reali previsioni sul futuro, la legge è stata approssimativamente rispettata non solo per quanto riguarda il numero di transistor, ma anche per altre evoluzioni di componenti informatiche (velocità di calcolo, dimensioni memorie, ecc.) che seguono un trend di crescita di tipo esponenziale. In Figura 4 vediamo

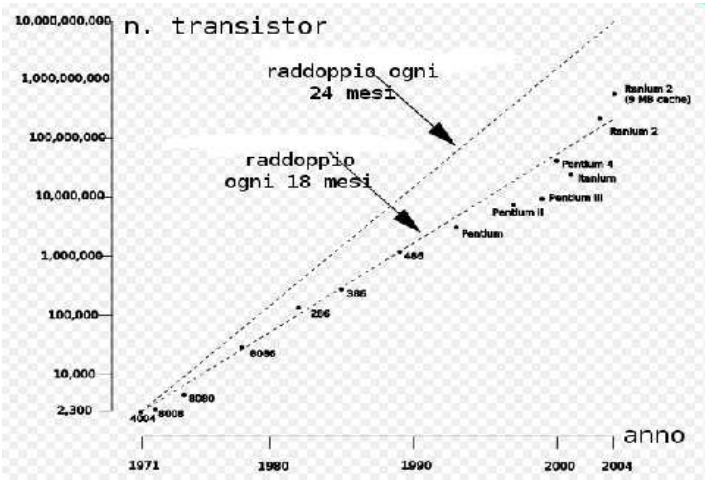

Figura 4: Evoluzione dei microprocessori della famiglia Intel

l'andamento del numero di transistor per la famiglia di processori Intel.

L'attività del microprocessore è quella di eseguire, a computer acceso, un ciclo di operazioni continuo (sempre attivo fino allo spegnimento fisico della macchina), che implica tre principali fasi, eseguite ciclicamente:

- Lettura dell'istruzione (FETCH): l'istruzione viene trasferita dalla memoria al registro istruzioni
- Decodifica dell'istruzione (DECODE): viene identificata l'istruzione tra quelle dell'insieme a disposizione
- Esecuzione (EXECUTE): vengono eseguite le operazioni previste dall'istruzione

Questo, nell'organizzazione a bus vista prima significa che il cosiddetto **bus di sistema** è suddiviso in realtà in tre parti distinte:

- una monodirezionale dal processore alla memoria detta bus degli indirizzi,
- una bidirezionale dal processore alla memoria e viceversa detta bus dei dati
- una monodirezionale dal processore alle altre unità funzionali detto bus dei comandi.

Vediamo in dettaglio l'esecuzione delle fasi del ciclo: all'inizio l'unità di controllo fornisce alla memoria l'indirizzo della cella contenente la prima istruzione da eseguire: la richiesta di ricerca dell'istruzione viene realizzata mediante la scrittura dell'indirizzo nel registro MAR e l'attivazione di un segnale di controllo sul bus che determina l'azione di lettura. La memoria seleziona la cella contenente l'istruzione, corrispondente all'indirizzo inviato sul rispettivo bus e pone il contenuto sul bus dati cosicché può venire copiato sul registro MDR. Il comando viene copiato sul registro IR. A questo punto è completata la fase di fetch e l'unità di controllo aggiorna il registro PC al valore corrispondente al comando successivo. L'unità di controllo esamina il contenuto del registro istruzioni e determina il codice operativo, cioè le operazioni da svolgere (decode). Infine vengono eseguite le varie fasi dell'istruzione mediante i segnali di controllo, fasi che possono prevedere la
lettura di altri dati dalla memoria (operandi) e il trasferimento di risultati dai registri alla memoria centrale. Finita l'esecuzione, si torna alla fase di fetch. Il tempo di ciascun ciclo dipende ovviamente dal tipo di istruzione; per velocizzare l'attività del calcolatore spesso si usa la tecnica del pre-fetch, il nuovo comando viene già letto prima che finisca l'esecuzione del precedente.
Il sicnronismo delle attività di lettura, scrittura ed interpretazione dei segnali deve essere

garantito da un orologio che determini la cadenza temporale delle operazioni elementari. Il “clock” della CPU è quello che svolge questo compito, ed è caratterizzato da una frequenza, cioè l’inverso della durata di ciascuna operazione elementare. Tanto più alta sarà la frequenza, tanto maggiori saranno le capacità del calcolatore di svolgere compiti velocemente, per cui nell'evoluzione dei microprocessori le frequenze sono cresciute rapidamente da pochi KHertz (KiloHertz) ai moderni valori dell'ordine dei GHz (GigaHertz), cioè del miliardo di operazioni elementari per secondo.
L'aumento continuo del numero di transistor e della frequenza del ciclo per migliorare le prestazioni delle macchine ha chiaramente limiti fisici dovuti anche al consumo di energia ed al calore dissipato, per cui nelle ultime generazioni di microprocessori si è passati invece allo sfruttamento del parallelismo con le cosiddette CPU multi-core.
Risolvendo problemi di sincronizzazione non indifferenti, infatti, si può suddividere il calcolo richiesto al microprocessore a più unità (core) integrati sullo stesso chip e che accedono alle stesse memorie. Questa svolta verso il parallelismo che migliora le prestazioni purché i programmi utilizzino elaborazione parallela (attraverso un meccanismo che prende il nome di multi-threading), è arrivata oggi anche ai PC domestici con i recenti chip Intel e AMD.
L’efficienza e le caratteristiche del microprocessore dipendono, ovviamente, dal *set di istruzioni* che è in grado di eseguire (che corrispondono quindi ai codici operativi) e da come queste operazioni sono implementate sui circuiti logici. La programmazione delle istruzioni viene detta “microprogrammazione”, ed è il più basso livello di programmazione presente sulla macchina. Due microprocessori che hanno lo stesso set di istruzioni possono essere compatibili anche se la loro struttura interna è differente e possono quindi eseguire gli stessi programmi, cosa non vera nel caso generale. I personal computer “IBM compatibili” attuali utilizzano processori Intel o compatibili costruiti da concorrenti come AMD.

1.3 La memoria

Abbiamo parlato della memoria come di un oggetto su cui, durante il suo ciclo, la CPU scrive e legge delle “parole”, copiandole da e su alcuni suoi registri spostandole attraverso il bus. In essa sono memorizzati sia i dati da elaborare, sia il programma che li elabora, e su di essa vengono scritti i risultati dell’elaborazione.
La memoria principale del calcolatore è una lista di caselle, tra le quali se ne seleziona una mediante una serie di bit (quelli che vengono trasmessi dal bus indirizzi) su cui leggere o scrivere. Essa è realizzata fisicamente mediante dispositivi a semiconduttore in grado di mantenere o variare valori di tensione in funzione di input esterni (flip-flop, condensatori).
La capacità della memoria sarà pari, per quanto abbiamo detto, alle dimensioni della parola (corrispondente in genere a quella del registro della CPU) per il numero di caselle disponibili. L’unità che viene usata per l’estensione della memoria è il Byte, sequenza di 8 bit, dal momento che le dimensioni delle parole caratteristiche delle CPU sono sempre multipli di 8 bit. Multipli caratteristici del byte sono Il Kilobyte (KB) = 2^{10} byte = 1024 byte, Il Megabyte (MB) = 2^{20} byte = 1024 KB byte, il Gigabyte (GB) = 2^{30} byte = 1024 MB, il Terabyte (TB) = 2^{40} byte = 1024 GB che significa oltre mille miliardi di byte!
Tornando alle operazioni di lettura e scrittura, la Figura X mostra come vengano “indirizzate” le caselle della memoria stessa: il bus di controllo manda il segnale che seleziona l’operazione di lettura, scrittura, il bus indirizzi posta al dispositivo di memoria il contenuto del registro MAR della CPU, ed infine viene scritta nella casella il contenuto del registro MDR che è presente nel bus dati (operazione di scrittura) oppure viene posto sul bus dati il contenuto della casella che passa al registro (operazione di lettura)

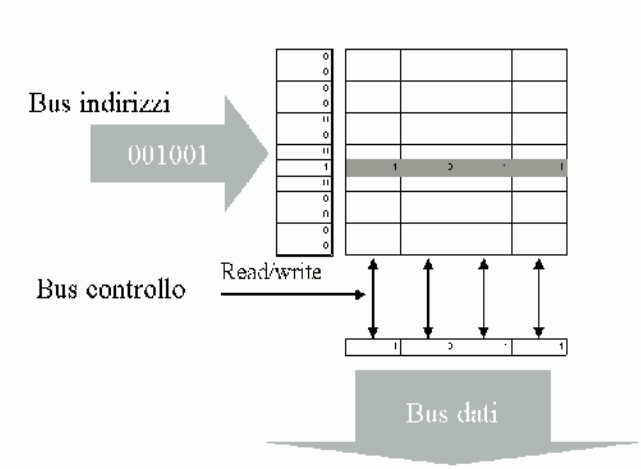

Figura 5: Lettura e scrittura dalla memoria: sul bus indirizzi viene indicata la casella per la scrittura/lettura, sul bus di controllo il comando da eseguire, ed i dati vengono trasferiti sul bus dati

Esistono vari tipi di dispositivo di memoria. Per quanto riguarda la tecnologia, possiamo avere, per esempio le memorie di tipo ROM (Read Only Memory), che possono essere solo lette; queste possono essere non modificabili o programmabili dall'utente con opportune metodologie, ma non ovviamente durante l'utilizzo del calcolatore. Esempi di memorie ROM programmabili sono PROM, EPROM, EEPROM e così via.
Le memorie su cui la CPU può scrivere e leggere si chiamano ad "accesso casuale" o RAM (Random Access Memory, significa che si può accedere ad ogni locazione direttamente mediante l'indirizzo, senza bisogno di scorrere tutte le locazioni precedenti). Anche qui possiamo avere varie sigle a seconda della tecnologia adottata, come SRAM, DRAM, CMOS. Le RAM dinamiche (DRAM) si basano su condensatori invece che su transistor (ne basta uno usato come interruttore); dato che la carica sui condensatori si mantiene per pochissimo tempo, devono essere "riscritte" periodicamente con un operazione che viene detta "refresh". Le RAM statiche (SRAM) utilizzano invece un circuito con più transistor e risultano quindi molto più costose (e ingombranti), ma anche molto più veloci.

I parametri che caratterizzano le prestazioni della memoria sono:

- le dimensioni cioè la capacità in numero di byte, ovviamente una grande capacità incrementa le potenzialità della macchina.
- il tempo di accesso (tipicamente dell'ordine delle decine di nanosecondi) che è cruciale per il calcolo ad alta velocità.
- la volatilità (cioè dopo quanto si cancella il segnale). Tipicamente le memorie RAM sono tutte di tipo volatile, cioè non mantengono il segnale se non alimentate da corrente, mentre ROM, Flash, dischi magnetici mantengono il dato anche a calcolatore spento. Le RAM devono essere sempre alimentate per mantenere il segnale, quelle dinamiche anche riscritte a intervalli regolari.
- Il costo per bit, che è un parametro cruciale per rendere un tipo di memoria preferibile per i produttori.

Ogni dispositivo di memoria ha le sue caratteristiche peculiari che vengono sfruttate al meglio dai progettisti dei sistemi. La memoria principale, quella indirizzata dalla CPU e che nei moderni PC ha dimensioni di centinaia di KiloBytes, è realizzata con RAM dinamiche perché, sebbene più lente, sono più piccole ed economiche delle RAM statiche.
La RAM statica per realizzare un sistema per velocizzare l'accesso ai dati, ossia la memoria cache. Essa consiste in una duplicazione su supporti ad accesso più veloce di

parte del contenuto della memoria principale. Se, come si cerca di fare, questa parte di memoria contiene le locazioni più spesso utilizzate, le prestazioni del computer risultano in media molto superiori.
Essendo il suo contenuto una copia della memoria principale la cache non aumenta il numero di indirizzi a disposizione della CPU.
Le parti della memoria principale che vengono copiate nella cache vengono scelte sfruttando il cosiddetto “principio di località” per cui è estremamente probabile che in tempi vicini il processore acceda a regioni di memoria contigui (località spaziale) e visitate di recente (località temporale). Mantenendo nella memoria cache regioni contigue di memoria utilizzate di recente dai programmi, le operazioni successive utilizzeranno dati che con grande probabilità verranno letti nella memoria veloce senza dover aspettare il tempo di accesso alla memoria principale.
Nei casi in cui il dato non sia presente nella cache, si ha il cosiddetto “cache miss” e si deve accedere comunque alla memoria principale. Saranno quindi necessari opportuni algoritmi che provvedano alla gestione di questi casi, e provvedano anche a mantenere coerente il contenuto della cache con il contenuto della memoria principale e ad aggiornare via via il contenuto della cache. Il meccanismo della cache può essere realizzato anche su più livelli con lo stesso concetto; nei PC si distinguono in genere cache di primo e secondo livello, la prima (L1) in genere integrata nel microprocessore, la seconda (L2) esterna.

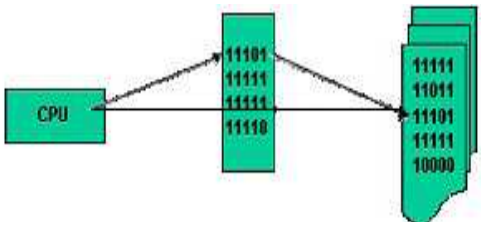

Figura 6: Il meccanismo della memoria cache: parte del contenuto della memoria principale viene ricopiato in una memoria di più veloce accesso

Memoria principale e cache non esauriscono le tipologie di dispositivi di memorizzazione nei calcolatori. Gli altri, in generale le cosiddette “memorie di massa” sono in genere considerati esterni al modello di Von Neumann e quindi considerabili in tale schematizzazione come “dispositivi periferici”.
Quello più significativo è una vastissima memoria non volatile (dell’ordine dei Gigabyte) detta memoria di massa o secondaria, su cui vengono immagazzinati i dati ed i programmi non utilizzati dalla CPU. Essa ha tempi di accesso molto più lunghi rispetto alle RAM e si basa in genere su dischi magnetici.
Data la non volatilità essa serve principalmente per salvare dati e programmi degli utenti e memorizzare il sistema operativo da caricare all’accensione del computer. In alcuni casi, come vedremo in seguito, può anche essere utilizzata analogamente alla memoria principale per accrescere le dimensioni del casellario visto dalla CPU per permettere l'esecuzione di programmi che allochino quantità di memoria superiori alla RAM disponibile.

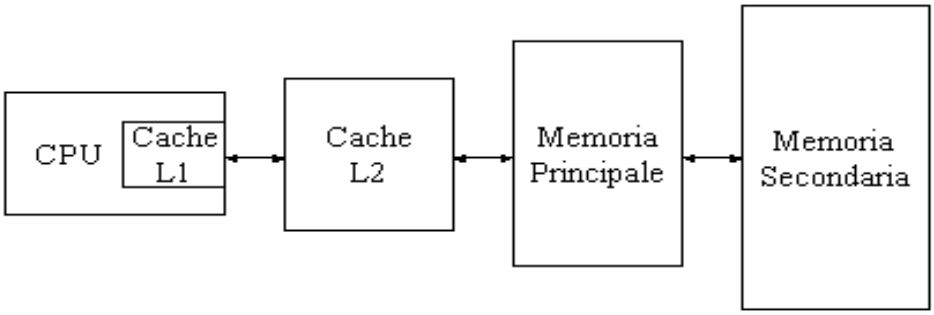

Figura 7: Gerarchia delle memorie sui moderni calcolatori: a sinistra abbiamo i dispositivi più piccoli e veloci, a destra quelli più lenti e capaci

Il complesso dei dispositivi di memoria che si usano nei calcolatori definiscono la cosiddetta "gerarchia delle memorie", in cui si ha il passaggio da memorie piccole e veloci (a sinistra in Figura 7), "vicine" alla CPU, a memorie sempre più grandi, ma lente (a destra in Figura 7).

1.4 Gestione dell'Input/output

I dispositivi di Input/output sono, nella schematizzazione di Figura 2, collegati al bus di sistema. Ma come dialogano periferica, memoria e CPU attraverso il bus?
Innanzi tutto nei sistemi di controllo dei dispositivi periferici o di I/O devono esserci dei registri per contenere i dati da mandare o ricevere dal bus, definire lo stato (lettura, scrittura, ecc.). Devono esserci, poi, unità di controllo ed un sistema di sincronizzazione. Deve esserci, insomma, un sistema "intelligente" di gestione dell'input/output che parli un linguaggio "concordato" (protocollo) con la CPU. Questa, a sua volta, gestisce le procedure di lettura e scrittura sui dispositivi secondo le necessità dei programmi che sta eseguendo.
Due sono le strategie tipiche con le quali può essere gestito l'I/O dalla CPU:

- il polling, cioè la CPU a cadenze prefissate "va a vedere" se le periferiche sono attive e devono inserire o prelevare dati;
- l'interrupt, con la quale sono gli stessi dispositivi periferici a segnalare alla CPU la loro richiesta di inserire o prelevare dati. In questo caso, la CPU dovrà possedere un sistema di gestione delle richieste di interruzione al fine di poter mantenere il suo funzionamento, soddisfare alla richiesta del periferico e riprendere la sua attività senza perdere dati o bloccare programmi in esecuzione. Si tratta di compiti piuttosto complessi, gestiti nei calcolatori a livello di sistema operativo.

L'utilizzo continuo della CPU per gestire il trasferimento dei dati rallenta in molti casi le prestazioni del calcolatore. Per evitare di coinvolgerla continuamente, può essere utile introdurre un secondo sistema di gestione detto "controller" DMA (Direct Memory Access) che fa in modo che le periferiche possano direttamente "scrivere" in memoria, mentre la CPU può continuare ad operare (senza ovviamente mettere dati in conflitto sul bus). La CPU viene interrotta solo per dare il via al trasferimento, dopodiché "lascia" il bus al DMA, che lo libera a trasferimento finito. Vedremo comunque meglio I meccanismi di gestione delle memorie quando ci occuperemo del software di base del PC, cioè dei sistemi operativi (capitolo)

1.5 Dal modello al PC fisico

Dove troviamo la struttura che abbiamo visto nei nostri PC? Quello che vediamo è, in genere, una scatola di metallo (il case) con un alimentatore e varie porte di uscita che lo collegano a tastiera, mouse e eccetera, slot per dischi, pulsanti. Dove si trova l’architettura di Von Neumann vista in teoria? Dove sono il bus, le memorie, la CPU? Per trovarli occorre aprire il case (Figura 8). Si vede subito che il cuore del PC è una grossa scheda che monta un gran numero di circuiti integrali, detta scheda madre
Essa contiene uno od anche più microprocessori, i bus di comunicazione, degli alloggiamenti per la memoria, i connettori per le unità esterne e svariati altri oggetti. Le memorie RAM (Random Access Memory) che costituiscono quella che abbiamo visto essere la “memoria principale” del calcolatore sono in genere piccole schede che contengono la circuiteria adibita alla memorizzazione. Abbiamo anche visto che la memoria principale deve essere costituita da RAM di tipo dinamico, esistono tuttavia vari

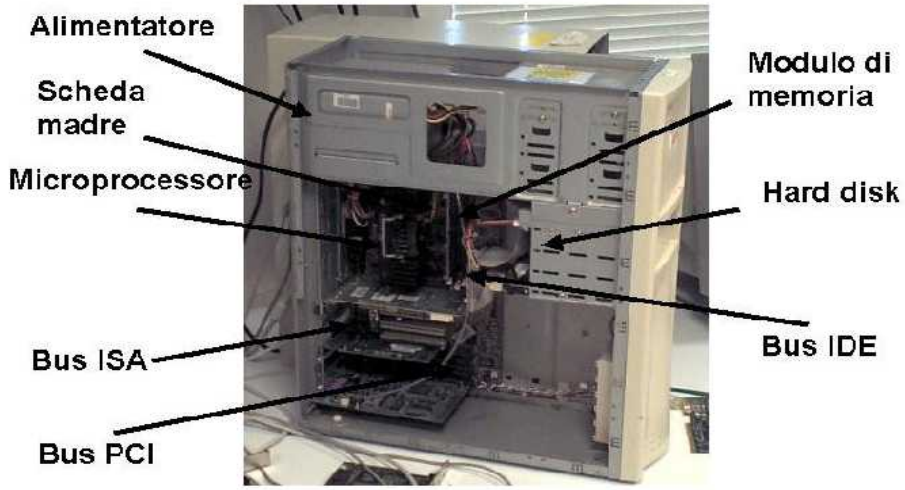

Figura 8: Un PC Desktop aperto mostra le parti di cui si compone e che corrispondono a quelle dello schema descritto

tipi anche di queste memorie, che si sono evolute negli anni: oggi si usano in genere memorie di tipo sincrono (cioè le operazioni vengono scandite da un clock periodico) e dinamico (SDRAM), dette DDR (Double Data Rate) in quanto le operazioni di lettura scrittura avvengono su entrambi i fronti (ascendente e discendente) del segnale di clock che le regola.

1.6 I dispositivi periferici

Se proseguiamo nella nostra ispezione all’interno del calcolatore oltre agli elementi visti troviamo, a completamento della “materializzazione” del sistema visto prima in astratto, anche tutta una serie porte e di connessioni esterne che costituiscono la connessione con quelle che avevamo chiamato nel modello astratto unità di Ingresso Uscita (Input/Output).
Nella pratica si tratta del disco rigido per la memoria di massa e di tutte le altre periferiche che ci permettono quotidianamente l’utilizzo del computer (tastiera, mouse, video, ecc.). Come si può osservare in Figura 8, tutto si collega alla scheda madre tramite i numerosi

spinotti ed alloggiamenti che vi sono integrati: il bus IDE per il disco rigido, slot per le schede esterne (video, audio, ecc.), porte per tastiera, mouse, microfono, ecc.
Occorre notare che l'evoluzione del calcolatore ha portato a massimizzare il disaccoppiamento tra i dispositivi "esterni" e la CPU, creando circuiterie "intelligenti" di connessione e gestione degli stessi. Infatti, se la CPU dovesse gestire completamente anche i periferici si avrebbe un degrado delle prestazioni. Nei calcolatori più moderni, i periferici contengono addirittura processori dedicati, alloggiati fisicamente al loro interno.
I dispositivi periferici, come abbiamo detto, vengono in genere distinti tra:

- Unità di ingresso, che permettono di introdurre nel calcolatore dati, programmi o di comunicare con dispositivi normalmente non compatibili col calcolatore fungendo così da interfaccia: ad esempio unità di ingresso sono tastiera, mouse, convertitori analogico-digitali, microfoni, scanner, lettori di codici a barre, telecamere, e così via.
- Unità di uscita, che permettono di visualizzare in qualche modo i risultati di elaborazione: esempi sono ovviamente il monitor, la stampante, il plotter, ecc.

Alcune unità possono essere anche bidirezionali: ad esempio i dispositivi di memorizzazione esterna, come nastri, dischi magneto-ottici, ecc.
Le unità esterne sono di varia natura e sono prodotti dai più svariati costruttori. Come è possibile garantire che si adattino tutte al nostro PC? La risposta risiede nei termini: interfaccia e standard.

- Un'interfaccia è la connessione del computer con l'esterno. Si compone di un aspetto hardware (la connessione fisica coi suoi spinotti maschio e femmina, e i suoi piedini), e di protocolli con cui i dati sono trasmessi attraverso di essa.
- Standard significa che i produttori di PC e periferiche si adattano ai tipi definiti di interfaccia di modo che non dovrebbero sussistere problemi di connessione. Esistono organismi di standardizzazione (ISO, IEEE, ecc.) che definiscono le specifiche dei vari aspetti del collegamento una volta che la tecnologia viene proposta e sperimentata e una volta definito lo standard, i produttori devono adeguarsi ad esso.

Come vedremo, poi, chi programma le applicazioni del calcolatore non deve utilizzare conoscenze specifiche del dispositivo per utilizzarlo, ma agire con le stesse modalità (ad esempio richiamando funzioni standard del sistema operativo) indipendentemente dalla periferica montata.
Questa interfaccia che consente l'"astrazione" nell'uso delle periferiche è realizzata attraverso i cosiddetti "driver", componenti software che sono in genere forniti direttamente dal produttore, che si integra col sistema operativo e permettono quindi allo steso di gestire correttamente l'unità utilizzando i suoi meccanismi standard (dunque un driver è specifico per un dato sistema operativo).

1.7 Bus e schede di espansione

La scheda madre nelle illustrazioni viste sopra mostra connettori o "slot di espansione" che permettono di inserire schede per estendere le potenzialità del calcolatore o collegare periferiche tramite di esse. Tali slot sono collegati ai bus. Nei PC possono esistere più bus, collegati in modo opportuno tra loro, l'evoluzione ha portato al succedersi di vari standard per i bus di sistema e locali (IDE, EIDE i primi, a soli 8, 16 bit e basse frequenze, poi PCI, AGP sempre più veloci e potenti per poter smistare la grande quantità di informazione elaborata dai processori e dalle schede grafiche). Schede video, audio, dischi rigidi, dispositivi di rete, ecc., sono inserite negli appositi connettori e si interfacciano poi direttamente con le unità esterne. Le schede devono contenere l'apposita logica di controllo che faccia dialogare la periferica con il PC nei modi prescritti dal protocolli del bus scelto.

Quando si inserisce una nuova periferica occorre normalmente spegnere il computer, collegarla, e installare il software che la gestisca. Tuttavia nei moderni sistemi operativi si parla di tecnologia “plug and play”, cioè il sistema “riconosce” automaticamente la periferica e installa automaticamente il driver per gestirla, per cui il dispositivo viene automaticamente configurato senza spegnere la macchina.
Per fare questo intervengono il BIOS (Basic Input Output System, programma che gestisce le operazioni base su periferici e che risiede in una ROM), che verifica la presenza del nuovo dispositivo ed il sistema operativo, che deve trovare sul disco, scaricare dalla rete o chiedere l’installazione del driver corretto all’utente.

1.8 Porte di collegamento

Sul retro dei PC esistono in genere le porte cui collegare i cavi di collegamento ad alcuni dispositivi esterni standard. Per ciascuna delle interfacce codificate esiste un opportuno connettore. Storicamente i PC hanno in genere presentato le porte seriali (RS232) e parallele (CENTRONICS).

Porta parallela

Come suggerisce il nome, abbiamo più canali ove i bit vengono trasmessi in parallelo (8 canali). Occorrono più fili ai due lati, e per questa complicazione la si usa in genere unidirezionalmente (es. per la stampante). I cavi di collegamento hanno anche lunghezza massima minore.

Porte seriali

Sono canali di comunicazione a 1 bit.
I bit vengono trasmessi in serie a distanza temporale costante, che definisce il bit rate, cioè il numero di bit trasmessi nell’unità di tempo. I bit sono trasmessi in pacchetti di dimensione fissata, con possibile presenza di bit estranei al segnale utilizzati per controllo di errore (parità). Gli standard delle porte seriali di vecchia generazione, che definiscono i livelli di tensione del bit, la temporizzazione, ecc. sono l’ RS 232 e RS432.
Anche per quanto riguarda le connessioni esterne c'è stata comunque un'evoluzione delle tecnologie e degli standard.
Lo standard *Universal Serial Bus (USB)* è utilizzato da tutti i calcolatori di nuova generazione. Esso consente di collegare fino a 127 dispositivi in serie, è molto più veloce delle vecchie seriali (12 Mbps) e consente di gestire facilmente i collegamenti per tutti i dispositivi che non richiedano il trasferimento di grosse quantità di dati (dischi rigidi ed ottici). Si utilizza per mouse, stampanti, scanner, modem. Ha inoltre gli ulteriori vantaggi di utilizzare la tecnologia “plug and play” (i dispositivi collegati vengono subito individuati dal sistema), non richiedere lo spegnimento del PC per il collegamento, di utilizzare cavi sottili e di trasmettere anche la corrente per l’eventuale alimentazione della periferica.
Queste caratteristiche sono condivise anche dalle porte Firewire (IEEE 1394), standard creato dalla Apple, ma anch’esso universalmente diffuso. Questa connessione è peraltro molto più potente e costosa, e viene utilizzata soprattutto per dispositivi come videocamere digitali, dischi, ecc. Può collegare fino a 63 dispositivi in serie.

1.9 Alcuni dispositivi periferici

Monitor

Figura 9: Monitor CRT, in via di estinzione...

Per la visualizzazione di ciò che il computer elabora si utilizzano apparecchi detti “monitor”. Fino a poco tempo fa si trattava principalmente di monitor a tubo catodico CRT, mentre oggi non solo i computer portatili, ma anche i desktop utilizzano monitor a cristalli liquidi (LCD). Questi hanno eliminato alcune problematiche sanitarie legate all'uso dei monitor CRT e riguardanti le rilevanti emissioni di radiazione di questi ultimi.
In passato si utilizzavano anche monitor alfanumerici (cioè in grado di visualizzare solo caratteri), in bianco e nero o monocromatici. La qualità del monitor è in genere espressa dalla sua ampiezza (diagonale in pollici) e dal suo dot pitch (distanza diagonale tra i fosfori) che determinano la risoluzione numero di pixel rappresentati), dalla frequenza di refresh (numero di volte in cui l’immagine viene sostituita al secondo) e da altri parametri.
Per i monitor a cristalli liquidi, la corrente tecnologia detta TFT (thin film transistor) a “matrice attiva” consente di ottenere immagini stabili (l’”attivo” si riferisce a una memoria che mantiene la tensione e non fa sfumare le immagini tra un ciclo e l'altro) e di buona qualità. Siccome uno dei problemi della tecnologia LCD era la lentezza nella risposta dei cristalli, un altro parametri rilevanti per determinarne la qualità visiva su sequenze in movimento è la velocità di risposta, in genere dell'ordine della decina di ms. Altri parametri rilevanti sono la luminosità (misurata in candele/$m^{2,}$ cioè la luce che può emettere) ed il rapporto di contrasto (cioè il rapporto tre luminosità di pixel bianchi e neri, ad esempio 1000:1) Un altro parametro rilevante e che aveva valori limitati nei primi modelli è l'angolo di visuale (ovviamente più è ampio, e meglio è).

Scheda grafica

Ciò che si vede sul monitor, dipende da una scheda montata sul PC, detta scheda video o scheda grafica. Su di essa è in genere presente un convertitore digitale analogico (RAMDAC) che genera il segnale che viene trasferito poi al monitor (se il monitor ha ingresso analogico), a partire da una mappa digitale che viene “scritta” sulla memoria presente nella scheda stessa (in caso di connessione digitale, per ora non diffusissima, vengono trasferiti direttamente i bit). La dimensione di questa memoria presente nella scheda grafica determina:la risoluzione ed il numero di colori rappresentati (palette)

Figura 10: Scheda grafica

Ad esempio, per avere 320 x 200, 256 colori, ognuno (8,8,8) bit: occorrono 64.000 byte, 768 per la palette. Per avere 640 x 480, 16 colori: 307.200/2 = 153.600 byte. Per avere 1024 x 768, 16.777.216 colori (true color, ovvero 24 bit): 2.359.296 byte. Esistono degli standard per risoluzioni e colori, ad esempio CGA, VGA, SVGA (1024x768), XGA.

Mentre in origine la scheda serviva soltanto come buffer dell'immagine da visualizzare ed interfaccia con lo schermo, al giorno d'oggi è stato trasferito sulle schede tutto il peso computazionale della generazione ed elaborazione grafica, specialmente per quanto riguarda la rappresentazione di scene “virtuali” in tre dimensioni a partire da modelli.
Le schede video moderne, in pratica, contengono “Coprocessori Grafici” che “calcolano” la bitmap dei colori a partire da descrizioni della scena, togliendo questo compito dalla CPU. In pratica, una scena 3D viene “descritta” alla scheda mediante i poligoni che la compongono, la posizione delle luci, ecc. La descrizione delle scene ed i comandi all'hardware della scheda grafica vengono in genere realizzati dai programmatori utilizzando opportune librerie o interfacce di programmazione specificatamente create, le due interfacce standard si chiamano *OpenGL* e *DirectX.*
La potenza e la quantità di memoria gestita dalle recenti schede video e l'architettura parallela delle stesse (che permette di realizzare contemporaneamente calcoli su un'intera matrice di dati) fa sì che tali schede vengano ormai utilizzate non solo per la grafica, ma anche per compiti di calcolo generico.
La scheda grafica è in genere collegata al monitor per via analogica, con un connettore parallelo che trasporta le componenti video ed audio (VGA). I moderni monitor e schede grafiche tuttavia supportano spesso la connessione digitale (DVI), che permette di riprodurre esattamente il segnale di uscita della scheda sul monitor senza rischi di distorsioni od errori dovuti alla trasmissione sul cavo.

Tastiera

È un insieme di tasti, connessi ad interruttori. La circuiteria individua il tasto premuto ed invia il codice al sistema, che determina il carattere ASCII. E' possibile modificare la tabella di conversione via software per adattarsi a diverse lingue e convenzioni. La tastiera ha anche dispositivi di memoria (buffer) per risolvere problemi di ripetizione di tasti e di rimbalzo. Una curiosità: la disposizione “qwerty…” dei tasti non è, come si potrebbe pensare, studiata per accelerare la battitura, ma la si deve ai problemi meccanici delle prime macchine per scrivere, che presentavano problemi di tasti che si incastravano nel caso di battitura consecutiva di tasti vicini. Il produttore (Remington) decise allora di utilizzare una disposizione di tasti atta ad evitare, appunto, l’utilizzo consecutivo di tasti contigui. Quando il problema meccanico fu risolto, non era più possibile passare a tastiere più efficienti, nonostante esse fossero note a causa dell’ormai capillare diffusione delle tastiere “qwerty” ed all’adattamento ad esse dei dattilografi.

Sistemi di puntamento

Figura 11: Mouse a due tasti

Sono unità di input per sistemi grafici; differiscono per la tecnologia utilizzata e la differente ergonomia dell'interazione con l'operatore umano. Il più diffuso è sicuramente il mouse. Esso trasmette la variazione di posizione (coordinate x e y) e mediante la pressione di tasti invia prefissate sequenze di caratteri Esistono mouse di differenti tipi, a pallina di gomma a LED e fotorivelatore su tavoletta tramata, a LED senza tavoletta tramata. Quest’ultimo tipo si basa sull’acquisizione dell’immagine che sta sotto il mouse stesso da parte di un minuscolo sensore CCD (in pratica una telecamera). Il LED serve per illuminare la regione ripresa.
Sulla pallina di gomma che muove due rulli si basa anche la trackball, utilizzata da particolari applicazioni (es. postazioni pubbliche, alcuni vecchi portatili e alcuni

videogiochi. Anche il joystick si usa soprattutto per i videogiochi, e si basa su una leva che preme dei pulsanti quando inclinata, fornendo le indicazioni di spostamento. Il touchscreen è un monitor sensibili al tocco del dito, ma ha risoluzione in genere modesta, basandosi sul fatto che il dito interrompe due fasci ortogonali di luce infrarossa tra fotoemettitori e fonorivelatori. Si usa in ambito industriale e per postazioni dimostrative o servizi per il pubblico.

Stampante

Mentre per i primi personal computer erano disponibili stampanti lente rumorose e costose, la tecnologia moderna consente di trovare ottime stampanti a prezzi bassi. Ciò ha fatto in modo che, curiosamente, l'era dell'elettronica che si supponeva portare verso una riduzione nell'utilizzo di carta e documenti stampati, abbia invece portato un costante incremento nell'utilizzo di carta come è emerso da un recente studio commissionato da un noto produttore di stampanti. Tornando ai dispositivi, esistono varie tecnologie di stampa: Ad impatto: martelletto su nastro inchiostrato: qualità buona, ma lenta; pochi colori.

- Ad aghi: insieme di aghi; la qualità dipende dal numero di aghi, dalla loro distanza e precisione; lente ma grafiche; monocromatiche; costi bassi, ormai desuete come le precedenti.
- A getto d'inchiostro: gli aghi sono sostituiti da ugelli che spruzzano gocce di inchiostro; qualità migliore; possibilità del colore
- Termiche: comprendono diversi tipi: aghi che bruciano carta termo-sensibile, calore che fa evaporare sostanze che si depositano sulla carta
- Laser: un raggio laser forma l'immagine della pagina su un cilindro fotosensibile che si carica elettrostaticamente nei punti colpiti da maggior intensità. Sul cilindro si deposita il toner, che viene trasferito a caldo sulla carta.

I parametri che caratterizzano le prestazioni della stampante sono la risoluzione (dot per inch) e la velocità (pagine per minuto).

Il plotter

Si basa su un meccanismo che muove una o più penne su un foglio. Serve più che altro per cartografia e disegno architettonico, progetti, permettendo un elevata risoluzione delle linee.

Lo scanner

Serve per acquisire immagini raster. Ve ne sono di manuali, da passare sul foglio di carta e piani, dove il foglio viene posto sul piano dove sotto un vetro il sensore scorre linearmente. Ormai non esistono più scanner a passate multiple, come si usava una volta per migliorare la risoluzione. La risoluzione ottica è la reale densità di punti generati dal sensore; la risoluzione interpolata aumenta artificialmente la densità dei punti con un algoritmo matematico. Gli scanner vengono di solito associati a programmi per riconoscere i caratteri di testo (OCR) in modo da trasformare documenti cartacei in documenti elettronici.

Figura 12: Scanner piano

Modem

Il modem (modulatore/demodulatore) collega il computer alle reti mediante conversione D/A (A/D) e collegamento a provider attraverso la rete telefonica. La velocità tipica dei modem attuali è di 56 Kbit/s. Per i collegamenti via ISDN e ADSL si utilizzano schede che

sono spesso chiamate impropriamente modem, in quanto non operano una modulazione del segnale, ma trasmettono e ricevono segnali digitali.

Scheda audio
Genera gli effetti sonori: si occupa di convertire un flusso di dati digitali in un segnale analogico riproducibile da altoparlanti. Contengono spesso anche sistemi di acquisizione per convertire l'input analogico di un microfono in uno stream di bit. Le architetture moderne contengono un sintetizzatore di suoni (solitamente usato per generare suoni in tempo reale senza usare la CPU e chip sonori per migliorare la qualità del suono stesso. Parametri caratteristici delle schede audio sono il numero di canali (segnali distinti) che sono in grado di gestire (es. 1 = mono, 2.0 = stereo, 2.1 = stereo + subwoofer, dolby 5+1, ecc)

Memorie a supporto magnetico

Figura 13: Disco rigido

Quasi tutti i sistemi di memorizzazione precedenti al CD si basavano sulle proprietà di un materiale ferromagnetico come l’ ossido di ferro depositato su un supporto inerte. Una corrente positiva o negativa orienta il materiale che costituisce le areole. La lettura della polarizzazione si ottiene facendo transitare le areole sotto una spira. Si arriva a centinaia di bit per mm^2. La memorizzazione rimane in assenza di alimentazione. Si differenziano attraverso la forma del supporto : dischi flessibili, dischi rigidi, nastri. Il floppy disk ha supporto flessibile, il disco rigido ha supporto metallico.
Entrambi hanno tracce concentriche e ogni traccia è suddivisa in settori mediante la formattazione
Si può accedere in lettura o scrittura solo ad un intero settore, pertanto il trasferimento avviene a blocchi. I Floppy hanno dimensioni tipiche 5"1/4 o 3"1/2, e contengono fino a 1.44 Mb; hanno un elevato tempo di accesso, capacità limitata
Gli hard disk sono immersi in gas pressurizzato, hanno alta velocità ed alto *tranfser-rate*; in un PC possono esserci anche più dischi in parallelo, collegati al bus di sistema.
I dischi rigidi hanno particolare importanza perché vengono in genere utilizzati per la memorizzazione off line di tutti i dati e programmi dei PC (memoria di massa). La loro capacità di memorizzazione è cresciuta notevolmente negli anni fino agli attuali centinaia di Gigabytes.
Per accedere ai dati devono prima posizionare il braccio nella posizione corretta (*tempo di ricerca*) e poi girare per portare la testina sul settore esatto, questo implica un tempo di *latenza*. Il dato da valutare complessivamente è il *tempo medio di accesso*, dipendente dai parametri precedenti, indice del tempo necessario per ciascun accesso a settore contenente dati (ogni apertura di file ne implica molte). Negli hard disk Ide-Ata i tempi di accesso sono attorno agli 8 millisecondi (ms) mentre le unità di tipo SCSI raggiungono anche i 5 ms.
Il *transfer rate* descrive poi la velocità di trasferimento dati in MB/s, che è proporzionale alla velocità di rotazione (tipicamente 7200 rpm).
Abbiamo detto che l’unità di memoria trasferibile coincide con la dimensione del settore.

Se la quantità di dati dell'unità logica di informazione (che può essere un testo, un programma, un immagine, ecc.) da trasferire supera la capacità del settore, i programmi di gestione del sistema operativo devono spezzettare i dati e concatenarli. Una catena si dice *file.* In Windows esiste un indice del disco (FAT) che contiene le informazione sui settori allocati e liberi. Per i sistemi server, poiché i dischi rigidi sono delicati e possono rompersi in caso di urto o spegnimento/accensione brusca, si usa duplicare l'informazione (mirroring, RAID) su più dischi.
Sui nastri magnetici il materiale magnetico è disposto su nastri di plastica avvolti su bobine e le informazioni memorizzate in blocchi intercalati da zone non magnetizzate. L'accesso è sequenziale, per cui molto lento. Per questo le applicazioni tipiche sono backup o archiviazione dati

Memorie di tipo ottico

I dischi ottici sono derivati dai CD usati per riproduzioni audio. La tecnologia si basa su deformazioni permanenti della superficie del supporto (materiale plastico) prodotte da raggi laser. Le variazioni di tensione accumulata su un fotorivelatore consentono di ricostruire l'informazione. Ne esistono di varie tipologie. I tipici CD/R sono memorie WORM (Write Once Read Many), esistono però anche I CD/RW, riscrivibili. La tipica capacità è 640 Mbyte, che corrisponde, per i CD audio, a 74 minuti. Oggi sono la forma di memorizzazione su mezzo rimovibile più diffusa per l'elevata affidabilità, la grande capacità ed il costo irrisorio. Negli apparecchi di scrittura/lettura si indica in genere la velocità con cui possono essere letti/scritti in funzione dei primi modelli (4x, 16x, 32x). La moderna tecnologia laser ha consentito di realizzare su supporti simili i DVD, che possono contenere fino a 17 Gb e vengono utilizzati anche per la memorizzazione di film. Anche in questo caso esistono differenti standard, quelli oggi introdotti sui PC consentono di memorizzare e leggere tipicamente 4.7 Gb.

Memorie a stato solido

Recentemente sono diventati di largo uso oltre che di basso costo e di notevole capacità le memorie permanenti e riscrivibili basate sulla fisica dei semiconduttori, generalmente riferite come memorie "flash", evoluzione di diversi tipi di dispositivi di memoria programmabile a stato solido.
Utilizzando principi della meccanica quantistica, questi dispositivi possono essere scritti e letti permanentemente, conservano i dati senza alimentazione e sono di dimensioni contenute. Sono per questo molto adatte per l'interscambio di dati (per cui hanno sostituito i vecchi dischetti magnetici) ed i dispositivi portatili come macchine fotografiche, cellulari, media player, ecc.
Esistono vari formati di scheda di memoria Attualmente esistono varie tipologie di flash memory card che si differenziano per il consorzio di produttori e standard di dimensioni, collegamenti e capacità: Compact Flash, SmartMedia, MultiMediaCard, Memory Stick ,Secure Digital, TransFlash, xD-Picture Card.
Particolare interesse per il personal computing hanno le cosiddette USB memory Flash, ovvero le ben note "chiavette USB" che sono appunto accessibili dal pc in lettura/scrittura mediante l'interfaccia USB.
Calcolatori che non hanno necessità di grosse quantità di memoria di massa possono addirittura sostituire l'hard disk con memorie flash (solid state disks), come accade per esempio nei recenti ultraportatili (Asus EEE, Acer Aspire One).

1.10 Storia del calcolatore

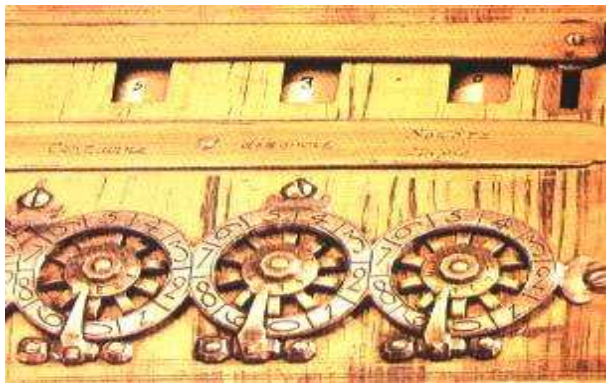

Figura 14: Pascaline

Come abbiamo visto il moderno computer non è altro che una sofisticata macchina per il calcolo automatico, che è in grado di eseguire ripetutamente operazioni aritmetiche e logiche.
Le prime macchine per il calcolo automatico furono apparecchiature meccaniche. Il primo esempio che si ricorda è quello costruito nel 1642 dal matematico e filosofo francese Blaise Pascal, la cosiddetta Pascaline. Essa era in grado di eseguire le addizioni tenendo conto automaticamente dei riporti ed eseguiva sottrazioni e moltiplicazioni. Negli anni seguenti altri studiosi proposero modifiche o nuove macchine, ma spesso solo da un punto di vista teorico, ricordiamo Leibnitz (1671) e Poleni (1709).
Il primo progetto in qualche modo definibile come antenato del calcolatore moderno fu la macchina analitica (Analytical Engine) progettata dal matematico inglese Charles Babbage nel 1832. Essa poteva, infatti, risolvere teoricamente qualsiasi problema descrivibile come una sequenza d’operazioni matematiche. Teoricamente, perché Babbage non riuscì a realizzarla in pratica a causa delle limitazioni della tecnologia del tempo, soltanto dopo la sua morte il figlio ne realizzò alcune parti.

Il primo calcolatore prodotto in serie risale al 1810 ad opera di Charles Thomas, ne furono costruiti circa 1.500 esemplari.
La prima applicazione pratica del calcolo automatico risale invece al 1890, per il censimento di New York e Baltimora quando lo statistico americano H. Hollerith utilizzò schede perforate per registrare i dati; le schede venivano lette da dispositivi ed elaborate meccanicamente. L’uso di questa macchina ridusse il tempo di elaborazione delle schede dai sette anni del censimento precedente a due anni e mezzo, nonostante l’aumento dei dati.

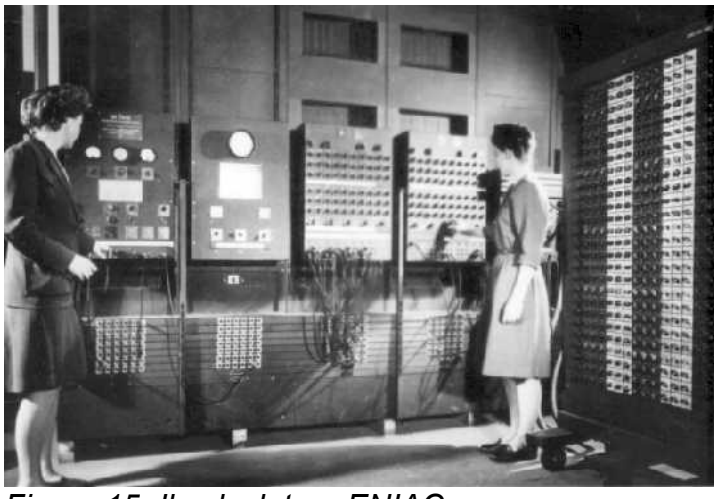

Figura 15: Il calcolatore ENIAC

La scheda perforata era stata introdotta da Jacquard per comandare i telai di tessitura. Hollerith fondò anche una società per la produzione dei lettori di schede perforate; la società prese il nome di IBM (International Business Machine).
Un altro grosso passo avanti nella storia delle macchine di calcolo lo si deve all’ingegnere tedesco Konrad Zuse che che realizzò nel 1938, calcolatrice meccanica in cui introdusse l;uso del sistema binario ed il primo calcolatore elettromeccanico: lo Z3 (1941) ottenuto sostituendo i dispositivi meccanici con relais.
Lo Z3 era in grado di eseguire le quattro operazioni elementari e di estrarre la radice

quadrata con tempi di esecuzione delle operazioni dell'ordine del secondo.
Nel 1944 venne realizzato all'Università di Harvard il Mark1, calcolatore elettromeccanico capace di eseguire un'addizione in 300 millisecondi e di calcolare alcune funzioni quali quelle trigonometriche, il logaritmo decimale e l'esponenziale.
Solo nel 1946 fu realizzato il primo calcolatore elettronico presso l'Università di Pennsylvania: chiamato ENIAC (Electronic Numerical Integrator and Calculator), si basava su valvole termoioniche (per la memoria dati) ed era mille volte più veloce dei suoi predecessori.
Sull'ENIAC i programmi erano costituiti da un insieme di connessioni tra fili; era quindi molto difficile programmare compiti piuttosto complessi.
Da notare che ENIAC pesava 30 tonnellate e occupava una superficie di oltre 100 metri quadri, nonostante ciò la memoria era di 20 numeri di 10 cifre.
Nel 1950 Fu realizzato il primo calcolatore a programma memorizzato, l'EDVAC (Electronic Discrete Variable Automatic computer) ove furono applicate l' idea di John Von Neumann, matematico ungherese, secondo cui il programma poteva essere posto nella memoria dati del calcolatore.
I computer che seguono questa architettura, cioè tutti I calcolatori moderni vengono anche chiamati macchine di Von Neumann e il ciclo di esecuzione delle istruzioni viene chiamato ciclo di Von Neumann.
Il primo calcolatore prodotto su scala industriale fu l'UNIVAC1(1951),
Nel 1955 l' IBM presenta il modello IBM-704 con 32000 byte di memoria e la possibilità di programmare anche in un linguaggio di alto livello (Fortran).
La cosiddetta seconda generazione dei calcolatori, risale alla fine degli anni 50 è caratterizzata dalla sostituzione delle valvole con i transistor, frutto delle ricerche dei Bell Labs alla fine degli anni 40. I transistor consentirono di ottenere computer più piccoli, affidabili e potenti.
In questo periodo cominciarono a nascere ed evolversi I "sistemi operativi", per rendere semplice ed efficiente l'uso dei calcolatori.
La terza generazione dei calcolatori vede invece l'introduzione dei circuiti integrati (chip). Il circuito integrato fu inventato da J. St. Clair Killey della Texas Instruments nel 1958, il primo calcolatore commerciale a circuiti integrati fu l'IBM serie 360.
In questo periodo vengono introdotti alcuni concetti base dei moderni sistemi operativi (vedi Capitolo 4) come la multiprogrammazione e il time-sharing.
La quarta generazione inizia nel 1971 con il primo microprocessore, il 4004, realizzato dall'italiano Federico Faggin presso la INTEL. Un microprocessore è una unità di elaborazione centrale (CPU) realizzata in un unico chip, di costo molto basso.
In questo periodo vengono introdotte anche le memorie a semiconduttori meno ingombranti, più veloci e meno costose delle precedenti memorie a nuclei.
Nel 1976 nasce il primo personal computer: l'Apple I, che utilizzava il microprocessore 6502 della MOS Technology.

Nel 1980 L' IBM introdusse sul mercato il PC-IBM (Personal Computer) con microprocessore 8088 Intel e 64 Kbyte di memoria. Il PC-IBM utilizzava il sistema operativo DOS che divenne poi uno standard mondiale per i personal computer.

Nel 1981 venne annunciato il primo laptop computer, l’Epson HX-20 della Epson, di peso inferiore a un chilo e mezzo, dotato di un piccolo display a cristalli liquidi. Da allora la potenza dei calcolatori e la complessità dei sistemi operativi non ha smesso di crescere, anche se non ci sono state, in realtà, rivoluzioni clamorose nelle architetture di calcolo.

Figura 16: Apple I, il primo Personal Computer

Vi è stata una netta convergenza nell’evoluzione di workstation scientifiche e professionali e computer domestici, che ormai sono sostanzialmente indistinguibili. Dal punto di vista dell’utilizzo del calcolatore la più grande rivoluzione la si e avuta con l’avvento dell’era di internet, in cui I calcolatori sono collegati in una rete mondiale di comunicazione e possono scambiarsi dati e programmi.

Nonostante ciò l’annunciato avvento del “network computing” con calcolatori privi di memoria a lungo termine che utilizzano solo programmi scaricati dalla rete, non si è verificato. Le tendenze più recenti riguardano l’evoluzione dei PC portatili, ormai paragonabili alle macchine desktop, ed all’evoluzione delle tecnologie wireless, che sta portando ad una convergenza tra i sistemi di elaborazione e quelli di telefonia.

Inoltre sta assumendo importanza sempre più rilevante il paradigma del cosiddetto “Grid Computing”, consistente nel consentire l’utilizzo di dati e risorse di calcolo distribuite in luoghi diversi a ciascun utente, che resta ignaro di dove queste si trovino, ma potendo così avere a disposizione risorse enormi per il calcolo.

2 Calcolo automatico, algoritmi, programmi

Pur avendo visto come è fatta la macchina, almeno per quanto riguarda l'hardware (l'insieme dellle componenti fisiche) siamo ancora ben lontani dal capire come sia possibile che essa consenta tutte quelle incredibili applicazioni che abbiamo oggi a disposizione. Per il momento, infatti, come era per i primi calcolatori, l'utilizzo di questa architettura si limiti all'esecuzione di una serie di operazioni elementari, aritmetiche e logiche sui dati memorizzati che avvengono secondo un programma prestabilito.
Tutto molto distante dai problemi che risolviamo con i nostri computer. Per trasformare il nostro sistema in qualcosa che ci serva nella vita reale, abbiamo da compiere alcuni fondamentali passi: codificare i dati che ci interessa elaborare in numeri binari ed implementare le operazioni che dobbiamo fare su di essi a partire da quelle fondamentali del mircroprocessore.
E, ovviamente, sapere come quali siano le metodologie ed i limiti dell'elaborazione automatica dell'informazione cosa che viene studiata dalla teoria degli algoritmi.

2.1 Algoritmi e programmi

Per quanto abbiamo visto il calcolatore è un elaboratore automatico d'informazione, che esegue operazioni su oggetti (dati) per produrre altri oggetti (risultati).

L'esecuzione di azioni viene richiesta all'elaboratore attraverso opportune direttive, dette istruzioni. Il calcolatore che abbiamo descritto è detto "elettronico" in quanto gli oggetti che riceve in ingresso, elabora e fornisce in uscita sono segnali elettrici, livelli di tensione e la loro elaborazione avviene mediante circuiti elettronici a semiconduttore.
Si parla anche di elaboratori digitali: il significato che viene dato a questi livelli di tensione non è legato, infatti, al loro valore assoluto, ma al loro essere superiori od inferiori a una determinata soglia, nel qual caso al livello di tensione si attribuisce il valore 1 o 0. I dati in ingresso ed in uscita nei calcolatori elettronici sono quindi sequenze di valori 0 o 1 (bit, da BInary digiT).

Al di là della complessità raggiunta da problemi e soluzioni che stanno alla base del software che utilizziamo oggi, anche le macchine che usiamo quotidianamente lavorano facendo sequenze di operazioni su dati binari.
Uno degli scopi dell'informatica consiste quindi nel cercare di capire quali siano i problemi affrontabili dalle macchine e lo studio di come si possa impostare la loro soluzione suddividendola in passi fondamentali eseguibili dalla macchina. Si parla in questo caso di studio degli algoritmi: ne riassumiamo nel seguito gli aspetti principali.

Supponiamo di avere un determinato problema e di volerlo fare risolvere automaticamente alla macchina calcolatrice che abbiamo a disposizione. Per far questo dovranno essere verificate alcune condizioni:

- La soluzione del problema deve essere nota.
- I passi che la compongono debbono poter essere tradotti in un linguaggio comprensibile dal calcolatore.
- I dati in ingresso devono essere codificati in dati adatti al calcolatore (in generale numeri binari).

Se ciò è vero, possiamo trovare opportuno metodo di elaborazione che mettendo in sequenza "azioni elementari" eseguibili dalla macchina trasformi i dati iniziali nei corrispondenti risultati finali desiderati.
La "sequenza di azioni elementari" così definita è detta anche *algoritmo.*

Non tutti i problemi, neppure tutti quelli, ben definiti, della matematica o della geometria possono essere risolti o tradotti per il calcolatore. Questo può essere dovuto al fatto che non si conosce neppure la soluzione del problema (ad esempio trovare tutte le coppie contigue di numeri primi) oppure perché non esiste un metodo automatizzabile.
Questo non è stato certo limitante per il diffondersi delle applicazioni del calcolo automatico: con il calcolatore è stato comunque possibile risolvere problemi molto complessi ed oggigiorno i personal computer non si limitano a risolvere qualche semplice problema matematico o logico, ma sono programmati per simulare o riprodurre operazioni complesse e sostituiscono l'attività umana in una serie di compiti sempre più rilevante.
Tutto questo non deve farci però dimenticare che tutte le attività svolte dal calcolatore nascono sempre dalla scrittura di algoritmi in opportuni linguaggi descrittivi (quella che si chiama "programmazione"), che vengono poi "tradotti" nei linguaggi a basso livello del calcolatore mediante quelle che vengono chiamate "compilazione" o "interpretazione".
Dunque non deve stupire che magari qualche volta i programmi che utilizziamo falliscano o si blocchino, o magari ci siano errori nelle elaborazioni, la colpa, contrariamente a quanto viene talvolta detto da non addetti ai lavori, non è mai del computer, ma di errori nella scrittura degli algoritmi, nella programmazione o nella selezione dei dati.
Il lavoro degli informatici cerca di ridurre questi errori e quindi discipline fondamentali per chi studia l'informatica sono sicuramente lo studio degli algoritmi, dei linguaggi e delle tecniche di programmazione. Senza addentrarci in queste materie, ricordiamo soltanto le proprietà principali degli algoritmi.

Il concetto di algoritmo (termine che deriva dal nome di un matematico arabo, Al Khowarizmi, che contribuì alla fondazioni dell'algebra) non è legato necessariamente al calcolatore elettronico, ma ad un qualunque "esecutore". Quest'ultimo è un entità che, in generale, dovrà essere in grado di compiere "azioni elementari" e di interpretare un linguaggio con cui queste azioni gli sono ordinate.
Per il resto, l'esecutore di un algoritmo può benissimo essere una persona, una sistema meccanico o di qualunque altro tipo. Se, per esempio, istruisco un mio conoscente ad andare a pagare una bolletta al mio posto, spiegandogli dettagliatamente dove deve andare, cosa deve scrivere, eccetera, sto facendo null'altro che istruire un esecutore a eseguire una serie di operazioni elementari che lui sa eseguire, cioè gli sto fornendo un algoritmo.
L'importante, ai fini dell'ottenimento dello scopo (nell'esempio avere la bolletta pagata) è che la mia descrizione dell'algoritmo possieda alcune proprietà fondamentali, cioè, ad esempio:

- Non ambiguità: le istruzioni devono essere univocamente interpretabili dall'esecutore
- Eseguibilità: l'esecutore deve essere in grado di operare ciascun passo in tempo finito
- Finitezza: l'esecuzione deve terminare per ogni configurazione dei dati in ingresso

Queste proprietà sono necessarie sia nel caso della bolletta, sia per la programmazione del computer. Se le proprietà sono rispettate, cosa peraltro non facilmente dimostrabile nei problemi reali, posso mandare tranquillamente il mio compito in esecuzione.
Nello studio teorico si considerano, per semplicità, delle macchine "astratte" che in qualche modo modellano le macchine calcolatrici reali; le più note sono Macchina di Turing e la Macchina RAM (con memoria ad accesso casuale). Dall'analisi di tali modelli di esecutore si possono ricavare teoremi riguardanti la calcolabilità di determinate

espressioni, e quindi sulla risolubilità di problemi. Fu per esempio l'analisi di un famoso problema sulla macchina di Turing a mettere in evidenza che nella logica matematica esistono classi problemi "non decidibili", per i quali non si può trovare soluzione. Una congettura comunemente accettata è che la classe di problemi matematici risolubili con una macchina astratta sia comune per tutte le macchine di questo tipo.

Non ci addentreremo nello studio di queste problematiche, per le quali rimandiamo a testi specifici, e andiamo invece ad analizzare come si possa passare da un problema risolubile alla sua esecuzione automatica, cioè analizziamo un attimo quello che gli analisti e programmatori fanno nella pratica del loro lavoro.

2.2 Analisi ed implementazione degli algoritmi

L'analisi del problema deve mettere in evidenza quali siano i dati, quale la soluzione da ottenere e con quali vincoli. Si definiscono cioè le precondizioni rispettate dai dati e le postcondizioni che devono rispettare le soluzioni. A questo punto si passa alla sintesi dell'algoritmo, che può essere libera o eseguita secondo ben codificati metodi progettuali. Una volta nota la soluzione la si dovrà descrivere in maniera chiara e comprensibile all'uomo per permettere ad altri di analizzarla e modificarla, per poi, in un secondo tempo "tradurla" in un linguaggio comprensibile al calcolatore. Il passaggio dalla descrizione dell'algoritmo alla scrittura e traduzione del programma viene spesso chiamato "implementazione".
Facciamo un esempio: supponiamo di voler scrivere un programma che calcola la media dei voti riportati da uno studente. Dovremo, oltre a fare in modo di codificare in bit i numeri reali e i caratteri che compongono il nome, saper risolvere il problema di calcolare la media (banale: somma diviso numero di esami) e scrivere le operazioni che portano ad essa in modo da farle eseguire al calcolatore. Tutto qui, anche se, ovviamente, se dovessimo realizzare tutto noi avendo a disposizione soltanto l'hardware della macchina sarebbe complicatissimo.

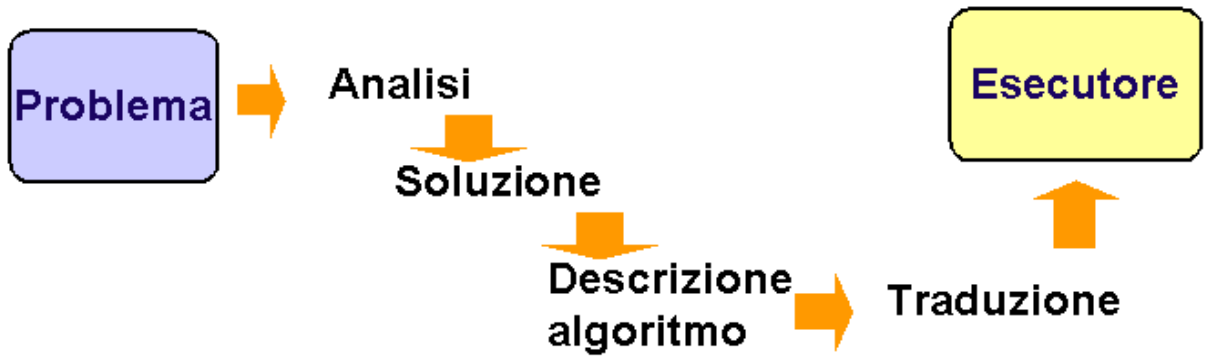

Figura 17: Dal problema allo sviluppo di programmi

Un calcolatore elettronico, di per sé, comprende soltanto serie di istruzioni per il microprocessore scritte in quello che viene detto "linguaggio macchina", cioè una serie di codici di operazioni (una sequenza di byte) e di operandi (dati scritti anche essi in binario). Ed è in grado di lavorare su dati numerici di tipo e dimensione limitati. Abbiamo visto che in un calcolatore elettronico la CPU ha un set di istruzioni a ciascuna delle quali corrisponde un codice: questo vuol dire che un programma eseguibile lo è soltanto per quella determinata CPU, o per una CPU con le stesse operazioni e gli stessi codici. Un

programma eseguibile, in linguaggio macchina, è eseguibile solo su un calcolatore con una determinata CPU e consiste quindi in una sequenza di codici operativi (numeri binari che corrispondono a una determinata operazione della CPU) e di operandi (dati codificati in codice binario. Poiché una simile sequenza non è molto comprensibile, la si indica con un linguaggio simbolico in cui stringhe di caratteri indicano le operazioni e le locazioni di memoria (es. ADD R1 R3 significa addizione dei contenuti di due registri).
Programmare un applicazione moderna lavorando a questo livello avrebbe comunque sempre una complessità proibitiva.
Fortunatamente i programmatori, al giorno d'oggi, non hanno bisogno di conoscere neppure questo linguaggio, in quanto esistono linguaggi evoluti attraverso i quali si possono scrivere gli algoritmi che poi vengono trasformati in linguaggio macchina dai compilatori, oppure controllati e interpretati linea dopo linea da programmi che vengono detti "interpreti". Questi linguaggi di programmazione vengono detti "di alto livello" (Pascal, Basic, C, C++, ecc) e ricordano più il linguaggio naturale umano che il linguaggio macchina.
Utilizzando questi linguaggi i programmatori lavorano a livelli "alti" del sistema operativo (vedi capitolo 4) e danno istruzioni a "macchine virtuali" in cui l'hardware è utilizzato indirettamente attraverso le funzioni del software di sistema.

Il compito dei programmatori di oggi è facilitato anche dall'avere a disposizione ulteriori strumenti, come ambienti di sviluppo, sistemi di "debugging" (per trovare facilmente gli errori), "editor" per scrivere il codice in maniera assistita, eccetera.
Inoltre moltissimi tipi di dati, anche immagini, suoni, ecc possono essere codificati in modo standard per essere elaborati dalla CPU e le librerie di programmi di sistema (incluse nei cosiddetti sistemi operativi) si occupano di gestire le parti del computer a basso livello (memoria, input/output, ecc.)
Insomma generazioni di programmatori hanno fatto sì che chi vuole realizzare applicazioni per il computer possa farlo ad alto livello, scrivendo liste di comandi complessi su dati di ogni genere e questo ha facilitato moltissimo lo sviluppo di applicazioni varie per l'utente generico.

Le macchine che oggi utilizziamo non si caratterizzano quindi soltanto per un'architettura hardware, ma anche per vari tipi di programmi di gestione che costituiscono i cosiddetti sistemi operativi (ci si riferisce a hardware + sistema operativo con il termine *piattaforma* di sviluppo). Questi sistemi operativi forniscono l'interfaccia all'utente per interagire con le risorse della macchina in maniera astratta ed intuitiva, ed ai programmatori la possibilità di scrivere i programmi che realizzano applicazioni complesse mediante metodologie di sviluppo semplici ed organizzate.

3 Codifica dell'informazione

Abbiamo stabilito che un calcolatore elettronico è una macchina per eseguire sequenze di operazioni elementari che risolvono problemi ben formulati (algoritmi). Esso si basa su circuito a semiconduttori ed all'ingresso accetta dati digitali cioè sequenze o insiemi di stati di alta/bassa tensione, in pratica numeri binari. L'unità fondamentale di informazione, cioè un singolo valore 0 o 1, prende il nome di BIT (BInary digiT). Sui questi dati tutte le operazioni eseguite vengono realizzate a partire da operazioni booleane (AND, OR, NOT…) implementati mediante circuiti a transistor.
Tutte le informazioni introdotte nei calcolatori devono quindi essere trasformati in sequenze binarie prima delle elaborazioni. In generale non lo sono. All'interno dell'elaboratore rimangono sempre memorizzate in forma binaria (valore alto/basso di tensione, interruttore aperto/chiuso, zona magnetizzata/non magnetizzata, riflettente/non riflettente, ecc.). Le memorie dei calcolatori immagazzinano sequenze binarie sotto queste forme. Per dare gli ordini di grandezza della quantità di informazione si utilizzano i vari multipli del bit, ad esempio, nell'ordine:

- Byte: 8 bit
- Kbyte (Kilo)= 210byte = 1024 byte
- Mbyte (Mega)= 1024 Kbyte
- Gbyte (Giga)= 1024 Mbyte
- Tbyte (Tera)= 1024 Gbyte

Numeri, lettere, immagini e tutte le informazioni che vogliamo elaborare automaticamente vengono anch'esse comunemente rappresentate mediante sequenze di simboli (che in termini tecnici prendono il nome di stringhe) appartenenti ad un alfabeto di dimensioni però ben maggiori di 2 (il caso dei simboli binari rappresenta semplicemente il caso minimale, cioè il numero minimo di simboli sufficiente per rappresentare informazione).
La codifica dei dati in binario consisterà quindi nel creare una corrispondenza biunivoca tra l'alfabeto da codificare e stringhe di "bit" 0 e 1.
poiché con N bit si possono creare 2^N stringhe differenti, vorrà dire che per codificare con stringhe binarie un insieme di simboli (alfabeto) di C elementi (C prende il nome di cardinalità dell'alfabeto) dovremo utilizzare un numero di bit tali che 2^N sia maggiore o uguale a C.
Se $2^N = C$ la codifica si dice non ridondante, altrimenti la si dice ridondante.
La ridondanza (cioè l'uso di un numero di bit maggiore dello stretto necessario) non è necessariamente uno spreco, ma può servire per rilevare o correggere errori. In effetti sia nelle memorie che nelle trasmissioni in rete di dati digitali, può accadere che il segnale di tensione che codifica lo 0 o l'1 possa essere errato a causa di qualche difetto. Se si usano bit di controllo o si duplicano dati il rischio che questi errori possano causare problemi è ridotto.
Un metodo di rilevazione di errori molto usato, ad esempio nelle memorie è il controllo di parità. Ai bit necessari (cioè non ridondanti) per codificare si aggiunge un ulteriore cifra binaria tale che il numero di '1' sia pari (parità pari) o dispari (parità dispari). Un eventuale cambiamento di valore di un bit dovuto a malfunzionamento della memoria potrebbe essere quindi rilevato.
Facciamo un esempio: il valore in memoria è 1101101 (7 bit). Aggiungo il bit di controllo: deve essere "1" se voglio che il numero di uno sia pari: lo aggiungo in fondo e il mio dato diventa 11011011. Se ora un bit cambiasse valore, per esempio il quarto per cui il mio numero diventa 11001011, quando faccio una verifica noto che il numero di "1" è dispari, c'è quindi stato un errore. Ovviamente il controllo di parità definito così rileva, ma non

permette di correggere gli errori. Accorgersi però del problema può essere sufficiente per evitare danni.

3.1 Codifica dei caratteri alfabetici

Un esempio di codifica consistente nella semplice corrispondenza simbolo-stringa binaria è quello dei caratteri dell'alfabeto.

32	33	34	35	36	37	38	39
	!	"	#	$	%	&	'
40	41	42	43	44	45	46	47
(	)	*	+	,	-	.	/
48	49	50	51	52	53	54	55
0	1	2	3	4	5	6	7
56	57	58	59	60	61	62	63
8	9	:	;	<	=	>	?
64	65	66	67	68	69	70	71
@	A	B	C	D	E	F	G
72	73	74	75	76	77	78	79
H	I	J	K	L	M	N	O

Tabella 1: Codice ASCII

Nei calcolatori essi vengono spesso rappresentati mediante sequenze di 7 bit, utilizzando il codice ASCII (American Standard Code for Information Interchange), riportato in Tabella 1. 7 bit corrispondono a 128 simboli diversi, che spesso non bastano a rappresentare tutti i simboli (ci sono caratteri particolari dipendenti dal paese, ecc.) per cui esiste un ASCII esteso (8bit) ove peraltro l'ottavo bit può essere usato anche per il controllo di parità.
Per gestire vocabolari più ampi si sono poi evolute codifiche con un maggior numero di bit che consentono di rappresentare molti caratteri speciali (e per esempio alfabeti di varie lingue, anche con migliaia di caratteri come il cinese).
Il sistema attualmente supportato dalle applicazioni web e da molti sistemi operativi è detto Unicode e permette di utilizzare fino a 21 bit permettendo anche rappresentazioni efficienti che utilizzano l'utilizzo di codici più compatti per i caratteri più frequenti.

3.2 Sistemi numerici

Per quanto riguarda la rappresentazione dei numeri per l'elaborazione al calcolatore, se ci potessimo limitare ai numeri naturali basterebbe convertire decimale in binario e potremmo lavorare.

Decimale e binario sono, lo ricordiamo, notazioni posizionale in base 10 e 2;. notazione posizionale in base r significa che il numero è espresso da stringhe di un alfabeto di r cifre (che formano un insieme d={0,1,…,r-1}) nelle quali la posizione delle cifre è determinante: scrivere cioè:

$$N = xyz$$

significa che

$$N = x*r^2+yr+z$$

dove x, y, z corrispondono a cifre dell'insieme d={0,1,…r-1}. È importante notare che qualsiasi sistema posizionale a base fissa è irridondante in quanto ad ogni combinazione di cifre diversa corrisponde effettivamente un numero diverso.

La numerazione che utilizziamo comunemente è semplicemente una notazione posizionale in base 10, cioè con r=10, d={0,1,2,3,4,5,6,7,8,9}.

Il sistema binario usato all'interno del calcolatore è un caso particolare di notazione posizonale,con r=2 d={0,1}.

Altri sistemi usati in informatica sono l'ottale (r=8, d ={0,1,2,3,4,5,6,7}) e l'esadecimale: (r=16,d= {0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F})

Nel caso dei numeri interi, quindi, i numeri in base 10 devono essere convertiti in e da base 2 per l'elaborazione al calcolatore. L'operazione è molto semplice: ad esempio, utilizzando la definizione:

$$1010_2 = (1*8 + 0*4 + 1*2 + 0*1)10 = (8+2)_{10} = 1010$$

Un poco più complesso, ma sempre molto semplice è il passaggio inverso. Un metodo consiste nel dividere il numero per 2, mettere il resto come prima cifra del valore binario, dividere il risultato ancora per 2 e mettere il resto come cifra più significativa del risultato e dividere ancora il resto, e così via, finché il risultato non è 1. Provare per credere.
Nel rappresentare i numeri interi sul calcolatore, occorre tenere conto del fatto che le macchine hanno vincoli spaziali: esiste quindi un massimo valore rappresentabile, dipendente dal numero di bit utilizzati. Con n bit si può rappresentare al massimo il numero 2^n-1. Al contrario, se si vuole rappresentare un numero intero X, è necessario avere a disposizione un numero di bit pari a n = INT($\log_2$ (X+1)).
Nel sistema binario, le operazioni aritmetiche si fanno esattamente come nel sistema decimale,. Ad esempio la somma si fa cifra a cifra con il riporto alla posizione successiva:

```
0 + 0   =  0
0 + 1   =  1
1 + 0   =  1
1 + 1   =  0              con riporto di 1
1 + 1 + 1 =  1            con riporto di 1
```

Nell'esempio sotto vediamo la stessa operazione eseguita in binario e decimale:

10010110	150_{10}
+11011010	$+218_{10}$
_____	_____
101110000	368_{10}

Abbiamo detto che i numeri interi rappresentati in binario sul calcolatore hanno limiti, superiore ed inferiore. Se le operazioni che facciamo devono generare numeri al di fuori da questi limiti, il risultato del calcolo automatico sarà errato: gli errori di questo tipo si chiamano Overflow e Underflow, a seconda del fatto che il risultato sia maggiore del limite superiore o inferiore al limite più basso.
Si può avere overflow se il risultato delle operazioni richiede un numero maggiore di bit di quanto disponibile: ad esempio se abbiamo 4 bit, ed eseguiamo la seguente operazione,

otteniamo un risultato errato:

```
1100 + 0111 = 1011
```

che in decimale significa

```
12+7=11
```

Per avere il corretto risultato occorrerebbe un bit in più:

```
01000 + 00111 = 110011
```

Gli errori di overflow e underflow vanno chiaramente tenuti in conto dai programmatori per evitare di ottenere risultati completamente errati e devono quindi essere sempre prevenuti.

Per passare dai numeri interi ai naturali (negativi compresi, quindi), esistono diverse possibilità: la più intuitiva è utilizzare modulo e segno: Il bit più significativo non conta per il numero ma solo per il segno: positivo (0) e negativo (1).
Ad esempio con 5 bit:

```
10011 = -3
00011 = +3
```

Questo metodo ha un problema evidente:esistono due rappresentazioni per lo '0': ad esempio con 5 bit 10000 e 00000. Inoltre la realizzazione delle operazioni non è banale: addizionando bit a bit +3 e –3 si otterrebbe 10110, che è interpretato come –6.
Conviene usare il *complemento a 2*: se il numero di bit a disposizione è *n*, col complemento a 2 si rappresentano i numeri tra $-2^{(n-1)}$ e $2^{(n-1)} - 1$. Come? I numeri positivi nel range di interesse (cioè inferiori a $2^{(n-1)} - 1$) si rappresentano normalmente, invece si sostituisce -B con $(2^n - B)$.
La cosa interessante è che così la differenza tra due numeri positivi A e B, si esegue sommando con le regole del calcolo binario $A+(2^n - B)$
La sottrazione si esegue mediante una somma. Nel caso precedente, per +3 e –3 avremmo 00011 e 11101, sommano bit a bit otteniamo 00000, come volevamo.
La codifica dei numeri diventa ovviamente problematica quando si tratta di rappresentare nel calcolatore i numeri reali. E' evidente che non si possono rappresentare numeri che variano nel continuo con stringhe binarie, per cui quello che normalmente si rappresenta sono in realtà numeri frazionari che approssimano con un errore noto il numero reale interessato.
L'idea più semplice sarebbe di usare un bit per il segno, un numero fisso N per il numeratore ed un numero D di bit per il denominatore. Questa soluzione presenterebbe il problema di limitare molto o il massimo numero rappresentabile o il numero di cifre significative del numero espresso in forma decimale, a seconda che si utilizzino più bit per il denominatore o per il numeratore.
La soluzione al problema è stata trovata con l'utilizzo della rappresentazione in *virgola mobile (floating point)*.
In questa rappresentazione i numeri espressi nella forma:

$$A = M * B^E$$

dove M viene detta mantissa, B: base ed E esponente.
Sono necessari un segno sia per la mantissa che uno per l'esponente (esponenti

negativi=numeri frazionari): si sceglie arbitrariamente una forma normalizzata ove $0 \leq M < 1$, l'esponente è espresso in complemento a 2 (in binario), la mantissa è espressa in modulo e segno. Effettuare le operazioni aritmetiche su questi numeri è abbastanza semplice: per moltiplicazione si moltiplicano o si dividono le mantisse in modo consueto si sommano o si sottraggono gli esponenti e si normalizza.
Per somma e sottrazione: si uguagliano gli esponenti, si sommano le mantisse. Si noti che la precisione del numero è data dal numero di cifre della mantissa e si utilizzano in genere due tipi di floating point: a singola precisione, e doppia precisione, cioè con doppia lunghezza della mantissa (range invariato, precisione raddoppiata). Lo standard di rappresentazione dei reali sui calcolatori si chiama IEEE 754 Floating point, ed ha mantissa di 23 bit (52 per doppia precisione), 1 bit di segno ed esponente a 8 bit (11 doppia precisione)

3.3 Altri tipi di informazione

In molte delle applicazioni che usiamo sul calcolatore, non lavoriamo su caratteri o numeri, ma elaboriamo informazione più complessa come immagini, suoni, misure fisiche, mappe geografiche e così via. Sono dati che in genere vengono generati da conversione di dati analogici, cioè misure fisiche di valori che possono assumere valori variabili con continuità nello spazio o nel tempo in numeri binari.
Ad esempio il suono acquisito dal microfono dovrebbe codificare la potenza dell'onda acustica che lo investe al variare del tempo, l'immagine acquisita da una telecamera dovrebbe rappresentare la misura del campo (bidimensionale) dell'intensità luminosa che raggiunge il sensore della camera stessa.
Per trasformare tali segnali (con le necessarie approssimazioni) in numeri binari che li rappresentino, sono necessarie due operazioni dette campionamento e binarizzazione.
Il campionamento consiste nell'acquisire il segnale soltanto per valori discreti di spazio e tempo.
La binarizzazione consiste nel rappresentare il numero reale che rappresenterebbe la misura fisica con un numero binario limitato in valore dalla quantità di bit utilizzati.
La Figura 18 mostra un esempio: se si suppone che il grafico a sinistra rappresenti un campo di grandezze fisiche che variano lungo l'asse orizzontale (ad es. intensità sonora in funzione del tempo), il campionamento acquisirà i valori dello stesso nei soli istanti t=t0+dt e la binarizzazione codificherà ciascun campione con una stringa binaria (rappresentante un approssimazione del valore della misura dell'intensità fisica). Il campionamento chiaramente introduce un errore dipendente dal "passo" di campionamento, cioè dal range di valori fisici codificati in un unico valore binario.

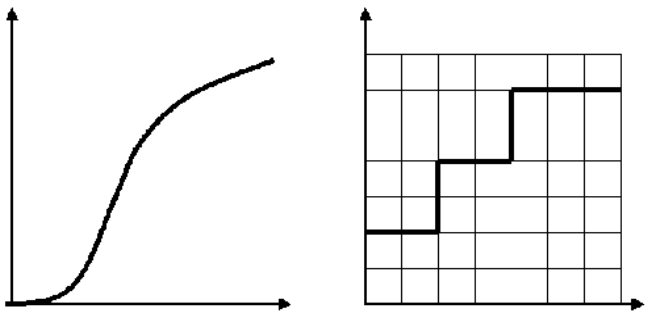

Figura 18: Esempi di segnale: analogico (sinistra) e digitale (destra)

Codifica audio

Il suono può essere trattato come un qualunque segnale analogico, cioè variabile con continuità nel tempo. I microfoni acquisiscono un segnale elettrico, cioè un valore di tensione variabile nel tempo e proporzionale all'intensità dell'onda sonora. Per memorizzarlo e renderlo elaborabile al PC, occorre campionarlo a tempi discreti e approssimarne il valore continuo con un valore numerico binario in un determinato sistema di codifica numerica in stringhe di bit (ad esempio 8, 16 o 24 bit). Si parla in questo caso di codifica PCM (pulse code modulation).

Il segnale digitalizzato sarà in generale differente dall'originale, la differenza sarà tanto minore quanto più elevata sarà la frequenza di campionamento e quanto più sarà elevato il numero di bit usati per la codifica dei valori di intensità.
La frequenza di campionamento ω_c è il parametro critico: se è sufficientemente alta, esiste un teorema (di Shannon) che garantisce che se la massima frequenza ω_{max} presente nel segnale è tale che $2\omega_{max} < \omega_c$ non vi è nessuna perdita di informazione dovuta al campionamento. Per questo il formato digitale utilizzato dai CD audio è campionato a 44.1 KHz, che corrisponde a poco più del doppio della massima frequenza percettibile dall'orecchio umano. Data l'elevatissima frequenza, l'audio PCM non compresso occupa moltissima memoria. Per cercare di ridurre il problema (che renderebbe ardua la trasmissione del suono sulle reti di calcolatori) un modo può semplicemente essere quello di rinunciare alla qualità, riducendo la frequenza di campionamento e tenendo basso il numero di bit (per la comprensibilità del parlato per esempio, bastano 8 bit e 8 Khz, ma la soluzione migliore è utilizzare tecniche avanzate di compressione che descriveremo brevemente in seguito).

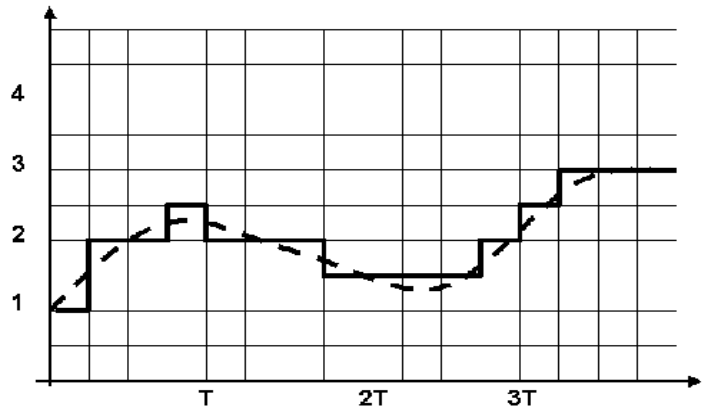

Figura 19: Binarizzazione del segnale: da un segnale continuo (linea tratteggiata) si passa ad uno discreto e campionato temporalmente.

Immagini

Le immagini che elaboriamo sui calcolatori vengono in genere suddivise in due tipi: raster e vettoriali. Le immagini raster sono matrici di *pixel* (picture element), vale a dire matrici di valori livelli di grigio (n bit) o componenti di colore 3*n bit. Sono immagini raster quelle acquisite da macchine fotografiche digitali, scanner, e così via e tutte le immagini sono

forzatamente convertite in raster al momento della rappresentazione sullo schermo, che mostra appunto una matrice di livelli di colore.
Nella grafica vettoriale le linee che compongono i disegni non sono realizzati come semplici pixel colorati, ma sono composti da una serie di oggetti. Esiste un numero limitato di oggetti o primitive standard definiti da equazioni: rette, poligoni, circonferenze, ecc. Tutte le altre forme sono generate da composizioni di queste primitive. La rappresentazione in memoria, dunque, non avviene punto per punto, ma oggetto per oggetto e ciascuno di questi è sintetizzato da una formula e da alcuni parametri o proprietà.
Un rettangolo, ad esempio, è rappresentato da parametri come altezza, larghezza, spessore della linea perimetrale, tipo di riempimento, colore e tipo delle linea perimetrale, colore e tipo del riempimento, trasparenza. Inoltre, per collocare l'oggetto rettangolo correttamente nell'ambito del disegno, è necessario stabilire le coordinate dell'angolo in alto a sinistra ed il livello di appartenenza. Quest'ultimo parametro permette la sovrapposizione di oggetti senza che questi, sul video, si fondano. La visualizzazione viene effettuata come se fosse composta da innumerevoli livelli immaginari sovrapposti, su ognuno dei quali è presente un solo oggetto. Gli oggetti che si trovano sui livelli superiori coprono quelli che si trovano sui livelli inferiori.
Se si realizza una linea, una curva od una figura geometrica con un programma di disegno ad oggetti, questa non sostituisce la parte di disegno già esistente, bensì vi si sovrappone. Inoltre, trovandosi su un livello a se stante, può essere sempre individuata, spostata, modificata, annullata o portata in un livello sottostante al disegno preesistente.
Il disegno ad oggetti trova applicazioni in campi molto disparati: disegno pittorico o illustrativo, disegno geometrico, disegno tecnico.
È chiaro che, mentre per un'immagine raster la rappresentazione sul display o sarà direttamente dipendente dalla matrice dei pixels, per quanto riguarda un grafico vettoriale a oggetti ciò che appare all'utente dipenderà da un apposito software utilizzato per la rappresentazione (rendering).
Tornando invece alle immagini aster la codifica delle immagini a colori può essere fatta con modalità differenti. La codifica *true color* prevede l'uso di 24 bit di cui 8 bit per il rosso (256 livelli), 8 per il verde, 8 per il blu. A volte si introducono altri otto bit per la correzione da effettuare per compensare le caratteristiche del monitor o per definire la trasparenza (canale alpha). Una codifica alternativa è quella con colormap: per ogni pixel sono usati solo 8 bit, i cui valori però non corrispondono a componenti di colore, ma sono indici che puntano ad una tabella (colormap) dove si trova la terna RGB da rappresentare (Figura 20). Questo permette nel caso il numero di colori dell'immagini sia ridotto di ottenere una rappresentazione più compatta.

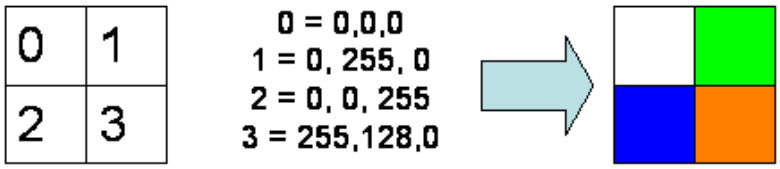

Figura 20: Utilizzo di colormap: i valori di ciascun pixel puntano ad una tabella dove si trovano le terne RGB corrispondenti ai colori.

Compressione dei dati

Dati come suoni ed immagini occupano, con le codifiche viste, enormi quantità di memoria. Questo ovviamente ne renderebbe difficile il trattamento e la trasmissione. Fortunatamente esistono tecniche di *compressione* che ne possono ridurre efficacemente l'ingombro senza evidenti danni per l'utilizzatore.

Queste tecniche sfruttano o la ridondanza dei dati o la possibilità di eliminare parti non utili di esse per ottenere fattori di compressione (rapporto tra la dimensione del dato compresso e quella dell'originale) spesso anche molto elevati.

Il suono campionato alla qualità dei CD audio, cioè 44100 Hz – stereo, quindi doppio canale e a 16 bit (65536 valori di intensità), occupa per ogni secondo 172 Kb. Una file contenente una canzone di 3 minuti occuperebbe quindi oltre 30 Mb. E' noto che i file di tipo mp3 possono contenere 3 minuti di qualità approssimabile al CD, occupando solo 2 Mb. Com'è possibile? La risposta risiede, in questo caso in una tecnica piuttosto complessa che decompone il segnale in bande di frequenza ed poi elimina dal segnale quelle frequenze che studi di psicofisica hanno dimostrato di essere poco rilevanti per l'orecchio umano.

In questo caso la compressione ha eliminato parte dell'informazione, si dice allora che essa è di tipo *lossy*, cioè tale che non è possibile risalire al dato originario a partire da quello compresso. Tecniche lossy ad hoc sono utilizzate allo stesso modo per la compressione di immagini o video, eliminando del tutto dal segnale le parti meno rilevanti ai fini della comprensione della scena.

Le tecniche che non degradano l'informazione si dicono invece *lossless* e devono essere applicate ogniqualvolta per motivi legali o di applicazione sia necessario preservare tutta l'informazione cercando però di ridurne le dimensioni (si pensi ai testi, alle immagini mediche)
Non tutte le tecniche di compressione sono intuitivamente complicate: facciamo brevemente cenno ad un paio di idee usate per comprimere i dati in casi semplici:

Una tecnica può consistere, ad esempio, nel codificare con meno bit i simboli (valori) più frequenti. Se in un testo la lettera è molto frequente mentre la q è rara potrei pensare di codificare la prima con due bit e la seconda anche con dodici, anziché tutte con otto. Allo stesso modo se in un immagine un certo colore è frequente lo posso codificare con pochi bit invece di 8 o 24. Questa idea è alla base della codifica di Huffman, algoritmo molto noto ed utilizzato.

Oppure si possono sfruttare le ripetizioni: posso pensare, ad esempio, di sostituire dei caratteri con all'indicazione del numero di successive ripetizioni dello stesso 0002222 = &30&42, dove il simbolo & è stato introdotto come separatore. Si può notare che in questo caso la dimensione viene ridotta, mentre nel caso non vi fossero ripetizioni la dimensione della stringa raddoppierebbe.
Questi metodi cenno sono evidentemente lossless e si possono applicare in modo piuttosto generico.

I metodi lossy, che eliminano informazione inutile come nel caso visto dei suoni, sono in genere specifici al tipo di dato. Nel caso delle immagini si utilizzano, ad esempio, metodi che tagliano alle alte frequenze, codificando l'immagine con le sue componenti di una

trasformata discreta (es. coseno) e eliminando le alte frequenze (Jpeg). Altro modo di comprimere le immagini può consistere nella riduzione del numero dei colori, mappando con misure di somiglianza i colori originali in una tabella (colormap) opportunamente ridotta. I più noti formati di immagini che usano sui PC (es. gif, jpg, bmp, png) implementano tutti metodi di compressione ottenuti combinando in modo vario quelli dei tipi sommariamente descritti.

4 Sistemi operativi e applicazioni software

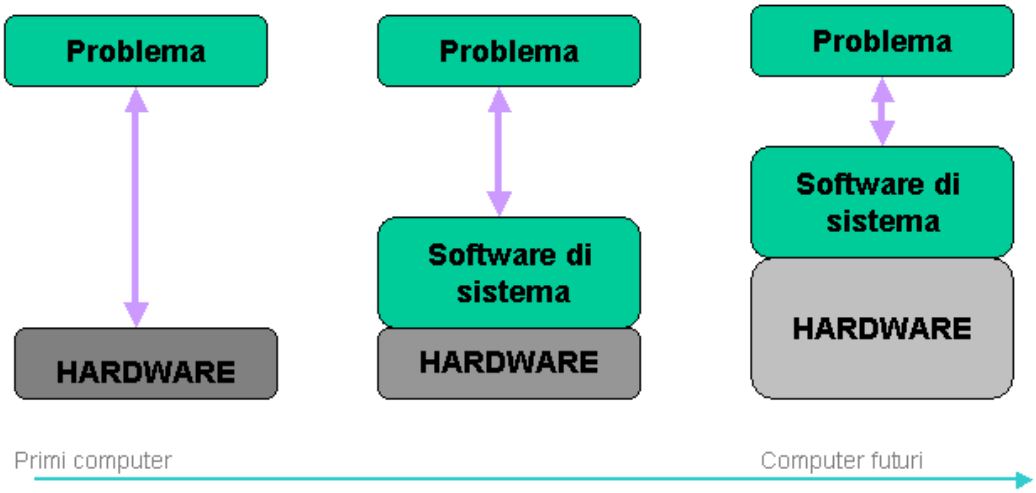

Figura 21: L'evoluzione del software di sistema, assente nei primi computer, dominante in quelli attuali.

Abbiamo già detto che istruzioni e dati memorizzati nella macchina di calcolo costituiscono il "software" del calcolatore. Nell'ottica di utilizzare un calcolatore per risolvere un problema, introduciamo nel calcolatore i dati, scriviamo l'algoritmo risolutivo e lo traduciamo nel linguaggio compreso dall'esecutore. Un calcolatore, di per sé, comprende soltanto serie di istruzioni per il microprocessore scritte in quello che viene detto "linguaggio macchina", cioè una serie di codici di operazioni (una sequenza di byte) e di operandi (dati scritti anche essi in binario). Abbiamo visto che in un calcolatore elettronico la CPU ha un set di istruzioni a ciascuna delle quali corrisponde un codice: questo vuol dire che un programma eseguibile lo è soltanto per quella determinata CPU, o per una CPU con le stesse operazioni e gli stessi codici. Un programma eseguibile, in linguaggio macchina, è eseguibile solo su un calcolatore con una determinata CPU e consiste quindi in una sequenza di codici operativi (numeri binari che corrispondono a una determinata operazione della CPU) e di operandi (dati codificati in codice binario. Poiché una simile sequenza non è molto comprensibile, la si indica con un linguaggio simbolico in cui stringhe di caratteri indicano le operazioni e le locazioni di memoria (es. ADD R1 R3 significa addizione dei contenuti di due registri).
Tuttavia i programmatori, al giorno d'oggi, non hanno bisogno di conoscere neppure questo linguaggio, in quanto esistono linguaggi evoluti attraverso i quali si possono scrivere gli algoritmi che poi vengono trasformati in linguaggio macchina dai compilatori, oppure controllati e interpretati linea dopo linea da programmi che vengono detti "interpreti". Questi linguaggi di programmazione vengono detti "di alto livello" (Pascal, Basic, C, C++, ecc) e ricordano più il linguaggio naturale umano che il linguaggio macchina.
Il compito dei programmatori di oggi è facilitato anche dall'avere a disposizione ulteriori strumenti, come ambienti di sviluppo, sistemi di "debugging" (per trovare facilmente gli errori), "editor" per scrivere il codice in maniera assistita, eccetera.
I linguaggi di alto livello hanno avuto un'evoluzione caratterizzata dal susseguirsi di diverse "filosofie" di programmazione.
Da quella iniziale basata sul semplice approccio sequenziale, si è passati all'approccio

"procedurale" in cui il programma è scomposto in una serie di moduli richiamati anche più volte dalla routine "principale", al moderno approccio ad oggetti, dove l'accento è messo sulla struttura e sull'organizzazione dei dati elaborate e sulle interazioni tra le strutture stesse. Anche in questo caso rimandiamo chi volesse approfondire l'argomento a leggere testi specifici, ma, soprattutto, a cimentarsi con la programmazione direttamente sul calcolatore, visto che nulla è più istruttivo della pratica e dell'apprendimento per tentativi ed errori.

4.1 Il software di sistema

Nei primi calcolatori, l'unico software era costituito dall'algoritmo per risolvere il singolo problema ed i dati di partenza. Con il tempo sono stati inseriti nei calcolatori dati e programmi "di base", atti a semplificare l'utilizzo della macchina. Sono stati inizialmente introdotti programmi per eseguire il caricamento dei dati e dei programmi, poi per la scrittura dei programmi. Poi sono stati introdotti sistemi per gestire al meglio le risorse evitando conflitti. A poco a poco si sono sviluppati una serie di programmi che separano sempre più l'utente dalla macchina fisica, che rendono possibile scrivere programmi in maniera facilitata, farli eseguire in maniera rapida e interattiva senza doversi preoccupare di dover scrivere le procedure per allocare la memoria necessaria, per accedere ai dispositivi di ingresso e uscita e così via.
Questi programmi costituiscono il software di base del calcolatore, detto anche sistema operativo. I sistemi operativi sono insiemi di programmi che permettono di utilizzare le risorse del calcolatore per scopi generali e senza doverne conoscere l'architettura ed il funzionamento a basso livello (cioè "vicino" all'hardware, utilizzando il linguaggio macchina, ad esempio). I programmi che lo compongono sono installati in genere dal fornitore, alcuni di essi vengono caricati in memoria all'accensione e si trovano normalmente in esecuzione, altri sono richiamati all'occorrenza.
Obiettivo primario di un sistema operativo è eseguire i programmi utente e agevolare la soluzione dei problemi dell'utente. Quando i sistemi operativi non esistevano, per risolvere i problemi l'utente doveva interagire direttamente con l'hardware. Con l'avvento del software di sistema, reso anche maggiormente necessario da un hardware sempre più complesso ed eterogeneo, la "distanza" tra utente e macchina si è drasticamente ridotta, i comandi che vengono dati e le interfacce con cui la si utilizza reagiscono con linguaggi che somigliano più a quello umano che alle istruzioni del microprocessore. Anche il tipo di problemi per la soluzione dei quali ci si affida al calcolatore sono cambiati e sono diventati vari e numerosi, dalla sola esecuzione di calcoli o algoritmi si è passati alla scrittura di testi, alla generazione di grafica, alla scrittura/esecuzione di musica, alla comunicazione remota e così via.

4.2 Stratificazione dei linguaggi

Abbiamo accennato a diversi linguaggi di programmazione dei calcolatori "vicini" al modo di operare dell'uomo oppure più vicini al modo di operare della macchina. Nei moderni calcolatori ne esistono diversi e si usa schematizzare la loro stratificazione con il concetto di "macchine virtuali" o di schemi a livelli. Cosa significa? Significa che quando si opera ad un determinato livello (ad esempio di programma utente, o di programmazione in C++, o di programmazione in linguaggio macchina, si hanno a disposizione delle risorse ed un linguaggio, cioè una serie di comandi ed una serie di operazioni che non corrispondono realmente alle operazioni elementari che compie la macchina stessa, ma che servono ad

utilizzare le risorse del livello.
I comandi “astratti” che si hanno ad un determinato livello definiscono una macchina “virtuale”, virtuale appunto perché non esiste fisicamente ed i suoi comandi sono implementati in termini di operazioni fatte dalla “macchina virtuale” di livello immediatamente inferiore. Ma chi agisce ad un determinato livello non necessita di sapere nulla di cosa accade al livello inferiore. Per rendere chiaro l’esempio, vediamo quali sono considerati i livelli degli attuali calcolatori:

- L6 Linguaggi altissimo livello/applicazioni
- L5 Linguaggi alto livello (C,C++,Fortran…)
- L4 Linguaggio assembler
- L3 Sistema operativo
- L2 Linguaggio Macchina
- L1 Microprogrammazione
- L0 Hardware, logica digitale

Questa divisione è solo una schematizzazione, comunque, piuttosto artificiosa e che va presa con beneficio di inventario. Quello che deve essere però chiaro sono il concetto di livello, macchina virtuale e di astrazione delle risorse, per cui chi utilizza un’applicazione o scrive un programma non deve conoscere il processore su cui girerà l’applicazione, non deve sapere come si ottiene la possibilità di scrivere in memoria, non deve sapere quali sono le istruzioni della CPU, ed ovviamente neppure come queste sono microprogrammate sui circuiti logici.
La reale attività della macchina avviene soltanto al livello più basso. Per trasformare un comando “virtuale” di un linguaggio di livello alto in reali azioni della macchina si passa per una serie di “traduzioni” che possono avvenire tutte prima dell’esecuzione (si parla allora di “compilazione” o direttamente alla chiamata del comando (si parla allora di “interpretazione”).
Altra cosa da notare è che ai livelli più alti, come si vede, corrisponde la semplicità d’uso e la vicinanza all’utilizzatore, ai livelli più bassi la vicinanza alla macchina e l’efficienza.
Dalla posizione nello schema a livelli, si capisce che il sistema operativo è qualcosa che “nasconde” i livelli più bassi legati alla macchina, a quelli più alti a cui accedono in genere gli utilizzatori più o meno specializzati.

4.3 Moduli del sistema operativo

Da quello che abbiamo visto, quindi, il sistema operativo costituisce lo strato di software che si trova tra le risorse della macchina e le applicazioni o i programmi dell’utente. Il suo compito è di rendere il più possibile semplice ed efficiente l’utilizzo delle risorse del calcolatore. Il sistema operativo sarà quindi composto da una serie di moduli che devono occuparsi dei seguenti compiti:

- L’esecuzione dei diversi *programmi* che, quando sono caricati in memoria e accedono al processore vengono chiamati *processi.*
- La gestione della memoria e di come questa sia attribuita ai programmi
- La gestione del salvataggio a lungo termine delle informazioni utili sul supporto non volatile (l’hard disk) attraverso la scrittura e lettura dei file.
- L’accesso alle risorse di eventuali utilizzatori diversi
- L’utilizzo delle **periferiche di input/output** con gestione di code e priorità.
- Fornire **interfacce utente** “user friendly” per l’interazione con l’utilizzatore.
- Rendere eventualmente possibile l’accesso da parte di più utenti, anche contemporaneo, mantenendo la sicurezza dei dati di ciascuno.

Nella pratica si utilizzano quasi sempre quelli che vengono definiti “programmi applicativi”:

compilatore di programmi, text editor etc. Il confine tra sistema operativo e programmi applicativi non è ben definito e nelle distribuzioni tipiche dei sistemi operativi più comunemente usati (Linux, Windows, MacOS…) si trovano centinaia di programmi applicativi vari già inclusi. Ma, in una visione stretta, il vero e proprio cuore del sistema operativo è costituito dal programma che comunque deve essere in esecuzione (kernel o nucleo) e che gestisce i processi e la memoria e nasconde ai processi utente l'hardware fisico, come nello schema di Figura 22.
Nel seguito descriveremo sommariamente i moduli fondamentali del sistema operativo:il gestore dei processi ed il gestore della memoria, il gestore dei file, il gestore dei dispositivi di input/output.

Figura 22: Il sistema operativo "nasconde" l'hardware agli utenti e crea l'ambiente di esecuzione per i programmi applicativi.

4.4 Gestione dei processi

Un processo è un programma in corso di esecuzione. Mentre un programma è una pura descrizione di un algoritmo, il processo è un'azione che sta evolvendo nel tempo, cioè un'entità attiva e dinamica.
Nei primi computer non c'era nessuna interazione dopo il caricamento di programmi e dati (in genere mediante scheda perforata): veniva lanciata l'esecuzione e se ne attendeva il termine. Questa modalità prende il nome di esecuzione batch. In questa modalità, il sistema operativo si occupava solo del caricamento dei dati, dell'allocazione della memoria e della scrittura dell'output e non c'era, quindi, bisogno di "gestione" dei processi: un solo processo poteva risiedere in memoria e questo non poteva essere interrotto dall'utente, ad esempio per cambiare dei dati. Inoltre tutte le volte che il programma richiamava una periferica, per esempio una stampante, la CPU rimaneva inattiva fino al termine dell'operazione, con perdita di grosse potenzialità di calcolo. La prima miglioria fu quindi realizzata mediante il cosiddetto buffering: la scrittura di dati in ingresso o in uscita su apposite aree di memoria quando la periferica non era pronta senza interrompere l'attività della CPU. In questo modo si realizzò anche il cosiddetto spooling: operazioni simultanee su più periferiche. I monitor di sistema vennero introdotti per verificare l'esecuzione del programma durante la sua stessa esecuzione.
Nei sistemi più recenti si passò all'organizzazione in memoria di più processi, alternati in memoria (multiprogrammazione). Quando la velocità dei calcolatori ha permesso di interrompere con frequenza elevata il programma in esecuzione si è ottenuta l'interattività.

In gergo tecnico, si usa classificare i sistemi operativi in base alla possibilità di gestire i processi in memoria come:
Mono tasking: capaci di gestire un solo processo: in tali sistemi non è possibile sospendere l'esecuzione di un programma per assegnare la CPU a un altro; erano ovviamente mono-tasking i primi sistemi operativi (es MS-DOS), lo sono ancora sistemi real time.

Multi tasking: capaci di eseguire più processi, mediante interruzioni e salvataggi in memoria
Time sharing: sistemi multitasking in cui l'alternanza dei processi non è causata solo da interruzioni esterne, ma viene programmata dividendo il tempo in quanti e facendo un cambio di processo (context switch) al termine di ogni quanto. Se il quanto è piccolo l'evoluzione dei processi appare parallela.
La gestione contemporanea di molti processi implica un grosso sforzo di progettazione del sistema operativo. Vediamo, quindi, come avviene l'evoluzione dei processi in un calcolatore moderno (time sharing), in cui, nonostante ci sia un solo microprocessore che opera sequenzialmente, abbiamo l'impressione che diverse attività avvengano in parallelo (possiamo, ad esempio, programmare ascoltando musica, ecc.) e con le quali possiamo interagire.
Innanzitutto abbiamo detto che si possono avere diversi processi in competizione per accedere alle risorse, in primo luogo alla CPU. Ciascun processo sarà caratterizzato dal suo stato, e dovrà poter essere interrotto per consentire lo svolgimento degli altri processi simulando l'esecuzione in parallelo e per consentire l'interazione con l'utente. Ad ogni processo sarà quindi associato un identificativo ed una serie di informazioni sul suo stato memorizzate in un apposita tabella (PCB Process Control Block). Essa dovrà contenere le informazioni utili al proseguimento della sua esecuzione, quindi ai valori di:

- stato di esecuzione
- contatore di programma
- registri della CPU
- informazioni sullo scheduling
- informazioni sulla gestione della memoria
- conteggio delle risorse
- stato dell'I/O

Quando si crea/distrugge un processo, si crea/distrugge il relativo controllore.
Ciascun processo avrà quindi il suo "ciclo di vita": verrà generato dall'utente, verrà caricato in memoria e posto in stato di "ready" (pronto ad essere eseguito) In questo stato il processo è potenzialmente in grado di avanzare. Dispone di tutte le risorse all'infuori del processore. Il nucleo mantiene una lista dei descrittori dei processi in questo stato detta "ready list".
Quando il job scheduler gli assegnerà la CPU, si dirà allora che il processo è in esecuzione (running). A seguito di una richiesta accolta di interruzione il processo potrà tornate nello stato di ready. Oppure potrà avere bisogno di dati o risorse al momento assenti od occupate, e passare quindi nello stato di waiting, in cui si pone in coda in attesa di un determinato "evento" che metta a disposizione i dati o la liberazione di un dispositivo di I/O.
Quando l'evento sarà segnalato o la coda del dispositivo libera, il processo tornerà nello stato "ready". Per la gestione di processi in multiprogrammazione si fa ricorso a delle code e a dei meccanismi di scheduling, cioè di programmazione dell'attività ed a meccanismi per gestire le interruzioni (interrupt). Si parla di scheduling a lungo termine per il caricamento in memoria dei processi e la loro messa in attesa delle risorse, di scheduling

a breve termine (job scheduling), per la gestione del tempo di CPU tra i processi. È` chiaro che per potere simulare il parallelismo tra processi ed avere l'interattività gli interventi del modulo di job scheduling (che vengono chiamati "cambiamento di contesto" o "context switch"), devono essere frequenti (ogni ~10 ms) e durare il meno possibile (il tempo ad essi dedicato ovviamente è tempo perso nell'esecuzione delle applicazioni).
Tra gli scopi dello scheduling c'è quello di massimizzare il tasso di servizio dei processi (*throughput)*, cioè il numero di processi serviti nell'unità di tempo.
Le interruzioni dell'attività del processo in corso non solo devono essere programmate, ma devono poter essere richieste, ad esempio, dai dispositivi di I/O, come del resto abbiamo visto nel capitolo 1.6 .
Il meccanismo di interruzione deve:

- Consentire il salvataggio del PC del Processo interrotto e trasferire il controllo ad una locazione fissa in memoria.
- Lanciare per ogni tipo di interruzione la relativa procedura di interruzione (Interrupt Routine o Interrupt Handler)
- Riconoscere la provenienza dell'interruzione
- Attribuire priorità e differenti strategie di risposta a ciascun diverso tipo di interruzione.

Il processo di interruzione si dice vettorializzato quando la richiesta contiene già una serie di informazioni quali la locazione della relativa routine di interrupt handling, la priorità, ecc.
Un problema che sorge è la necessità di non consentire l'interruzione di determinate attività (ad esempio delle routine di interrupt handling, altrimenti si perderebbero i dati sul processo interrotto). Si tratta, in generale di operazioni avviate dal sistema operativo e che, essendo più importanti di altre vengono dette istruzioni privilegiate. Altri esempi sono: smistamento del processore tra i processi, accesso ai registri di protezione della memoria, operazioni di ingresso e uscita, arresto del processore centrale. La differenziazione tra operazioni privilegiate e non può avere anche una seconda importante motivazione: avendo in memoria diversi
programmi, gli errori di uno possono influire sugli altri. Questo vuol dire che un buon sistema operativo deve impedire che i programmi scritti dall'utente ed eseguiti possano accedere "direttamente" alle risorse condivise (RAM, I/O devices.)
Tutto questo fa sì che, nei moderni calcolatori, si abbia un supporto hardware per differenziare almeno DUE modi di operazione.

- User mode –da parte di un utente.
- Monitor mode- gestita dal S.O.

L'avvio è in monitor mode, viene caricato il S.O. poi si avviano programmi utente. Il S.O. lavora sempre in monitor mode.
Le istruzioni privilegiate sono sempre eseguite in monitor mode
Abbiamo accennato al fatto che il sistema di scheduling della CPU (dispatcher) assegna quanti di tempo ai vari processi nello stato "ready" (prescindendo dalle interruzioni). Si parla in questo caso di "time sharing". L'assegnazione può avvenire con diverse strategie; ad esempio la "First come first served", con cui si attribuisce la CPU al primo processo che la chiede, la "Round robin" con attribuzione a turno, o quella basata su priorità. È grazie a questa alternanza che più processi possono essere contemporaneamente in memoria e, se l'alternanza è abbastanza veloce, dare l'impressione di essere eseguiti in parallelo, come avviene sui moderni PC. Inoltre la gestione veloce dei dispositivi di I/O congiunta con i processi multipli può efficacemente dare l'impressione di totale interattività con il sistema. Le interruzioni per le operazioni di I/O complicano ovviamente la gestione del sistema, creando la necessità di variare il programma di esecuzione dei processi da parte della CPU. Il dispatcher deve gestire le code di attesa per i vari dispositivi che possono

essere richiamati dai processi (ad esempio la stampante o il disco rigido) e pertanto deve essere accuratamente progettato per evitare situazioni di stallo (deadlocks), che si possono verificare se per esempio un processo è bloccato in attesa di una risorsa, ma occupa delle risorse di cui il processo servito dalla CPU ha bisogno, oppure situazioni di "starvation", in cui un processo non viene mai servito dalla CPU. Ovviamente il fatto che vi siano più processi contemporaneamente in memoria implica che possano dover comunicare cioè scambiarsi dei dati. Questo implica anche la gestione di aree di memoria condivisa e sistemi di sincronizzazione che impediscano l'accesso concorrente a tali aree. Ultima considerazione: il fatto che l'esecuzione di ciascun processo sia dipendente dal dispatcher, implica che nel sistema operativo non possa eseguire un processo in un tempo certo. Sistemi di calcolo in cui occorre avere vincoli temporali stretti (Hard real time), come quelli di robot industriali veicoli o velivoli non possono avere un sistema operativo time-sharing.

4.5 Gestione della memoria

Abbiamo visto nel capitolo come la memoria principale del calcolatore sia vista come un casellario di parole composte da un numero fisso di bytes e come esistano diversi tipi di dispositivi di memorizzazione.
Vogliamo ora vedere come il sistema operativo attribuisca la memoria ai vari processi in esecuzione. Innanzitutto il sistema operativo si occupa di:

- Assegnare a ciascun processo tutta la memoria di cui ha bisogno
- Rendere l'accesso a questa memoria il più possibile veloce
- Evitare che i diversi processi in memoria accedano alle stesse aree di memoria in modo non coordinato e che le scritture/letture di un processo danneggino gli altri processi.

Ovviamente il sistema operativo può utilizzare tutti i dispositivi di memoria che abbiamo visto nel capitolo 1.3, cioè una grande RAM, una piccola e veloce cache ed il disco rigido. Consideriamo la memoria principale come un casellario e dimentichiamoci della gestione della cache di cui abbiamo già un po' parlato. A parte un settore (supponiamo le prime N caselle) dove viene caricato il sistema operativo, nella restante parte di memoria si possono caricare i processi. Per poter gestire processi di dimensioni variabili, è preferibile non fissare a priori le aree di memoria riservate a un processo, si ha quindi una allocazione *dinamica* (Figura 25) Nella forma più semplice si possono avere associate a un processo un indirizzo base e un indirizzo relativo. Inoltre è necessario un limite per impedire che il processo scriva in regioni dove è già presente un altro processo. Se il processo crescesse oltre il limite, sarebbe necessario spostarlo in un'altra regione di memoria. Con processi multipli in memoria avremmo una situazione problematica e poco flessibile: anche con pochi processi in memoria, dopo un po' di tempo c'è il rischio di non trovare spazio contiguo sufficiente per l'allocazione di un nuovo processo. Supponiamo infatti che un processo finisca e lasci libero il suo spazio, mentre i processi che lo precedono e seguono negli indirizzi in memoria continuino a esistere (Figura 24). Un nuovo processo può occupare lo spazio che resta libero solo se ha dimensioni minori. Inoltre se ha dimensioni inferiori lascia comunque libero uno spazio non utilizzabile. Si parla, in tal caso di *frammentazione* della memoria.
Per risolvere questo problema occorre svincolare l'indirizzo logico della memoria indirizzata da un programma da quello fisico. La memoria associativa (con una tabella di conversione che a ciascun indirizzo logico faccia corrispondere un indirizzo fisico completamente diverso) ha un peso troppo grande (metà della memoria servirebbe per la tabella); un ragionevole compromesso lo si ha con il meccanismo della *paginazione*.

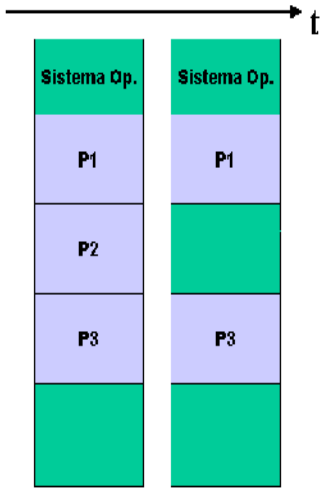

Figura 24: Allocazione dinamica nella memoria di processi multipli. Il processo 2 termina e lascia uno spazio libero, che difficilmente può essere sfruttato. Si parla di frammentazione della memoria.

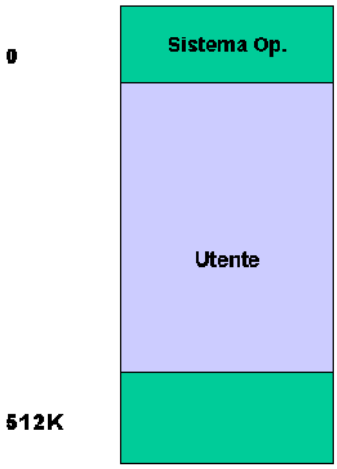

Figura 23: Allocazione statica: all'utente viene data una regione fissa di memoria dopo quella affidata al sistema operativo

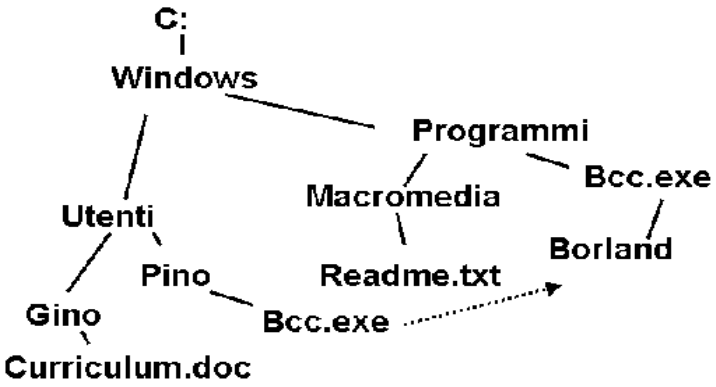

Figura 26: Struttura ad albero del file system di Microsoft Windows

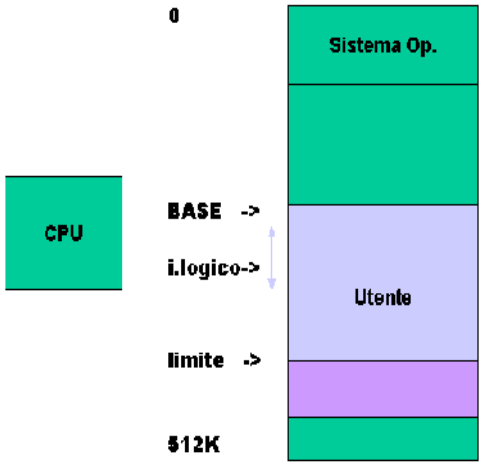

Figura 25: Allocazione dinamica: la memoria utilizzata dal programma utente è indirizzata da un indirizzo "logico"; quello fisico viene calcolato poi dalla CPU sommando tale indirizzo all'indirizzo base stabilito dalla CPU. Il limite superiore impedisce che la regione allocata vada a occupare regioni attribuite ad altri processi.

La paginazione consiste nel suddividere la memoria in settori di dimensione fissa detti, appunto, pagine. La memoria logica di ciascun processo viene fatta corrispondere alla memoria fisica pagina per pagina, cioè l'indirizzo viene suddiviso in due parti, la prima parte indica il numero di pagina, il secondo l'indirizzo all'interno della pagina. Il numero di pagina viene tradotto in una locazione fisica dal meccanismo di gestione, che è un sistema hardware detto Memory Management Unit. In questo modo si risolve il problema della

frammentazione (al massimo gli spazi non allocabili diventano di dimensione inferiore alla pagina). Esempio: supponiamo di avere indirizzi di 8 bit (256 caselle)
Ma la paginazione permette di fare di più, cioè di ottenere, utilizzando il disco rigido, uno spazio di memoria potenzialmente molto maggiore della memoria fisica. Infatti non c'è nessun bisogno di avere un numero di pagine uguale al numero di pagine contenute in memoria. Si possono avere molte più pagine tenendo quelle meno usate salvate in un'apposita area del disco rigido. Si parla in questo caso di memoria virtuale. Ovviamente questo fa sì che in alcuni casi il Memory Management Unit non trovi un indirizzo di memoria fisico corrispondente all'indirizzo virtuale. Si parla in tal caso di "*page fault*". Per leggere i dati richiamati bisogna caricare la pagina dal disco rigido. L'operazione di trasferimento di pagine dal disco rigido alla memoria o viceversa prende il nome di "swapping". Ovviamente l'accesso al disco rigido è estremamente lento (~10 ms contro i ~20 ns della memoria!), per cui in caso di page fault le prestazioni del sistema sono drammaticamente compromesse, ma almeno i processi possono continuare senza essere interrotti. Ovviamente non è tutto così semplice, ovviamente occorre gestire la collocazione delle tabelle, la programmazione del caricamento in memoria dei processi atto a minimizzare i page fault e ottimizzare l'esecuzione (long term scheduling o job scheduling), tuttavia i meccanismi fondamentali di gestione sono questi. Si può ancora notare che i sistemi operativi possono distinguere settori di memoria relativi al codice da quelle dei dati allo stack dei dati della cpu salvati dai processi e garantire anche l'accesso concorrente di più processi ad alcune aree di memoria (es. il settore del codice).

4.6 Gestione dei file

Abbiamo già visto che i file sono le unità di informazione organizzata salvate sulla memoria di massa (l'hard disk). Quando utilizziamo un sistema operativo, siamo abituati ad avere a che fare con oggetti dotati di nome e svariate proprietà (eseguibilità, leggibilità, identità del creatore, data di creazione, eccetera). Li richiamiamo con interfacce testuali, oppure li vediamo in interfacce desktop come icone, eccetera. Chi gestisce la corrispondenza tra queste unità logiche e i reali bit scritti sui settori del disco è il sistema operativo. Il modulo che si occupa di questo prende il nome di file system. Esso è lo strumento che permette all'utilizzatore della macchina

- la registrazione di dati su supporti di memorizzazione esterni al modello di Von Neumann
- la ricerca di dati su supporti esterni

L'entità base di dato memorizzato, come abbiamo detto prende il nome di file ed è in sostanza una raccolta di dati registrati su memoria di massa caratterizzata da un tipo, un nome, una dimensione, una data di creazione e, ovviamente una serie di dati memorizzati sulle tracce del disco rigido o su un altro dispositivo di memorizzazione permanente. Come avvenga questa mappatura è un problema di progetto del sistema operativo, non ci addentreremo qui a descrivere i metodi, ma citiamo quelli usati da Unix (Inodes, elementi memorizzati con possibili collegamenti a catena, la File Allocation Table (FAT) di DOS e Windows, l'NTFS di Windows NT. Il file system si occupa in genere anche di rendere indistinguibili i dati memorizzati su supporti diversi o anche in realtà residenti in macchine remote.
Come file troviamo memorizzati documenti testuali, programmi eseguibili in linguaggio macchina, immagini, eccetera. A parte informazioni di tipi diversi, i moderni sistemi operativi distinguono anche casi molto particolari di "file" come la *directory (cartella)*, che permette di strutturare l'archivio dei files ad albero, ed il link (collegamento), che permette di accedere agli stessi dati da differenti posizioni nell'albero senza dover duplicare l'informazione stessa. La directory è semplicemente un particolare tipo di file in cui sono

memorizzati tutti i nomi di una collezione di file e informazioni ausiliarie quali data e ora di creazione/ultima modifica, proprietario e diritti di accesso, dimensioni in byte.
La directory, quindi, anziché puntare ad informazione, punta una lista di altri files, cosicché a partire dall'origine del disco rigido, o meglio di una sua partizione (indicato in Microsoft Windows da una lettera, es. "C"), si trovano i vari files seguendo una serie di collegamenti che definiscono una struttura ad albero come quella disegnata in Figura 26.
I sistemi operativi distinguono in generale differenti tipi di file, assegnando per esempio degli attributi per distinguere se si tratta di file binari o di testo (codificati in ASCII), se sono eseguibili (i programmi) o non eseguibili (i file di dati). Alcuni sistemi operativi (es. Windows) associano tipi diversi a file prodotti con programmi diversi (es. Word, Excel, ...), ed in genere lo fanno attraverso le ultime lettere del nome (estensioni). Grazie a tali distinzioni è possibile per esempio fare in modo che nelle interfacce grafiche si attribuisca una diversa icona a ciascun tipo di file e sia possibile lanciare l'applicazione con cui è stato salvato il file semplicemente premendo il tasto del mouse quando il puntatore è sovrapposto sull'icona del file stesso. Altri attributi che vengono assegnati dal sistema operativo ai file riguardano data e ora di creazione e le informazioni sull'utente che ha creato il file e i diritti di accesso.
Questi ultimi indicano quali operazioni possano essere esercitate da ciascun utente sul file. In calcolatori che possono essere utilizzati da più utenti, infatti, i sistemi operativi possono fare in modo che, per esempio, utenti generici possano leggere i file prodotti da un altro utente, ma non modificarli, oppure abilitare o meno un gruppo di utenti all'esecuzione di programmi.

Nei sistemi operativi si distinguono spesso, relativamente a un file, le seguenti categorie di utenti:

- il proprietario
- il gruppo di lavoro (del proprietario)
- gli altri utenti

e diverse tipologie di permessi:

- di esecuzione
- di lettura
- di scrittura

Nei sistemi Unix e derivati, a ciascun file sono assegnati un proprietario ed un gruppo, e ciascun file può essere eseguibile o meno da parte del proprietario, del suo gruppo e da tutti gli utenti.

4.7 Gestione delle periferiche

Un ulteriore compito del sistema operativo è quello di mettere a disposizione dei processi le periferiche di input/output e di rendere l'attività dell'utente indipendente dal tipo e dalla marca dei vari componenti esterni. Deve quindi esistere un modulo apposito del sistema operativo adibito alla gestione dei sistemi di input e output e da questo modulo dipende la maniera in cui i periferici vengono riconosciuti vengono sfruttati.
Esistono vari modi con cui i sistemi mettono a disposizione le funzione delle periferiche agli utenti (ad esempio in Unix esistono tipi di file che corrispondono logicamente ai vari dispositivi e su cui si può leggere e scrivere). Per poterlo fare, tuttavia, devono esistere, per ciascun dispositivo, programmi appositi di controllo che facciano corrispondere le interfacce standard del sistema operativo con le effettive azioni del dispositivo stesso.

Questi programmi si chiamano *device drivers* e devono essere forniti dal produttore della periferica stessa. Essi rendono la periferica "virtuale" al processo (cioè il processo comunica con esso attraverso comandi standard indipendenti da esse e sono i driver a convertire gli ordini nelle reali azioni sui dispositivi). I driver quindi gestiscono anche le comunicazioni, i segnali in transito verso i dispositivi e devono gestire le eventuali conflittualità che nascono quando più processi vogliono comunicare con una periferica.

4.8 Interfacce utente

L'interfaccia utente è ciò che ci permette di interagire con la macchina. Mentre una volta l'interfaccia di un calcolatore poteva essere solo un lettore di schede perforate, nei moderni personal computer si interagisce principalmente mediante lo schermo, una tastiera ed il mouse. Nel calcolatore esistono programmi che vengono eseguiti e che ricevono gli input da tastiera e mouse, e visualizzano in corrispondenza delle varie azioni degli effetti. I primi personal computer, esattamente come le workstation professionali, avevano un interfaccia testuale. In tali interfacce l'utente può scrivere del testo con la tastiera vedendolo comparire direttamente sullo schermo. Un programma funge da **interprete comandi** e richiama le funzioni del sistema operativo (chiamate di sistema) corrispondenti alle parole chiave digitate. Allo stesso modo si possono far eseguire dal sistema anche i programmi utente richiamandone il nome dall'interfaccia, che prende usualmente il nome di "shell". L'uso di questo tipo di interfaccia è stato reso obsoleto dall'introduzione di interfacce grafiche e di software specifico per le varie applicazioni. Le interfacce a finestre, come quelle che si usano in Windows, MacOS e Xwindows (Unix), risalgono al 1980 e furono introdotte dalla Apple con il modello Macintosh. Queste interfacce rappresentano file e cartelle con simboli grafici su uno sfondo che rappresenta una virtuale "scrivania" (desktop in inglese). Si parla propriamente di "paradigma della scrivania". Grazie all'uso del mouse come strumento di puntamento, l'utente può spostare file e programmi da una cartella all'altra come se fossero oggetti sul tavolo e sempre usando i tasti del mouse stesso può fare eseguire i vari programmi (in generale con il tasto sinistro, mentre col destro si esaminano le proprietà del file e le possibili applicazioni. Nei primi PC

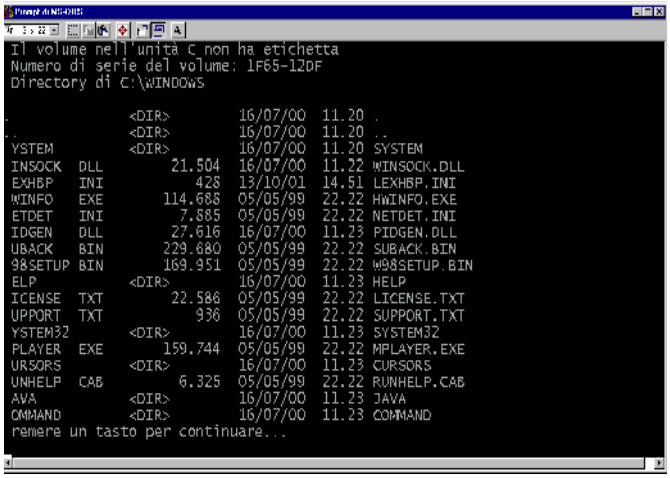

Figura 27: Shell di MS-DOS

Figura 28: Interfaccia a "desktop" di Microsoft Windows 2000

con interfaccia a finestre, comunque, il mouse aveva un solo tasto). I simboli grafici che caratterizzano cartelle e files sulla “scrivania” prendono il nome di “icone” e Il successo del sistema ha fatto sì che venisse riprodotto su tutti i sistemi operativi presenti in commercio.
La modalità di utilizzo dei programmi secondo quello che viene chiamato il paradigma di interazione WIMP (Windows, Icons, Menus, Pointer) è da anni il più utilizzato a livello di sistemi operativi ed applicazioni utente.
Altri tipi di interfacce (vocali, con touchscreen, ecc.) sembrano per il momento riservate a calcolatori utilizzati in particolari applicazioni (esempio integrati in stazioni di pubblico utilizzo, elettrodomestici, industriali, ecc.).

4.9 Classificazione dei sistemi operativi e sistemi operativi in commercio

A seconda delle caratteristiche di cui abbiamo parlato, si usa suddividere i sistemi operativi disponibili sui calcolatori in diverse categorie.
In base alla possibilità di interagire con la macchina possiamo avere sistemi:

- Batch (esecuzione di un processo non interattiva)
- Interattivi (ormai tutti i sistemi di uso comune): i processi possono essere interrotti o modificati dall'utente

In base alla gestione dei processi:

- Monoprogrammati (un solo programma caricato in memoria)
- Multiprogrammati (più di un processo contemporaneamente in memoria)
- Time sharing (processi che simulano il parallelismo alternando l'esecuzione dei processi)

Un’altra distinzione si fa per i sistemi **real time**, che devono garantire l’esecuzione dei processi in un tempo predeterminato e che quindi non possono agire in time sharing come quelli che abbiamo descritto.
A seconda della possibilità di distinguere tra più utenti, mediante aree disco separate, autenticazione con username e password, diritti di accesso dei file, ecc. i sistemi operativi si dicono

- Monoutente
- Multiutente

I primi personal computer di successo furono quelli derivati dai primi PC IBM con sistemi operativi prodotti dalla compagnia Microsoft. Questi si sono imposti per ragioni commerciali e di larga diffusione, non tanto per l'ottimalità del sistema. Il primo sistema operativo, MS DOS (Microsoft Disk Operating System), aveva infatti le seguenti, non ideali caratteristiche:

- monotask
- monoutente
- file-system gerarchico
- memoria limitata (1 MB / 640 KB)
- nessuna protezione

L’evoluzione del sistema, che cercava di imitare le caratteristiche di successo della piattaforma Macintosh, portò al sistema operativo Windows: esso aveva inizialmente le seguenti caratteristiche:

- multitask imperfetto

- monoutente
- stesso file system del MS-DOS
- interfaccia grafica a finestre e menù
- sistema a 16 bit
-

Microsoft nel frattempo sviluppò anche un sistema operativo professionale con caratteristiche migliori, Windows NT. Esso aveva le seguenti caratteristiche:

- multitask
- monoutente
- NTFS (NT File System)
- microkernel, thread
- sistema a 32 bit

Le workstation professionali e scientifiche, invece, già dagli anni '70 avevano sistemi operativi per molti versi più robusti, appartenenti alla "famiglia" Unix. Il sistema nacque presso gli AT&T Bell Labs e si mantenne all'avanguardia in quanto sviluppato nelle università, almeno per quanto riguarda il supporto efficiente alla programmazione ed all'esecuzione di programmi di calcolo. Esso introdusse, tra l'altro, caratteristiche come:

- multitasking
- multiutenza
- protezione accessi, diritti su file, sicurezza
- ottima integrazione in rete
- portabilità dei programmi

Le più recenti evoluzioni dei sistemi operativi Microsoft e Apple hanno fatto acquisire agli stessi molte caratteristiche di Unix, mentre un sistema di tipo Unix adattato all'architettura Intel da Linus Torvalds e detto per questo Linux, si è diffuso come alternativa ai sistemi Windows sui comuni PC, adottando anche interfacce grafiche user friendly simili a quelle di Windows che lo rendono estremamente competitivo anche per le applicazioni di interesse generale, anche per il fatto che si tratta di software libero che non richiede licenze d'uso.

4.10 BIOS

Il sistema operativo fin qui descritto e che risiede nel disco rigido del PC non può essere l'unico software presente nel calcolatore, anche perché non risiede in memoria alla sua accensione. All'accensione del computer si avviano delle routine esterne che devono controllare lo stato dell'hardware e delle periferiche e caricare i programmi del sistema operativo dal disco rigido alla memoria centrale. Questi programmi di avvio non possono risiedere nel disco rigido. Il cosiddetto ***BIOS*** (*Basic Input Output System*), consiste di una serie di routine che esaminano l'hardware, avviano il sistema operativo, sovrintendono al controllo delle periferiche di lettura/scrittura e alle operazioni di lettura/scrittura della RAM. Esso risiede in una memoria ROM montata sulla scheda madre e può esser caricato istantaneamente all'accensione.
Il BIOS e le componenti software similari, non sono in genere assimilati al sistema operativo, anche se sono effettivamente parte del software di sistema; essi prendono spesso il nome di **firmware**, ad indicare che è fornito dall'assemblatore del PC, a differenza del sistema operativo.

4.11 Programmi applicativi

L'uso del computer è ormai diffuso a tutta la popolazione e le applicazioni che se ne fanno

sono le più svariate. La maggior parte degli utilizzatori non si pone più problemi da risolvere mediante la programmazione, ma si limita ad utilizzare programmi già scritti per svariate applicazioni.
Esistono programmi per l'utilizzo domestico, aziendale e scientifico. Essi sono realizzati dai programmatori sfruttando le caratteristiche dell'hardware e del sistema operativo, per cui, in generale, sono specifici a una determinata architettura e a un determinato sistema operativo. Gli applicativi da ufficio più usati sono Word Processor, Fogli elettronici, database, programmi per generare presentazioni e disegni, ma anche client di servizi di rete come posta elettronica, Web, ecc.
Un Word processor è un programma che permette di scrivere, correggere, archiviare, stampare documenti. Il loro utilizzo costituisce un grande vantaggio rispetto all'usuale modo di scrivere, consentendo di fare correzioni modificare files in tempi successivi, trasmettere documenti senza stamparli, tagliare parti di testo ed incollarle in altra posizione, cambiare caratteri spaziature, colori, impaginare, eccetera. I word processor più diffusi, sono del tipo WYSIWYG (What You See Is What You Get) basati su un interfaccia che rappresenta sullo schermo, l'immagine della pagina come verrà stampata. Inoltre integrano le caratteristiche tipiche che una volta erano riservate a programmi di "desktop publishing" (programmi professionali per l'impaginazione), come incolonnamenti, caselle di testo grafici, figure, eccetera. Prodotti di questo tipo sono, ad esempio Microsoft Word, OpenOffice.org Writer.
Un foglio elettronico (spreadsheet), è un programma che consente di elaborare dati ed organizzarli. Esso si basa su tabelle, in ciascuna delle quali è possibile inserire dati in vari formati o funzioni che li elaborano. Le funzioni integrate nei programmi sono sia matematiche che statistiche ed economiche rendendo il loro utilizzo molto versatile. Inoltre essi danno la possibilità di scrivere le proprie funzioni e di inserire grafici basati sui dati nelle tabelle. Esempi tipici sono Lotus 1-2-3, Microsoft Excel e Openoffice.org Calc.
Programmi molto utili per la didattica e la comunicazione sono quelli che consente di generare presentazioni scientifiche e commerciali (slideshow). Essi permettono di creare diapositive con vari stili e sfondi, generare animazioni, definire tempi di transizione tra slide, inserire commenti sonori e altro. Esempi di questo tipo di programmi sono ad esempio Microsoft PowerPoint e OpenOffice.org Impress.
Le basi di dati o database sono archivi di dati organizzati anche se il termine database viene spesso utilizzato in modo impreciso per indicare i programmi che li gestiscono (DBMS, database management system) e che garantiscono per i dati sicurezza e rapidità di accesso, supportano varie modalità di ricerca, ed evitano duplicazioni e ridondanze.
La tecnologia dei database si è evoluta nel tempo, dall'organizzazione gerarchica, si è passati al modello relazionale dei sistemi più diffusi al giorno d'oggi, dove i dati sono organizzati in tabelle, che rappresentano le relazioni tra le varie entità definite dai dati.
L'uso delle basi di dati è molto diffuso per le applicazioni di rete, che servono molti utenti. I server possono essere interrogati ed aggiornati dai programmi clienti utilizzando apposite interfacce e linguaggi di "query" come SQL.
Sistemi di gestione di sever di basei di dati molto noti sono per esempio Oracle o Microsoft SQL Server, ma esistono anche ottimi prodotti open source (vedi capitolo seguente) come MySQL.
Un tipo di sistema di gestione di dati che può essere molto utile per diversi professionisti e ricercatori è costituito dai sistemi informativi geografici (GIS) che sono sostanzialmente dei database in cui è possibile georeferenziare l'informazione (cioè attribuire a ciascun campo dei valori collegati di longitudine e latitudine od altre relazioni spaziali) e di poter effettuare così ricerche territoriali.

Tornando alle semplici applicazioni utente, l''ampio uso di apparecchi digitali per generare immagini e video (fotocamere digitali, webcam, camcorder) ha reso molto popolari le applicazioni di fotoritocco ed anche quelle di montaggio audio-video. Altri programmi applicativi ormai utilizzati da gran parte degli utenti dei calcolatori sono poi ovviamente quelli utilizzati per l'utilizzo di servizi Internet (che vedremo meglio in seguito), come la posta elettronica e i browser web, ma anche lo streaming audio e video.
Per non parlare poi dei vari tipi di simulazioni e giochi (l'intrattenimento tramite i giochi per calcolatore ha ormai superato il cinema come giro d'affari e va considerato che anche le moderne console per videogiochi sono, in pratica, potenti calcolatori non dissimili da quelli descritti nei capitoli precedenti.

L'utilizzo di questi programmi è lo scopo principale della maggior parte degli utenti dei moderni sistemi informatici. Essi sono in generale dotati di interfacce utente ben progettate e di modalità di utilizzo specifiche all'applicazione che possono essere apprese in modo abbastanza semplice da chi ne ha bisogno per la propria attività e senza avere nozioni di come il programma od il computer funzionino.
Per questo molto spesso l'atteggiamento dell'utilizzatore di PC "ingenuo" consiste semplicemente nel pensare che gli utenti debbano solo farsi vendere a scatola chiusa una macchina che avvii il programma che gli occorre.
Quali vantaggi abbiamo ottenuto dal capire come questi programmi utilizzino le macchine attraverso i sistemi operativi, come possano codificare i dati salvati, come utilizzino la memoria ed il processore?
Molto più di quanto potrebbe sembrare. Chi usa un PC per far girare una determinata applicazione, deve sapere ad esempio quali versioni dell'applicazione sono disponibile per una data architettura ed un dato sistema operativo ed a che costi. Ha sicuramente vantaggi nel sapere che i programmi eseguibili funzionano solo su un determinato processore e con un determinato sistema operativo, che necessitano di una determinata quantità di memoria o di una scheda grafica con determinate caratteristiche. Può scegliere quindi, ad esempio, il computer adatto alle sue esigenze senza spendere più del necessario per caratteristiche o programmi che non userebbe.
Può decidere quale sistema operativo è più adatto per i programmi di cui fa utilizzo e per l'architettura scelta (ad esempio evitando di installarne uno che richieda troppe risorse se la macchina non le possiede.
Le sue competenze possono fargli capire a cosa è eventualmente dovuto un malfunzionamento o un eventuale crash del programma (ad esempio per problemi nella gestione della memoria associata processi da parte del sistema operativo). Può risolvere in maniera semplice problematiche relative a connessioni di periferiche, driver, utilià di sistema. Se un utente sa come funzionano i filesystem e come sono codificati i files non deve chiamare un tecnico solo perché magari vogliono salvare un file di testo in modo da renderlo leggibile per un altro programma. E così via.

Non c'è bisogno, quindi, di essere programmatori per trarre giovamento sia dal punto economico che della produttività nel proprio lavoro dalla conoscenza delle informazioni basilari che qui cerchiamo di riassumere.
Un'altra cosa che può sicuramente giovare a tutti gli utenti dei calcolatori è una minima conoscenza di come vengano prodotti i programmi applicativi, con quali licenze d'uso vengano distribuiti e così via, ed è il caso, quindi, di spendere due parole a proposito.

Come abbiamo visto i computer non sarebbero utilizzabili dall'utente senza software. Le macchine vengono in genere vendute con un sistema operativo prodotto da una azienda

che in genere non è quella che fornisce l'hardware. Come tutti sanno la grande maggioranza dei computer utilizzati dal grande pubblico ha microprocessori Intel/AMD con associati sistemi operativi prodotti da Microsoft, che detiene una posizione dominante sul mercato. Le alternative principali sono costituite dai sistemi operativi Apple, che però vengono venduti soltanto associati a macchine dello stesso produttore e dalle varie versioni (distribuzioni del sistema operativo “libero”) GNU/Linux.
Per quanto riguarda i programmi applicativi esistono molte più opzioni, anche se spesso gli utenti non ne sono sufficientemente consapevoli.
Sistemi operativi ed applicativi vengono venduti/distribuiti con una licenza d'uso, che è in pratica il contratto stipulato tra chi detiene I diritti intellettuali sullo stesso (copyright) e l'utente. Il software comunemente commercializzato dalle aziende con vincoli stringenti e costi associati alle licenze viene in genere detto software proprietario.
Esistono vari tipi di licenze commerciali che vincolano e limitano l'uso dei programmi sulla base della tipologia di acquisto o di politiche commerciali, anche se, ovviamente è difficile il controllo sulla diffusione dei programmi.
Copiare e utilizzare i programmi ove non consentito dalla licenza è ovviamente illegale e le case produttrici cercano di utilizzare sistemi per evitare l'uso non autorizzato dei programmi (codici di blocco, server di licenze, ecc), anche se esistono sempre gruppi di “cracker” che cercano di trovare i codici per sbloccare illegalmente le licenze (si parla spesso di software “craccato” o “piratato” (Notare che in termini da informatici vi è differenza tra cracker e il termine similare “hacker”, che non ha necessariamente una connotazione di criminalità).

Non tutto il software utile, però è necessariamente costoso e neppure necessariamente a pagamento. Esistono programmi che alcune aziende distribuiscono anche gratuitamente con vari tipi di licenze es shareware (versioni di prova), abandonware (vecchie versioni) ecc.
Ma soprattutto esiste un grandissimo movimento teso allo sviluppo e la diffusione di software non proprietario, cioè libero, creato con criteri diversi da quelli del profitto aziendale, ma non per questo meno efficiente o utile.
Il software libero nasce in ambito scientifico educativo ove è normale che una grande comunità collabori a progetti comuni. Così nono nati i sistemi operativi di tipo Unix, compreso Linux che è ormai una valida alternativa ai sistemi Microsoft, ma allo stesso modo grandi comunità lavorano allo sviluppo di programmi per i più disparati utilizzi.
Per quasi tutti i tipi di applicazioni utente citate sopra, esistono ormai ottime alternative di software “libero” ai costosi programmi commerciali, disponibili per tutte le piattaforme e sistemi operativi, proprietari e non.
Visto che anche sul fenomeno del software libero o delle licenze d'uso l'informazione non è sempre ottimale, e visti i grossi vantaggi economici e di efficienza che possono derivarne per l'utente, facciamo una breve panoramica su di esso per chiarire brevemente le idee e fugare le diffidenze spesso ingiustificate che spesso sono ancora diffuse.

4.12 Il software “libero”

Il cosiddetto software “libero” è software che pur essendo dotato di una licenza d'uso, si distingue dal normale software proprietario in quanto lascia in genere piena libertà di modifica e di uso dei programmi garantendo soltanto la tutela del copyright degli autori.
La definizione di software libero data dal suo pioniere Richard Stallmann e dalla sua Free Software Foundation dice che il software, per poter essere definito libero deve garantire le cosiddette quattro "libertà fondamentali":

- Libertà di eseguire il programma per qualsiasi scopo ("libertà 0")
- Libertà di studiare il programma e modificarlo ("libertà 1")
- Libertà di copiare il programma in modo da aiutare il prossimo ("libertà 2")
- Libertà di migliorare il programma e di distribuirne pubblicamente i miglioramenti, in modo tale che tutta la comunità ne tragga beneficio ("libertà 3")

Dato che si considera anche la disponibilità del codice sorgente nella definizione di software libero, spesso si usa anche la definizione software “open source”, che, però, non è del tutto coincidente: infatti con Open Source ci si riferisce a requisiti dichiarati dall Open Source Initiative che non è garantito rimangano costanti.

Si notino alcune cose: la comunità di sviluppo del software libero è sempre più ampia e non necessariamente composta da soli volontari, ma per alcuni progetti gode di forti sostegno economico. Il fatto di avere un ampia base di sviluppo e la disponibilità dei codici rende, specialmente per I sistemi operativi, i programmi più robusti, impedisce l'inserimento di programmi di controllo degli utenti da parte delle case, e presenta molti altri vantaggi. Per questo il software open source sta riscuotendo molto successo, non solo per la sua gratuità per l'utente, ma anche per la sua alta qualità.

Il software libero non coincide con il software gratuito: il software distribuito gratuitamente (freeware) può benissimo essere chiuso e non libero, mentre è possibile “vendere” il software libero, le libertà sopra citate non lo impediscono.

Questo di fatto accade, tra l'altro, ad opera di grandi aziende, che forniscono insieme ai programmi liberi, assistenza e manutenzione che per l'uso professionale sono sicuramente utili.

Aziende produttrici di software possono anche avere interesse a sponsorizzare la comunità di sviluppo del software libero in quanto poi possono utilizzare liberamente i suoi prodotti per le proprie attività commerciali.

Oltre al ben noto sistema operativo Linux, esistono moltissime applicazioni per l'utente che sono software libero, disponibili anche per i sistemi operativi commerciali.

Esempi famosi sono il server web Apache, il database Mysql, il client web Mozilla Firefox, la suite OpenOffice, usata tra l'altro per scrivere questo libro e molti altri. L'utilizzo di software open source può semplificare la vita di molti professionisti e consentire anche grossi risparmi a aziende e pubblica amministrazione.

Tornando invece al sistema operativo GNU/Linux, esso presenta molti vantaggi in termini di efficienza e sicurezza rispetto alle piattaforme rivali, le interfacce utente sono ormai simili a quelle Apple e Microsoft. L'equivalente dei sistemi forniti dalle aziende commerciali sono le cosiddette “distribuzioni”, cioè unione di kernel linux e pacchetti di programmi curate da società (ed in vendita) o da comunità di sviluppo che le rendono disponibili gratuitamente. I limiti maggiori alla diffusione di questi sistemi riguardano semplicemente la conoscenza del grande pubblico (ed anche le strategie commerciali che fanno in modo che tutti i PC siano in genere venduti preinstallati) e l'assenza di corrispondenti “free” di programmi di intrattenimento e gioco. Per l'uso professionale però la loro maggiore sicurezza/stabilità li fa spesso preferire.

4.13 Il software “maligno” o malware

Un'ultima nota riguardante il software utente riguarda le problematiche di sicurezza. Un altro effetto della complicata gestione dei programmi in esecuzione sulle moderne piattaforma, è che, specie nei sistemi operativi proprietari, non è possibile controllare

l'esecuzione dei singoli programmi ed anche dei file salati sul disco rigido.

Questo fa in modo che sui calcolatori possano girare anche molti programmi “indesiderati” ed essere salvati file non noti all'utente e non previsti dal sistema. Questi possono installarsi sui PC all'insaputa dell'utente utilizzando punti deboli nei sistemi di connessione di rete o di lettura dei file da dispositivi esterni e possono in genere creare vari pericoli: dal semplice controllo dell'attività dell'utente, al danneggiamento di file, al blocco del sistema.

Si usa in genere il termine virus per indicare un software che “contagia” dei files replicandosi e non agisce come programma a sé stante (tipologia di programma malevolo non più molto diffusa). Il termine worm in genere si usa per un programma indesiderato che si avvia all'accensione del computer, trojan (cavalli di troia) sono i software indesiderati che si celano all'interno di programmi apparentemente utili. I programmi detti “spyware” vengono utilizzati per il controllo dell'attività dell'utente, la diffusione di posta elettronica indesiderata (spam) o la redirezione a siti pubblicitari su internet.

In definitiva, comunque, nonostante il pericolo di danni sia spesso sopravvalutato, occorre prendere qualche precauzione utilizzando il PC per la connessione alla rete e scambiandosi programmi via supporti esterni. L'uso di software di difesa e controllo è sicuramente consigliato (ne esistono anche di gratuiti) , ma la difesa principale consiste nell'evitare di copiare software di dubbia provenienza, scaricare file da siti non sicuri senza usare opportuni accorgimenti, e, soprattutto, aprire messaggi di posta elettronica con allegati di cui non si conosce la provenienza.

5 Le reti di calcolatori

Lo sviluppo della tecnologia delle telecomunicazioni ha permesso la nascita delle reti informatiche, usate sia all'interno delle singole organizzazioni che, in modo più vasto, tra organizzazioni differenti (aziende, città, nazioni, ecc…)

Una rete di calcolatori è un insieme di sistemi autonomi e interconnessi.
Autonomo significa che ogni sistema è in grado di funzionare in modo indipendente e completo (non ha bisogno del supporto di altri calcolatori), interconnesso significa che deve essere in grado di scambiare informazioni (file, messaggi, comandi, ecc) con altri calcolatori.

I computer collegati in rete possono essere i più vari, di marche diverse tra loro e con diverse capacità di elaborazione (dal PC al mainframe); ciascuna macchina ha delle proprie risorse (dischi fissi, stampanti, periferiche di I/O, ecc…) che possono essere condivise anche da altri sistemi.

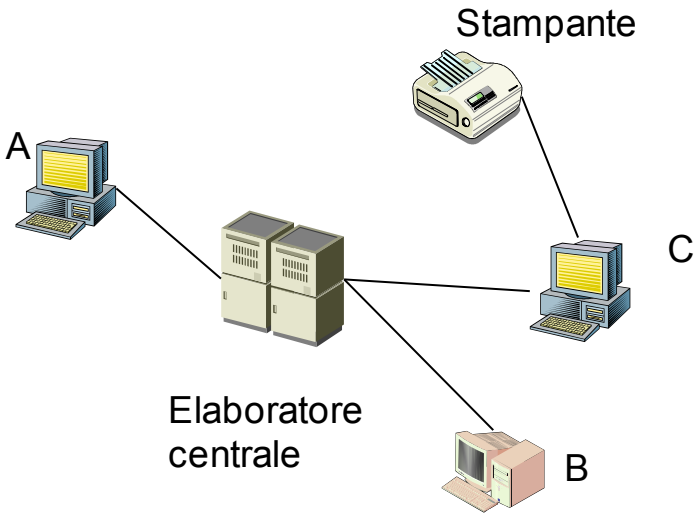

Figura 29: Rete locale di calcolatori per un ufficio.

Esistono varie tipologie di reti e varie scale di realizzazione, i meccanismi costitutivi e le problematiche di base restano comunque simili. Un semplice esempio di rete, in questo caso "locale", è quello in Figura 29. Come si vede è composta da diverse parti: un elaboratore centrale al quale potrebbe essere affidata la gestione della rete vari computer ognuno con proprie risorse, linee di comunicazione tra i vari sistemi. L'interazione tra le varie parti pone ovviamente molti problemi ai progettisti delle reti, ma i vantaggi nella loro realizzazione sono tali che l'avvento delle reti ha completamente rivoluzionato l'intero mondo dell'informatica negli ultimi vent'anni.

5.1 Motivazioni, vantaggi e svantaggi delle reti

Le reti informatiche vengono realizzate allo scopo di ottenere diversi vantaggi:
Velocità: una rete consente di impiegare in combinazione le risorse di diversi computer, e in questo modo può raggiungere una potenza di calcolo maggiore di quella di un mainframe (ad esempio: i cluster Linux sono un insieme di PC per il calcolo distribuito; ogni programma è suddiviso in parti che sono eseguite autonomamente da ogni PC del

cluster).

- Organizzazione: coordinamento delle operazioni compiute da utenti collocati in spazi remoti (posta elettronica, programmi collaborativi, etc…)
- Affidabilità: anche se un componente della rete non funziona, gli altri possono continuare a lavorare. Inoltre un calcolatore può prendere in carico le mansioni svolte da un altro andato fuori uso: le prestazioni del sistema non risentono di malfunzionamenti locali.
- Espandibilità: le risorse hardware possono essere aumentate in modo semplice.
- Condivisione di risorse: i calcolatori in rete possono facilmente condividere insiemi di dati, collezioni di documenti, dispositivi hardware. Questo consente di avere i dati sempre sincronizzati (l'ufficio acquisti di un'azienda avrà i dati di magazzino sempre aggiornati, con evidenti vantaggi per il lavoro). Inoltre l'organizzazione può risparmiare acquistando meno hardware (non serve più una stampante per ogni stazione, ma bastano poche stampanti condivise!) e meno software.
- Comunicazione: una rete può sostituire sistemi di comunicazione tradizionali (posta, telefono, TV, ecc…)

Ovviamente ci sono anche prezzi da pagare e problemi da risolvere, da valutare attentamente per ogni caso:

- Eterogeneità dei sistemi e dei loro collegamenti (con relative spese di gestione per interfacce di collegamento tra sistemi differenti e problemi di manutenzione)
- Insufficienza dei servizi: la tecnologia del software delle reti non è ancora sufficientemente matura.
- Saturazione della rete: i sistemi tradizionali di comunicazione (linee telefoniche) non erano progettati per il transito di grandi quantità di dati. Il carico eccessivo delle linee può disturbare e/o compromettere l'attività di un singolo PC (tempi di attesa lunghi, collisioni di dati sulla rete e perdita di informazioni).
- Sicurezza: più facile è scambiarsi informazioni e più difficile è garantirne la sicurezza

5.2 Mezzi trasmissivi

In una rete di calcolatori i computer possono essere distanti fisicamente gli uni dagli altri: nell'esempio in Figura 29 il computer C può essere in un ufficio, il B in un altro e l'elaboratore centrale in una palazzina a se stante, ma se parliamo di reti ampie (geografiche) le macchine possono trovarsi in continenti diversi. Le linee di interconnessione saranno di tipo differente a seconda del tipo di dati da trasmettere, della loro quantità e della distanza tra le stazioni. In una rete complessa saranno poi spesso necessari apparecchi ausiliari in grado di ripetere amplificare e distribuire i segnali.

Mezzo	Velocità	Larghezza banda	Distanza ripetitori
Doppino telefonico	**١٠-٤ Mbit/s**	**٣MHz**	**١٠-٢ Km**
Cavo coassiale	**٥٠٠ Mbit/s**	**٣٥٠ MHz**	**١٠-١ Km**
Fibra ottica	**٢ Gb/s**	**٢ GHz**	**١٠٠-١٠ Km**

Tabella 2: Tipologie di mezzi trasmissivi comuni

I mezzi trasmissivi impiegati nella costruzione di reti informatiche si distinguono principalmente in due tipi: guidati (cavi, fibre ottiche) e non guidati (segnali radio). Quantità e qualità del segnale che passa attraverso tali canali determinano le scelte delle

connessioni da realizzare. Occorre quindi considerare la larghezza di banda da cui dipende la capacità del canale (funzione anche del protocollo utilizzato per la codifica) l'attenuazione del segnale lungo lo stesso e le eventuali interferenze subite dall'esterno. Altro parametro dei mezzi trasmissivi è la latenza: il tempo che intercorre tra l'invio e la ricezione di un segnale di dimensione 0, cioè il tempo che serve per "iniziare" una trasmissione (diversa dalla velocità di trasmissione "a regime").
Tra i cavi abbiamo un largo utilizzo del doppino ritorto, cavo inizialmente dedicato alla telefonia, nelle tipologie non schermate (UTP), o schermate (STP, FTP). Le vecchie reti erano invece realizzate con il cavo coassiale ove i conduttori sono disposti in maniera concentrica, ne esistevano di due tipi "thick" (spesso) con 15 mm di diametro ed una capacità di trasporto fino a 200 Mb/s, e "thin" (sottile) con 3 mm. di diametro e adatti a trasmissioni fino a 10 Mb/s. Il mezzo guidato che sicuramente garantisce la maggior velocità di trasmissione è senza dubbio la fibra ottica. Si tratta di cavetti di vetro di dimensione micrometrica, resistenti ed elastici, formati da due materiali a diverso indice di rifrazione, che non permettono alla luce di uscire. Il segnale ottico si trasmette con pochissima attenuazione ed a velocità notevole, l'unico problema è l'interfacciamento con le apparecchiature che ridistribuiscono il segnale che rendono questo mezzo adatto solo alla trasmissione punto a punto.
Le trasmissioni mediante mezzi non guidati, cioè onde elettromagnetiche, vengono realizzate utilizzando antenne che possono essere allineate o meno (trasmissioni per uno o più ricevitori). Mentre un tempo erano riservate alla trasmissioni di grosse moli di dati p.es. via satellite, si stanno ora diffondendo anche per sistemi locali (reti "wireless").
Alcuni degli spettri di frequenze tipicamente utilizzati sono:

- 30 MHz-1GHz (trasmissioni radio)
- 2 GHz-40 GHz (microonde) collegamenti direzionali su satelliti
- 300 GHz-200 THz (Infrarossi) aree limitate

Alcune considerazioni vanno fatte anche su come i segnali da trasmettere vengano inseriti su questi canali trasmissivi. Evidentemente devono essere stabilite regole (definite in protocolli, come vedremo in seguito), ma si possono definire in generale tecnologie o modalità differenti di trasmissione. Per esempio si parla di trasmissione sincrona se i due terminali della trasmissione hanno un orologio comune e concordano su inizio e termine

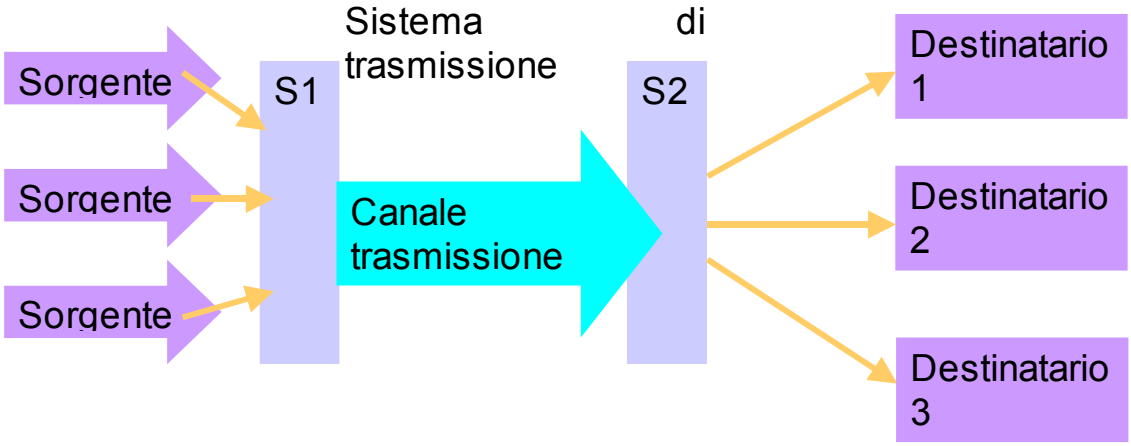

Figura 30: Trasmissione di più messaggi su un unico canale mediante multplexing/demultiplexing

della trasmissione, asincrona se non c'è orologio e l'inizio e la fine della trasmissione vengono comunicati con opportuni segnali.
Spesso poi, per sfruttare appieno un collegamento tra due punti, è opportuno far passare

per lo stesso canale diversi messaggi allo stesso tempo. Questo vene realizzato con una tecnica detta **multiplexing**. I diversi segnali vengono trasformati nell'unico che attraversa il canale da un apparecchio detto "multiplexer" e separati nuovamente all'arrivo a destinazione da un "demultiplexer". L'unione può essere fatta alternando i segnali nel tempo (multiplexing di tempo) o modulandoli su bande di frequenza diverse (multiplexing di frequenza).

Circuiti commutati e dedicati

Esistono diversi metodi per collegare due stazioni tra di loro, che dipendono dall'architettura dei circuiti. Il primo tipo di circuito è detto a commutazione, ed è quello che ugualmente viene usato nella normale rete telefonica. Questo sistema prevede che l'utente stabilisca la comunicazione (componendo il numero del destinatario), dopodiché la centrale di commutazione provvede a collegare i due apparecchi. In questo modo qualsiasi apparecchio della rete può collegarsi con un altro.
Il secondo tipo di circuito è chiamato circuito dedicato. Questo prevede un collegamento fisico diretto tra le due stazioni che vogliono comunicare. Non occorre alcuna fase per stabilire la comunicazione e non occorre alcuna centrale di commutazione.
Errore: sorgente del riferimento non trovataLa figura 31 mostra un esempio di rete commutata. Gli host, per esempio A,B,C, sono le postazioni di lavoro di utenti remoti o di reti locali remote e possono essere collegate seguendo vari percorsi e seguendo diverse politiche per quanto riguarda l'allocazione delle risorse di rete.

Circuiti a commutazione di pacchetto

I circuiti commutati si dividono in due tipi: *a commutazione di circuito* e *a commutazione di pacchetto*. Il caso della centrale telefonica appena presentato è quello di un sistema a commutazione di circuito, cioè la centrale predispone un collegamento fisico (momentaneo, cioè per l'intera durata della telefonata) e continuo (cioè non vi possono essere interruzioni durante lo svolgimento della comunicazione.
Questo tipo di sistema è adatto per trasportare segnali che per loro natura sono "continui" (ad esempio la voce… non riusciremmo a capire nulla se il segnale arrivasse a pezzetti in presenza di tempi morti), ma non è altrettanto adatto per il mondo dei calcolatori dove i

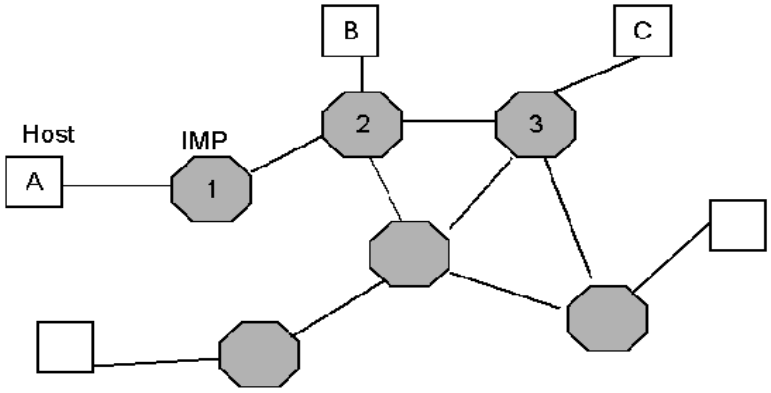

Figura 31: Esempio di rete commutata

segnali da trasmettere sono digitali (sequenze di bit).
Il sistema a commutazione di circuito, infatti, genera delle inefficienze. Il collegamento, essendo esclusivo, non può essere ottimizzato trasportando anche dati di altri utenti durante i "tempi morti" della comunicazione. Questo fatto causa una effettiva riduzione

dell'informazione che può essere trasmessa in rete (calcolatori che volessero dialogare tra loro troverebbero la linea occupata!).
Per questo motivo nel mondo dell'informatica si è preferito il sistema a commutazione di pacchetto. Tale sistema non prevede un collegamento continuativo e costante tra le stazioni interessate; i dati vengono trasmessi a blocchi (pacchetti), di grandezza prestabilita. La comunicazione avviene nel seguente modo: la stazione sorgente suddivide i dati in pacchetti di bit; in ogni pacchetto viene messo l'indirizzo del destinatario e un numero sequenziale. I pacchetti, indipendentemente, vengono trasmessi sulla rete e memorizzati da nodi intermediari (detti Interface Message Processor (IMP) e nel caso specifico del livello di internet router). In base alle caratteristiche del traffico di rete (presenza di ingorghi, etc…) i nodi intermediari possono instradare (routing) verso altri nodi i vari pacchetti in modo diverso. I pacchetti viaggiano di nodo in nodo fino a raggiungere la stazione di destinazione. È compito della stazione attendere l'arrivo di tutti i pacchetti e "riordinarli" in base al numero di sequenza per ottenere i dati originali. Si osservi che con un sistema di questo genere può succedere che arrivi prima alla stazione di destinazione un pacchetto partito dopo dalla stazione sorgente. È compito della stazione finale anche accorgersi di non aver ricevuto tutti i pacchetti e/o controllare che non vi siano stati errori di trasmissione (richiedendo, in tal caso, la ritrasmissione dei dati).
Tutto questo meccanismo di inoltro dei pacchetti, chiamato store and forward, viene utilizzato nella maggioranza delle reti di calcolatori.
I vantaggi di questo tipo di collegamento sono i seguenti:

- Gli utenti pagano in base alla quantità di dati effettivamente trasmessi.
- I collegamenti sono ottimizzati, in quanto condivisi da più utenti.
- Da una stessa stazione possono partire contemporaneamente più comunicazioni inviate a stazioni differenti.
- I nodi intermedi cercano automaticamente la via più efficiente sulla quale instradare i pacchetti (risolvendo anche problemi nel caso di guasti su alcune tratte di collegamento).

Gli svantaggi sono invece:

- La linea è condivisa e quindi diminuisce la velocità di trasmissione (rispetto ad una linea dedicata, la velocità media di trasmissione è più bassa).
- Ad ogni pacchetto vengono aggiunte informazioni di controllo e indirizzi che ne aumentano la dimensione, e diminuisce quindi l'effettiva velocità di trasmissione.

Si osservi che nel collegamento a commutazione di pacchetto non esiste una linea fisica permanentemente collegata tra due stazioni. La connessione fisica cambia continuamente in base al traffico della rete e, ad un livello di astrazione più alto, si dice semplicemente che tra le due macchine esiste una connessione logica (cioè si sa solo che la macchina A comunica con la macchina B, ma non è possibile definire, in un dato istante, un percorso fisico di collegamento).
I vantaggi della commutazione di pacchetto sono tali che anche la comunicazione voce ormai viene effettuata in modo più efficiente utilizzando i protocolli di rete dei calcoalatori (Internet Protocol) basati sui pacchetti e si può oggi telefonare a basso costo utilizzando la cosiddetta telefonia "over IP", cioè realizzata con la rete ed i protocolli di Internet.

5.3 Classificazioni delle reti.

Le reti si distinguono per dimensione, topologia e tecnica di trasferimento dei dati. Il collegamento tra computer avviene a diversi livelli, dando vita a reti di dimensioni diverse, che collegano gli elaboratori a distanze differenti in aree geografiche più o meno vaste.

Classificazione per dimensione:

LAN *Local Area Network* (rete in area locale).
Si tratta in pratica di reti che collegano computer collocati a breve distanza fra loro. Un'area locale può essere rappresentata da un'area aziendale con diversi uffici i quali hanno computer collegati tra loro.
MAN *Metropolitan Area Network* (rete in area metropolitana)
In questo caso sono collegati computer che si trovano all'interno di una determinata area urbana (esempio, gli uffici dell'anagrafe comunale sono sparsi nei vari quartieri, ma i loro computer sono collegati tutti alla sede centrale).
WAN *Wide Area Network* (rete in area vasta)
Questa rete copre un'area più ampia che comprende di regola il territorio nazionale, fino ad arrivare a collegare calcolatori collocati in diversi stati limitrofi.
GAN *Global Area Network* (rete globale)
È l'ultimo livello e consiste nella rete che collega i computer collocati in tutto il mondo, anche via satellite.

Classificazione topologica:

Topologia a stella
I computer sono collegati a una stazione centrale, la quale provvede a collegare due computer tra loro quando vengono trasmessi dati. Tutti i dati trasmessi da una stazione all'altra transitano nella stazione centrale.

Vantaggi:

- Per ogni stazione esistono le stesse condizioni di trasferimento, in quanto vi è solo la stazione centrale interposta tra le stazioni. Questo significa tempi di trasmissione uguali per tutti.
- In caso di avaria di una delle stazioni (esclusa la stazione centrale), le altre non ne risentono minimamente, e possono tranquillamente continuare il lavoro e le comunicazioni tra di loro.

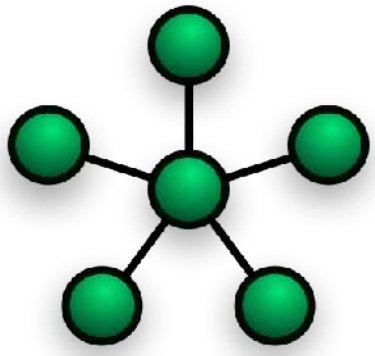

Figura 32: Topologia di rete a stella

Svantaggi:

- La rete diventa inutilizzabile in caso di avaria della stazione centrale.
- La stazione centrale raggiunge costi molto elevati nel caso di reti complesse con molti terminali.
- Una stazione centrale "sottodimensionata" rappresenta un collo di bottiglia che rallenta il lavoro dell'intera rete.
- Non è facile adattare questo tipo di rete alle caratteristiche logistiche di un'azienda con uffici dislocati disordinatamente. Se gli uffici da collegare sono distanziati tra loro, il collegamento tra stazione centrale e periferiche presenta delle difficoltà di carattere fisico (i cavi hanno dei limiti in lunghezza, ci sono problemi di attenuazione del segnale e problemi riguardanti i disturbi sulla linea).

Topologia ad anello

In questa configurazione non è necessaria una stazione centrale; qui le stazioni sono

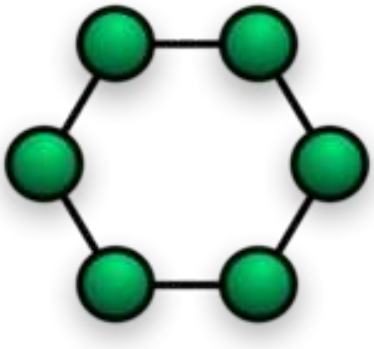

Figura 33: Topologia di rete ad anello

collegate in cerchio. Questo elimina alcuni svantaggi visti precedentemente (ma ne importa di nuovi). Nella struttura ad anello ogni stazione che riceve i dati trasmessi li memorizza e li passa poi alla stazione successiva, fino a che non raggiungono il destinatario.

Vantaggi:

- La rete è facilmente espandibile.
- Non è necessaria una stazione centrale (vengono meno i problemi relativi al malfunzionamento della stazione centrale)
- I tempi di attesa per il proprio turno di trasmissione possono essere facilmente calcolati e previsti.

Svantaggi:

- A meno che non vengano raddoppiati i collegamenti tra le stazioni, si rischia il collasso dell'intera rete in caso di avaria di un solo collegamento.
- I tempi di trasmissione aumentano all'aumentare del numero di stazioni presenti nella rete.
- Difficile adattamento alla logistica tipica di un'azienda.

Topologia a bus

La struttura a bus è del tipo **broadcast**: i dati trasmessi da una stazione arrivano, senza memorizzazioni intermedie, direttamente alla stazione destinazione. Questo comporta velocità di trasferimento maggiori rispetto alle reti ad anello e stella. Le reti con struttura a bus sono quelle più diffuse.

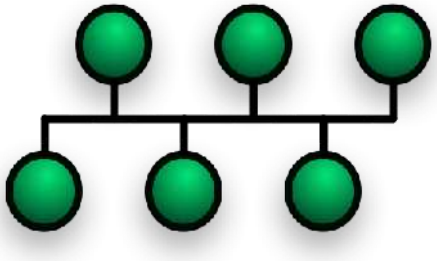

Figura 34: Topologia di rete a bus

Vantaggi:

- Il collegamento principale (bus) è passivo (è un semplice cavo che non compie alcuna operazione) e non è soggetto a malfunzionamenti.
- I costi per l'espansione della rete sono molto contenuti e l'operazione è semplicissima.
- È facilmente realizzabile il broadcasting (trasmissione diretta dei dati a TUTTE le stazioni della rete) o il multicasting (trasmissione a gruppi di calcolatori).
- La rete non risente delle avarie delle stazioni collegate.
-

Svantaggi:

- La linea di collegamento può essere occupata solo da una stazione trasmittente: non è possibile la trasmissione da più stazioni contemporaneamente (se ciò avviene si hanno delle collisioni di pacchetti e occorre un sistema per evitare e/o risolvere questa evenienza).
- La struttura a bus non è facilmente realizzabile con condutture in fibra ottica (che offrono prestazioni molto più elevate dei normali cavi coassiali solitamente impiegati).

5.4 Protocolli di comunicazione

La comunicazione tra due macchine avviene solo se tra i due sistemi esiste un insieme comune di regole per il passaggio delle informazioni (proprio come tra gli umani: per avere

una comunicazione due persone devono avere regole comuni, cioè un linguaggio).
Un protocollo è uno speciale insieme di regole che stazioni connesse usano quando comunicano. I protocolli esistono a diversi livelli che vanno da livelli fisici (hardware) fino a livelli logici (software) più "vicini" all'utente.
Cerchiamo ora di capire meglio quali informazioni sono contenute in un protocollo di collegamento. Abbiamo innanzitutto un'indicazione che riguarda l'apparecchio con il quale ci si vuole collegare, in pratica un indirizzo del destinatario dei dati che si stanno per spedire. Vi è poi indicata la velocità di trasferimento dei dati, in modo che sia la stazione ricevente che quella di trasmissione siano sincronizzate, altrimenti si avranno perdite di dati.
Infine, vi saranno vari altri dati riguardanti il controllo degli errori di trasmissione e dati utilizzabili per il ripristino dei dati perduti o modificati durante la trasmissione. Vi sono in particolare informazioni che indicano come una determinata stazione deve comportarsi in caso di errore: se occorre ripetere il trasferimento o se i dati possono essere ripristinati.
Errori nella trasmissione o distorsioni dei dati si possono verificare in casi di malfunzionamento della rete o di interferenze dall'esterno nella rete di trasmissione, che spesso è la rete telefonica, soggetta a disturbi, oppure ancora da un difetto di una delle stazioni, ricevente o trasmittente.
I dati appartenenti ai protocolli di trasmissione vengono elaborati in diversi stadi della trasmissione e vengono via via aggiunti alle informazioni provenienti dagli stadi precedenti. In questo modo si ottiene un sistema simile alle matriosche, quelle bamboline russe contenute una dentro l'altra.
Nel caso dei protocolli di collegamento, il primo protocollo viene inserito all'interno di un secondo protocollo più specifico, contenente altre informazioni; questo a sua volta viene inglobato in un ulteriore protocollo e così via fino ad arrivare allo stadio più basso, che rappresenta il collegamento fisico vero e proprio tra le due stazioni che sono entrate in comunicazione.
La stazione ricevente poi provvederà a "scartare" i singoli protocolli eseguendo le operazioni della stazione mittente a ritroso, e ai rispettivi stadi elminerà i dati che non servono al livello superiore, ripristinando infine le informazioni come sono state inviate dal livello più alto della stazione trasmittente.

Un esempio per chiarire: supponiamo che vi siano due filosofi, uno italiano e l'altro cinese, che vogliono comunicare scambiandosi le proprie idee. Il problema che sorge è duplice: innanzitutto essi non si capiscono per via della diversa lingua, e inoltre non sono esperti di comunicazione (non conoscono l'uso del telefono), quindi non sanno come fare a parlarsi.
Poiché tuttavia la necessità di comunicare è più forte, si decidono a costruire entrambi un sistema di comunicazione basato su tre stadi.
Il primo stadio è quello dei filosofi che riflettono e che producono le informazioni da scambiare.
Il secondo stadio è quello dei traduttori, che traducono le informazioni provenienti dallo stadio precedente in una lingua universale, ad esempio l'inglese.
Il terzo stadio è rappresentato dai fattorini e dagli impiegati postali che effettivamente creano il ponte tra un continente e l'altro, trasferendo le informazioni.
Le caratteristiche di questo sistema sono le seguenti: ognuna delle persone coinvolte percepisce il collegamento in modo orizzontale; ciò significa che il filosofo cinese che riceve le informazioni dal filosofo italiano le ottiene così come sono state pensate in origine, senza distorsioni o aggiunte. In questo modo, a entrambi i filosofi sembra di essere collegati direttamente, senza intermediari.
In realtà ogni collegamento, a eccezione dell'ultimo stadio, avviene in modo verticale: il

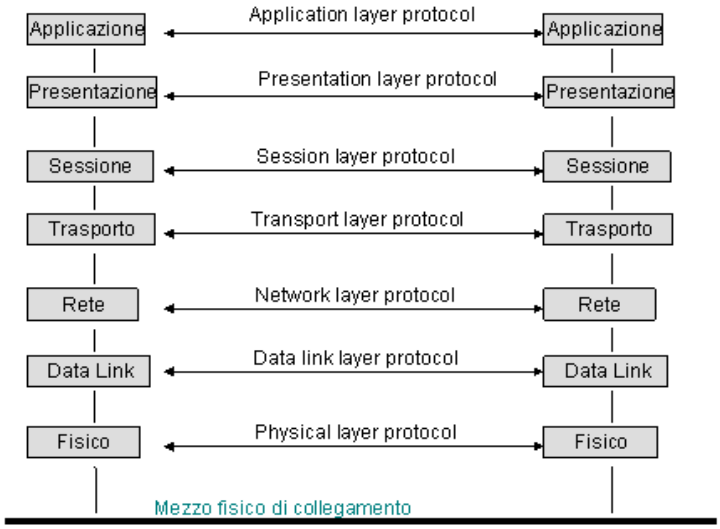

Figura 35: I sette livelli del protocollo ISO-OSI

filosofo italiano passe le informazioni al traduttore, che è un gradino più in basso, e quest'ultimo le passa a sua volta al fattorino che effettua la spedizione.
I protocolli di comunicazione sono pressoché indipendenti fra loro ciò significa che i due filosofi possono riflettere su qualsiasi argomento a loro piacere, usando un linguaggio proprio dei filosofi. Le loro *"frasi"* saranno contenute in un determinato protocollo. I traduttori a loro volta possono usare una lingua a loro piacimento, che può essere l'inglese, il francese, il tedesco o un'altra, indicandola nel protocollo specifico. Infine, i tecnici del livello "reale" possono usare il mezzo di comunicazione più idoneo: la posta, il telefono, il telegrafo o altro. Tutto ciò senza che le persone precedenti o successive risentano di questi cambiamenti nei protocolli dello stadio precedente o successivo. La comunicazione è del tutto trasparente.
In nessuno degli stadi più bassi (traduttori e tecnici) vengono interpretate le informazioni provenienti dallo stadio precedente; ciò significa che i traduttori si limitano a tradurre senza aggiungere, togliere o commentare nulla. I tecnici a loro volta si limitano a trasferire le informazioni dall'uno all'altro senza modificarle. In entrambi gli stadi ci si limita ad aggiungere solo le informazioni che fungono da chiave di lettura per i rispettivi stadi riceventi, in modo che questi possano ritrasformare le informazioni, portandole allo stadio iniziale.
Le persone riceventi si preoccupano di ritrasformare le informazioni togliendo a ogni stadio i dati contenuti nel protocollo specifico del proprio livello. In questo modo, quando le informazioni raggiungono il vertice (il filosofo), saranno ripulite da dati superflui, in modo che risultino tali e quali a quelle inviate in origine dall'altro filosofo trasmittente.
Tutti i protocolli, tranne quello dei tecnici, sono virtuali; ciò significa che non servono realmente al collegamento effettivo — fisico — tra trasmittente e ricevente. L'unico protocollo che effettivamente conta per il trasferimento vero e proprio è appunto quello dei tecnici.
Come si può notare da quanto spiegato in precedenza, i collegamenti tra computer non sono semplici da costruire, in quanto occorre tenere conto di molti fattori. Di queste difficoltà, comunque, l'utente finale non deve preoccuparsi, proprio perché il collegamento avviene in modo che si percepisca effettivamente soltanto ciò che interessa, mentre tutto il resto viene fatto negli stadi sottostanti. L'utente finale quindi non si dovrà preoccupare di elaborare un protocollo di traduzione o un protocollo di trasferimento fisico, né tantomeno

del mezzo usato. Il filosofo (utente) non deve fare altro che pensare e riflettere, producendo le informazioni: al resto pensano gli altri (il programma e l'hardware utilizzati per il collegamento). L'unica scelta lasciata all'utente riguarda proprio il tipo di hardware e software da impiegare, proprio come il filosofo cerca il traduttore che più gli va a genio o quello più capace.

Figura 36: L'esempio dei filosofi

5.5 Il modello ISO/OSI

Il collegamento e la cooperazione tra sistemi informatici che utilizzano sistemi operativi incompatibili tra loro è una delle principali esigenze del mercato attuale. I sistemi capaci di interagire tra loro, pur basandosi su sistemi operativi incompatibili, sono detti aperti quando permettono le comunicazioni in accordo con gli standard specificati nel modello generale Open System Interconnection (OSI). Questi standard sono stati definiti da una speciale commissione dell'International Standard Organization (ISO), ossia l'agenzia dell' ONU responsabile degli standard internazionali, inclusi quelli delle comunicazioni.
Questi standard sono nati come risposta alla diffusa esigenza di interconnettere tra loro sistemi incompatibili. La difficoltà di fondo consiste nel far comunicare tra loro due o più processi che usano, internamente, regole e tecniche diverse. Ci si è preoccupati quindi di definire le strutture dei dati trasmessi, le regole e i comandi per la gestione dello scambio dati tra applicazione o tra utenti, i meccanismi di controllo che assicurano uno scambio senza errori.
Il comitato ISO ha stabilito le regole e le opzioni per tali interazioni, definendo un modello

di riferimento. Un modello di riferimento è cosa diversa da un'architettura di rete:
un *modello di riferimento* definisce il numero, le relazioni e le caratteristiche funzionali dei livelli, ma non definisce i protocolli effettivi;
una *architettura di rete* definisce, livello per livello, i protocolli effettivi
Un modello di riferimento, quindi, non include di per sé la definizione di protocolli specifici, che invece vengono definiti successivamente, in documenti separati, come appunto accaduto dopo l'introduzione del modello ISO/OSI.
Tale modello suddivide le necessarie funzioni logiche in sette diversi strati funzionali, detti layer (livelli), come in Figura 41.
L'insieme dei 7 layer garantisce tutte le funzioni necessarie alla rete comunicativa tra sistemi, nonché una gamma molto ampia di funzioni opzionali (come ad esempio la compressione e la cifratura dei dati): in tal modo, si è in pratica suddiviso un compito complesso in un insieme di compiti più semplici.

Al fine di descrivere i 7 strati funzionali *(layer)* di cui si compone il *modello ISO/OSI,* esaminiamo velocemente i vari comportamenti di una stazione connessa ad una rete (di qualsiasi tipo):
Il *livello fisico* si preoccupa di normalizzare le caratteristiche dei mezzi fisici utilizzati dai gestori di reti pubbliche. In particolare sono definite dal CCITT le caratteristiche elettriche, funzionali e procedurali, e dall'ISO stesso le caratteristiche meccaniche. Le caratteristiche elettriche si riferiscono alle tensioni usate e alla modulazione dei segnali, quelle funzionali e procedurali riguardano il metodo e le regole di collegamento tra workstation e nodi di commutazione. Le caratteristiche meccaniche invece riguardano il tipo di porta usata e il collegamento dei vari apparecchi a tali porte. Un esempio di queste porte normalizzate sono le porte V.24, più conosciute come RS-232, e le porte V.25 e X.21.
Il *livello di collegamento dati (data link)* si preoccupa di trasmettere e ricevere dati privi di errori. La funzione di questo livello è quella di rilevare ed eventualmente correggere gli errori provocati dalla trasmissione dei dati nei mezzi fisici. Un altro compito di questo livello è quello di stabilire, mantenere e chiudere un collegamento virtuale tra due stazioni.
Il terzo livello, detto *di rete (network),* ha il compito di collegare due stazioni anche senza un collegamento fisico, ma mediante collegamenti virtuali. Il livello di rete inoltre ha il compito del controllo del flusso dei dati. Un'altra importante funzione del livello di rete è quella di collegare sottoreti differenti. Ogni stato ha infatti un sistema di rete pubblica, che può non coincidere con quelle di altri stati. In questi casi il livello di rete dell'OSI prevede l'aggiunta di servizi a quelle sottoreti che non raggiungono i livelli previsti dall'ISO e ne toglie a quelle sottoreti che invece ne hanno in eccedenza. In questo modo si rende omogeneo il collegamento tra computer di paesi diversi, così che, ad esempio, un computer in Italia può collegarsi tranquillamente con un computer negli Stati Uniti anche se le sottoreti dei paesi in questione hanno caratteristiche diverse quanto a servizi offerti. I collegamenti tra le due sottoreti avvengono mediante i gateway*,* cioè vie di accesso o corridoi. Questi gateway sono specifici calcolatori che provvedono a trasformare i protocolli provenienti da una sottorete negli equivalenti protocolli di un'altra sottorete.
Il *livello di trasporto* provvede a eliminare ulteriori errori nella trasmissione dei dati, conferendo in questo modo una buona affidabilità ai trasferimenti, indipendentemente dalla sottorete della quale si avvale l'utente. Tra i compiti del livello di trasporto vi sono il recapito nella sequenza corretta dei pacchetti e l'eventuale riordino. Il livello di trasporto si limita soltanto alla trasmissione dei dati, mentre dell'elaborazione e della modifica si occupa l'ultimo livello applicativo.
Con questo livello termina la descrizione dei livelli di comunicazione; il prossimo livello fa parte del secondo gruppo, dei livelli di elaborazione.

Il *livello di sessione* ha come compito principale quello di sincronizzare e organizzare i dati da trasferire provenienti dai livelli superiori. Ogni connessione di sessione può attivare più connessioni di trasporto. Un ultimo compito molto importante svolto dal livello di sessione è il ripristino del collegamento dopo possibili cadute. Si tratta della cosiddetta *risincronizzazione,* che permette di riprendere il trasferimento dei dati dal punto in cui era stato interrotto. In questo modo si elimina il problema delle ritrasmissioni di dati, contribuendo a contenere i costi di trasmissione.
Il *livello di presentazione* invece ha il compito di rendere possibile il trasferimento di dati tra computer con architetture differenti, che altrimenti non si comprenderebbero. Tramite appositi protocolli di conversione, stabiliti in fase di contrattazione o di formazione del collegamento, i dati di un computer vengono trasformati in codice comprensibile anche al computer remoto. Si pensi ad esempio a due computer, funzionanti uno con codice ASCII e l'altro con codice EBCDIC (un altro codice per la codifica dei caratteri alfanumerici).
Il *livello di applicazione* elabora i dati in accordo alle richieste dell'utente. È il livello più vicino all'utente e serve come interfaccia per comprendere i comandi che l'utente vuole impartire. Al livello di applicazione appartengono tutti i programmi (e protocolli) che servono a comunicare in rete e a cui l'utente ha diretto accesso (ftp, http, telnet, etc…).

6 Internet

L'utilizzo principale delle reti di calcolatori fatto dalla maggioranza degli utenti consiste nel collegamento dei PC domestici o di lavoro ad Internet. Ma cos'è Internet?
Sostanzialmente Internet (con la I maiuscola) è la prima ed unica rete di computer mondiale ad accesso pubblico.
Essa è una rete di reti che connette sottoreti di macchine di calcolo in tutto il mondo mediante un unico spazio di indirizzi e protocolli standard TCP/IP, rispettivamente protocollo di trasporto internetwork, e protocollo di network.
Quindi è una rete commutata come quelle che abbiamo visto prima, che utilizza l'efficiente tecnologia della commutazione di pacchetto per trasferire i messaggi tra gli host.
La rete è costituita da una serie di nodi (host) e di sottoreti eterogenee, collegate da apposite macchine (quelle che avevamo definito Interface Message Processors) che fanno da traduzione tra i messaggi ai vari livelli dell'architettura (bridge, router, gateway).

6.1 Connettersi ad Internet

L'utente domestico non è collegato direttamente alla rete, ma accede attraverso i cosiddetti Internet Service Provider, in genere tramite connessione telefonica via Modem oppure mediante connessione ADSL.

ADSL (Asymmetric Digital Subscriber Line) è una tecnologia che permette di trasformare la linea telefonica analogica (il tradizionale doppino telefonico in rame) in una linea digitale ad alta velocità per un accesso ad Internet ultra-veloce.

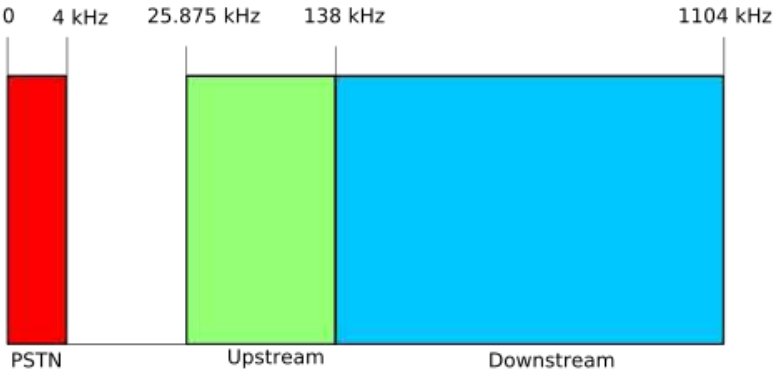

Figura 37: La suddivisione delle frequenze nei canali voce e dati utilizzata in ADSL.

La tecnologia ADSL utilizza la banda delle frequenze superiori a quella normalmente utilizzata per la voce (fonia) nella linea telefonica tradizionale per trasportare in forma numerica dati, audio e immagini.
Quando sulla linea telefonica è abilitata ADSL, la banda disponibile viene suddivisa in tre sotto-bande di frequenza distinte: una dedicata alla ricezione dei dati da Internet (downstream), una dedicata all'invio dei dati verso Internet (upstream) ed un'altra molto ristretta riservata al traffico voce.
Per utilizzare il servizio è richiesta l'installazione di un'apposita coppia di filtri ADSL alle estremità della linea telefonica (presso la centrale telefonica e l'abitazione). Ci sono due possibili tipi di connessione, che garantiscono entrambi l'utilizzo della linea sia in fonia che dati. Nel primo caso (il più comune nel caso di utenze domestiche) si installa su ogni presa

telefonica un dispositivo che prende il nome di *microfiltro* che permette di utilizzare in ogni diramazione sia il servizio ADSL che il comune telefono o fax. Il microfiltro separa il segnale telefonico standard dai diversi segnali che viaggiano sullo stesso cavo (le linee dati di ADSL): il telefono o il fax sono quelli usuali poiché a monte del filtro il protocollo di trasmissione è quello usuale (cioè il segnale è solo uno: quello telefonico!).

Nel secondo caso viene installato un dispositivo simile al microfiltro, detto *splitter***,** che si preoccupa di effettuare questo "taglio" di frequenze all'ingresso della linea ADSL nell'edificio. In questo secondo caso, la versatilità risulta essere ridotta, dato che la linea ADSL non arriverà necessariamente fino a tutti gli ambienti dove invece è presente la linea telefonica. Il lato positivo è dato dal fatto che non è necessario disseminare tutte le prese telefoniche con i microfiltri.
Nella *centrale di commutazione* come è avvenuto nella sede del cliente un filtro separa le frequenze della linea telefonica da quelle della linea dati. In questo modo, le frequenze telefoniche prendono la strada della *PSTN* (Public Switched Telephone Network / Rete Telefonica Pubblica a Commutazione, la rete telefonica "normale"), invece quelle dati (ADSL) vengono convogliate verso un dispositivo che prende il nome di *DSLAM* (DSL Multiplatore di Accesso) che fa da interfaccia di collegamento con la rete Internet.

Le velocità attualmente disponibili per le connessioni ADSL sono dell'ordine dei 4 Mbps in ricezione e 640 in trasmissione, ma stanno diventando disponibili le offerte *ADSL 2+*, a 12, 20 e 24 megabit. Si tenga però conto che le velocità reali sono in genere più basse perché la banda è frazionata fra molti utenti.

I principali vantaggi offerti da ADSL sono:

- Linea telefonica indipendente dalla linea dati: è possibile usare il telefono e contemporaneamente usare Internet.
- Velocità di navigazione e trasferimento file teoricamente molto più alte di quelle di una linea digitale ISDN.
- Costi contenuti (non richiede nell'abitazione l'installazione di linee telefoniche supplementari o speciali)
- Il collegamento sempre attivo consente la ricezione immediata degli e-mail e di essere subito avvisati appena arriva un messaggio.
- L'utente è permanentemente connesso a Internet, senza la necessità di attivare ogni volta la connessione via modem.

Proprio il fatto di essere sempre collegati a Internet rappresenta un fattore di rischio per gli utenti. Infatti una connessione continua li espone all'attacco di **hacker**, persone che cercano di accedere, senza autorizzazione, alle risorse dei computer (aziendali e non). Occorre quindi prevedere una protezione adeguata (ad esempio installare sui gateway aziendali dei **firewall**, cioè un insieme di programmi che prevengono l'accesso non controllato da parte di altri utenti).

6.2 Servizi di Internet

Internet viene principalmente utilizzata per fornire e fruire di servizi di vario tipo.
Un servizio di Internet è un'architettura software (cioè un insieme di programmi che collaborano tra loro per svolgere un determinato compito) che utilizza per la gestione e l'ordinamento dei pacchetti il protocollo IP.
In generale, esistono due architetture di base per i servizi di rete: client/server e peer to peer.

Architettura client/server

Un processo servente (server, o nella terminologia più moderna daemon) offre un servizio ai processi clienti (client, o nella terminologia più moderna agent), che sono i soli abilitati ad iniziare la comunicazione.

Un processo servente:

- aspetta richieste da un processo cliente
- può servire più richieste allo stesso tempo
- applica una politica di priorità tra le richieste
- può attivare altri processi di servizio
- è più robusto e affidabile dei clienti
- tiene traccia storica dei collegamenti (logfile)

Un esempio può aiutare a comprendere meglio. Supponiamo che un utente tramite il suo Browser (il client del web, es. Internet Explorer, Safari, Mozilla Firefox, ecc) voglia leggere il contenuto della pagina principale del sito http://elvira.univr.it/index.html
Sula macchina che corrisponde al dominio elvira.univr.it (che ha indirizzo IP 157.27.10.55) è attivo un programma servente che si chiama HTTPD (HTTP Daemon). Il cliente (Explorer, in questo caso) invia la richiesta "GET http://elvira.univr.it/index.html" e il server HTTPD cercherà il documento (file) corrispondente e lo invierà al cliente. Notare che il server non fa nulla finché non arriva una richiesta del client.
Il server HTTPD è sempre in ascolto e può servire più client contemporaneamente (cioè più utenti possono fare richieste contemporanee). Se richiesto da opportune istruzioni contenute nel documento (file) al quale si vuole accedere, HTTPD può attivare altri processi (es. il programma che conta il numero di accessi). Infine HTTPD registra ogni richiesta (il nome della macchina chiamante, NON l'identificativo dell'utente) in un opportuno file (logfile) che serve a fini statistici o per controllare il verificarsi di errori che impediscano agli utenti una corretta lettura dei documenti.

Architettura peer-to-peer

Si dice peer to peer una rete in cui i nodi sono tutti equivalenti. In tale rete due applicazioni su due computer diversi possono indifferentemente iniziare una comunicazione. Questo avviene per esempio nei ben noti applicativi di file sharing, anche se in genere i servizi sono spesso in realtà ibridi e usano magari un server per la ricerca dei nodi contenenti i file.
Esempio: un utente vuole scaricare dalla rete delle canzoni. Dopo essersi registrato on-line in un sito opportuno (pagando l'eventuale abbonamento scelto) ha l'accesso a dei cataloghi di canzoni mediante un programma fornitogli con l'abbonamento. Seleziona il nome del cantante e le canzoni che gli interessano dal menu interattivo. A questo punto il processo iniziato dall'utente termina e viene avviata una procedure per la quale in qualunque momento un nodo che contiene dati può iniziare autonomamente la trasmissione alla macchina dell'utente. Questa è una comunicazione peer-to-peer, poiché la comunicazione è stata iniziata una volta da una macchina (quella dell'utente) e una volta da un altra.

Un processo peer-to-peer:

- richiede una procedura di registrazione dell'utente che vuole accedere al servizio.
- può chiedere informazioni a server specifici (es. fornire dei cataloghi di prodotti)
- mette a disposizione sulla rete dei file dichiarati pubblici (cioè leggibili e scaricabili)

- può servire più richieste allo stesso tempo
- non è particolarmente robusto e affidabile
- di solito NON tiene traccia storica dei collegamenti

Si noti che anche altri servizi “classici” di Internet, come ad esempio il servizio di distribuzione delle news (protocollo NNTP, vedi dopo) sono organizzati in modo peer to peer.
I servizi e relativi protocolli per lo scambio di file (es. bitTorrent) hanno avuto grosso successo in quanto la distribuzione del carico relativo alla trasmissione di file in una serie di connessione indipendenti con molti nodi invece che con un unico server permettono di sfruttare molto più efficacemente la banda complessiva disponibile nella rete (in pratica di “scaricare” molto più velocemente i file). Si noti che l'uso degli applicativi di file sharing e dei protocolli di condivisione di file p2p non è illegale, ma potrebbe esserlo rendere disponibile materiale non libero.

Servizi sincroni ed asincroni

Un servizio di Internet può essere:

- Sincrono attività simultanea tra i siti
- Asincrono non richiede attività simultanea
- Esempi di servizi di Internet sono:
- Servizi di tracciamento: verifica dell’esistenza e connessione di un account o host su Internet
- Servizi di comunicazione: per scambio messaggi, flussi di dati o programmi fra due o più corrispondenti
- Servizi di cooperazione: condivisione e modifica di risorse (dati, programmi, documenti) fra più corrispondenti
- Servizi di coordinazione: attività concordata di persone, servizi o programmi

Servizi	Asincroni	Sincroni
Tracciamento	finger	Ping
Comunicazione	E-mail, news, forum,	Chat, streaming audio/video, conferencing
Cooperazione Coordinazione	Ftp, www, Wiki, e-commerce	File sharing, Multi user games,

Tabella 3: Servizi di Internet sincroni ed asincroni

Uniform Resource Locator

Un aspetto particolare del funzionamento di molti dei servizi di Internet è la tecnica di indirizzamento dei documenti, ovvero il modo in cui è possibile far riferimento ad un determinato documento tra tutti quelli che sono pubblicati sulla rete.

La soluzione che è stata adottata per far fronte a questa importante esigenza si chiama *Uniform Resource Locator* (*URL*). La 'URL' di un documento corrisponde in sostanza al suo indirizzo in rete; ogni risorsa informativa (computer o file) presente su Internet viene rintracciata e raggiunta dai nostri programmi client attraverso la sua URL. Prima della introduzione di questa tecnica non esisteva alcun modo per indicare formalmente dove fosse una certa risorsa informativa su Internet.
Una URL ha una sintassi molto semplice, che nella sua forma normale si compone di tre

parti:

```
protocollo://nomehost/nomefile
```

La prima parte indica con una parola chiave il tipo di server (o meglio il protocollo di comunicazione cui il server risponde) a cui si punta (può trattarsi di un server gopher, di un server http, di un server FTP, e così via); la seconda indica il nome simbolico dell'host su cui si trova il file indirizzato; al posto del nome può essere fornito l'indirizzo numerico; la terza indica nome e posizione ('path') del singolo documento o file a cui ci si riferisce. Tra la prima e la seconda parte vanno inseriti i caratteri '://'.
Un esempio di URL è il seguente:

http://www.liberliber.it/index.html

La parola chiave 'http' segnala che ci si riferisce ad un server *Web*, che si trova sul computer denominato 'www.liberliber.it', dal quale vogliamo che ci venga inviato il file in formato HTML il cui nome è 'index.html'.
Mutando le sigle è possibile fare riferimento anche ad altri tipi di servizi di rete Internet:

- 'ftp' per i server FTP
- 'gopher' per i server gopher
- 'telnet' per i server telnet.

Occorre notare che questa sintassi può essere utilizzata sia nelle istruzioni ipertestuali dei file HTML, sia con i comandi che i singoli client, ciascuno a suo modo, mettono a disposizione per raggiungere un particolare server o documento. È bene pertanto che anche il normale utente della rete Internet impari a servirsene correttamente.

Esempi di servizi: la posta elettronica

Il servizio di *posta elettronica* (*e-mail service*) permette la comunicazione asincrona uno-a-uno (cioè un utente con un utente) o uno-a-molti (cioè un utente con molti utenti). Quando un utente desidera inviare una lettera (*e-mail*) deve conoscere l'indirizzo (*e-mail address*) del destinatario. Un indirizzo di posta elettronica di Internet ha la forma: **nome@indirizzo-dominio-di-intenet**, per esempio ciccio@yahoo.it. Ci sono due modi per avere indirizzi:
Il fornitore della connettività TCP/IP (cioè l'Internet Service Provider) può attribuire anche un indirizzo di e-mail. Per accedere e gestire la propria casella di posta occorre un programma specifico (Netscape Messenger, Microsoft Outlook, Eudora, etc…) opportunamente configurato per accedere al server della posta del provider.
È possibile richiedere un indirizzo ad un fornitore di servizi web (es. hotmail.com). In questo caso si accede alla propria casella di posta direttamente tramite il programma di web-browser (Netscape, Explorer, etc…). Quando il sistema di posta è gestito tramite un sito web si parla di servizio di *webmail.*
Un schema semplificato del servizio di e-mail è quello di Errore: sorgente del riferimento non trovata.
Sono diversi i sottosistemi che intervengono per inoltrare un mail:

- il Mail User Agent (MUA) è un'applicazione cliente che interagisce con l'utente nella composizione del messaggio; riempie alcuni campi in modo automatico (indirizzo del mittente, data, Reply-to, riservatezza, priorità, durata, ecc.). I MUA più noti sono i programmi Outlook, Mozilla, Netscape Composer, Eudora, etc…
- il Message Transfer Agent (MTA) provvede al trasporto del messaggio, dialogando via rete con un MTA remoto; più MTA possono essere coinvolti nel routing. L'MTA è un daemon (cioè un programma servente, normalmente denominato sendmail) in

esecuzione su una macchina che viene chiamata server della posta (una macchina il cui indirizzo è fornito dall'Internet Provider al momento della sottoscrizione dell'abbonamento). Il daemon MTA legge il campo indirizzo del destinatario e consegna subito il msg se il destinatario è locale. Altrimenti coinvolge il sistema DNS per convertire l'indirizzo di email del destinatario nell'indirizzo IP del server della posta del destinatario.

SMTP (Simple Mail Transfer Protocol) è il protocollo di Internet per la posta elettronica.
Per l'invio della posta viene seguita questa procedura:
la posta elettronica viene spedita quando la macchina mittente (server della posta del mittente) ha stabilito una connessione TCP con la macchina destinataria (server della posta del destinatario). Su questa macchina c'è in ascolto il daemon MTA che "parla" il protocollo SMTP: accetta le connessioni in arrivo e copia i messaggi nelle caselle postali destinatarie. Se non è possibile spedire un messaggio, al mittente viene restituita una notifica di errore contenente la prima parte del msg non spedito. Dopo aver stabilito la connessione il mittente opera come un cliente e attende la risposta del server destinatario. Il server inizia spedendo una linea di testo che lo identifica e dice se è pronto o no a ricevere la posta. Se non lo è, il cliente rilascia la connessione e riproverà più tardi. Se il server può accettare la posta il cliente annuncia mittente e destinatario del msg. Se il destinatario è noto, il server dice al cliente di proseguire nell'invio. Il cliente quindi invia il messaggio e il server ne dà conferma.
Esiste la possibilità di raggruppare sotto un unico indirizzo, un insieme (lista) di indirizzi corrispondenti ad altrettanti utenti. Quando una persona invierà un messaggio a tale indirizzo esso verrà inviato a tutti gli utenti della lista. In questo caso si dice che la comunicazione avviene tramite *mailing list* (lista di utenti della posta).

Per come è stato progettato questo servizio, non esiste alcun sistema di controllo sull'avvenuto ricevimento dei mail e non esiste la possibilità di garantire dei tempi di consegna. Ciò significa che NON è possibile essere certi che un nostro mail sia stato inoltrato al destinatario e in che tempi. Basta che un server della posta si blocchi e il mail non viene consegnato. Il mittente non viene avvisato di questo evento e quindi non può sapere se deve ritrasmettere il messaggio oppure no. Per ovviare in parte a questo problema, alcuni MUA consentono di attivare un'opzione sul mail che sta per essere inviato che farà aprire una finestra di dialogo al ricevente per chiedergli se voglia o meno confermare l'avvenuto inoltro (viene chiamata spesso *ricevuta di consegna*). L'utente finale può anche decidere di non confermare (è una sua libera scelta) vanificando così la volontà del mittente, oppure per gli stessi problemi di affidabilità già spiegati, può succedere che la conferma NON raggiunga mai il mittente. In poche parole: il sistema di posta elettronica è inaffidabile.

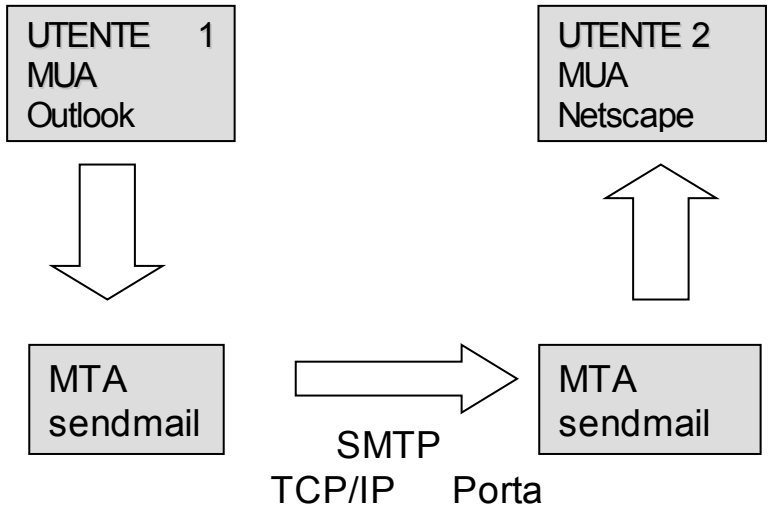

Figura 38: Trasferimento di un messaggio di posta elettronica

Per la lettura della posta esistono due metodi:

- *off-line*: in molti casi quando si lavora su un PC con connessione modem non si vuole gestire la posta "on-line", per risparmiare soldi e non tenere attiva una connessione telefonica quando non serve. Un protocollo usato per recuperare la posta da una casella postale remota è il *POP3* (*Post Office Protocol* – RFC 1225). Ha comandi per permettere all'utente di connettersi, disconnettersi, recuperare messaggi e cancellarli. Lo scopo di POP3 è recuperare la posta dalla casella remota e memorizzarla nella macchina locale dell'utente per leggerla e gestirla fuori linea (off-line), cioè senza il collegamento ad Internet attivo.
- *on-line*: un protocollo più sofisticato di POP3 è *IMAP* (*Interactive Mail Access Protocol* -RFC 1064), che è utile per chi vuole gestire la posta da computer diversi (esempi: azienda, utente che viaggia frequentemente) In questo caso il mail server conserva un deposito centrale accessibile da qualsiasi macchina cliente. A differenza di POP3, IMAP non copia la posta sulla macchina personale dell'utente perché questi può usarne parecchie:la gestione è quindi on-line (più costosa), ma più sicura (non vengono lasciate copie in giro delle proprie e-mail).

Esempi di servizi: i newsgroup

Un *newsgroup* è una bacheca elettronica dove è possibile inserire un proprio mail che verrà poi letto da tutti i fruitori del servizio. I newsgroup sono divisi in canali tematici e sono un'utile fonte di informazioni.
Ad esempio, se un utente vuole avere una precisazione su una legge può scrivere a it.diritto. Chi legge quel messaggio affisso in bacheca può decidere di rispondere e fornire delle delucidazioni. Oppure può aprirsi un confronto/dibattito tra più utenti su un argomento (es, sulle ultime elezioni, sulle modalità di preparazione di una ricetta, etc…)
Per accedere al servizio non occorre iscriversi. Alcuni newsgroup possono essere *moderati*, cioè i messaggi, prima di essere inseriti nella bacheca, passano il vaglio di un censore che controlla che siano attinenti all'argomento del newsgroup e che non siano ingiuriosi o illegali.
Ci sono newsgroup per tutti i generi e gusti:

- alt.alien.visitors
- it.hobby.umorismo
- comp.os.ms-windows
- it.comp.hardware.cd
- it.discussioni.sentimenti
- it.tlc.cellulari

solo per fare pochi esempi. La gerarchia .it è gestita da www.news.nic.it
Per accedere ad un newsgroup basta usare il programma MUA della posta (occorre configurare il programma indicando il news server a cui collegarsi; tale informazione è normalmente fornita dal provider presso il quale si è sottoscritto un abbonamento ad Internet). Esistono anche siti Web che tengono attivi dei link ai newsgroup e che consentono la lettura e scrittura di messaggi sui newsgroup.
Oggi i newsgroup tendono ad essere rimpiazzati da servizi web come social network specializzati (vedi in seguito) ma mantengono un pubblico affezionato tra gli esperti dei vari settori.

Esempi di servizi: il WWW (World Wide Web)

World Wide Web (cui ci si riferisce spesso con gli acronimi *WWW* o *W3*) è sicuramente il

servizio di maggiore successo su Internet. Il successo del web è stato tale che attualmente, per la maggior parte degli utenti, essa coincide con la rete stessa.
Sebbene questa convinzione sia tecnicamente scorretta, è indubbio che gran parte dell'esplosione del "fenomeno Internet" a cui si è assistito in questi ultimi anni sia legata proprio alla diffusione di questo servizio e che nel linguaggio comune l'errata equivalenza tra Internet e Web sia comunemente utilizzata.

La storia del World Wide Web inizia nel maggio del 1990, quando Tim Berners Lee, un ricercatore del CERN di Ginevra - il noto centro ricerche di fisica delle particelle - presenta ai dirigenti dei laboratori una relazione intitolata "Information Management: a Proposal". La proposta di Berners Lee ha l'obiettivo di sviluppare un sistema di pubblicazione e reperimento dell'informazione distribuito su rete geografica che tenesse in contatto la comunità internazionale dei fisici. Nell'ottobre di quello stesso anno iniziano le prime sperimentazioni.

Per alcuni anni, comunque, il World Wide Web resta uno strumento per pochi eletti. L'impulso decisivo al suo sviluppo viene solo agli inizi del 1993, dal National Center for Supercomputing Applications (NCSA) dell'Università dell'Illinois. Basandosi sul lavoro del CERN, Marc Andressen (che pochi anni dopo fonderà con Jim Clark la Netscape Communication) ed Eric Bina sviluppano una interfaccia grafica multipiattaforma per l'accesso ai documenti presenti su World Wide Web, il famoso Mosaic, e la distribuiscono gratuitamente a tutta la comunità di utenti della rete. World Wide Web, nella forma in cui oggi lo conosciamo, è il prodotto di questa virtuosa collaborazione a distanza. Con l'introduzione di Mosaic, in breve tempo il Web si impone come il servizio più usato dagli utenti della rete, e inizia ad attrarne di nuovi.

Il successo di World Wide Web ha naturalmente suscitato l'interesse di una enorme quantità di nuovi autori ed editori telematici, interesse che ha determinato dei ritmi di crescita più che esponenziali. Nel 1993 esistevano solo duecento server Web: oggi ce ne sono milioni.

Su World Wide Web è possibile trovare le pagine di centri di ricerca universitari che informano sulle proprie attività e mettono a disposizione in tempo reale pubblicazioni scientifiche con tanto di immagini, grafici, registrazioni; quelle dei grandi enti che gestiscono Internet, con le ultime notizie su protocolli e specifiche di comunicazione, nonché le ultime versioni dei software per l'accesso alla rete o per la gestione di servizi; ma è possibile trovare anche riviste letterarie, gallerie d'arte telematiche, musei virtuali con immagini digitalizzate dei quadri, biblioteche che mettono a disposizione rari manoscritti altrimenti inaccessibili; ed ancora informazioni sull'andamento della situazione meteorologica, con immagini in tempo reale provenienti dai satelliti, fototeche, notizie di borsa aggiornate in tempo reale e integrate da grafici... ma è meglio fermarci qui: ogni giorno nasce una nuova fonte di informazioni, ed ogni enumerazione sarebbe incompleta non appena terminata.

Naturalmente si sono accorte delle potenzialità del Web anche le grandi e piccole imprese: per molti analisti economici Internet è la nuova frontiera del mercato globale. Prima sono arrivate le grandi ditte produttrici di hardware e software, dotate ormai tutte di un proprio sito Web attraverso il quale fornire informazioni ed assistenza sui propri prodotti, annunciare novità, e (cosa assai utile dal punto di vista degli utenti) rendere disponibili aggiornamenti del software. Poi sono arrivate anche pizzerie e negozi di dischi, agenti

immobiliari ed artigiani della ceramica, librerie e cataloghi di alimentazione naturale... si vende via Internet, si acquista (in genere) con carta di credito.

Le caratteristiche che hanno fatto di World Wide Web una vera e propria rivoluzione nel mondo della telematica possono essere riassunte nei seguenti punti:

- la sua diffusione planetaria
- la facilità di utilizzazione delle interfacce
- la sua organizzazione ipertestuale
- la possibilità di trasmettere/ricevere informazioni multimediali integrate con il documento
- le semplicità di gestione per i fornitori di informazione.

Dal punto di vista dell'utente finale Web si presenta come un illimitato universo di documenti multimediali integrati ed interconnessi tramite una rete di collegamenti dinamici. Uno spazio informativo in cui è possibile muoversi facilmente alla ricerca di informazioni, testi, immagini, dati, curiosità, prodotti.
Dal punto di vista dei fornitori di informazione il Web è uno strumento per la diffusione telematica di documenti elettronici multimediali, decisamente semplice da utilizzare, poco costoso e dotato del canale di distribuzione più vasto e ramificato del mondo.

Tra i diversi aspetti innovativi di World Wide Web, come si accennava, i più notevoli sono decisamente l'organizzazione ipertestuale e la possibilità di trasmettere informazioni integralmente multimediali.

Ipertesti e multimedialità

Da diversi anni queste due parole, uscite dal ristretto ambiente specialistico degli informatici, ricorrono sempre più spesso negli ambiti più disparati, dalla pubblicistica specializzata fino alle pagine culturali dei quotidiani. Purtroppo questo non basta ad evitare la generale diffusa confusione che sussiste sul loro significato.
In primo luogo è bene distinguere il concetto di *multimedialità* da quello di *ipertesto*. Mentre il primo si riferisce ai mezzi diversi con cui può avvenire una comunicazione, il secondo riguarda la sfera più complessa della organizzazione dell'informazione.
Un documento si dice multimediale se può contenere più mezzi (media) di comunicazione differenti (video, audio, immagini, etc…).
Una pagina di un libro con delle figure è un documento multimediale (perché oltre allo scritto è presente un altro media, cioè le immagini).
Da notare che spesso, dagli addetti ai lavori, vengono considerati “media” diversi anche differenti stili nell'uso di un medesimo mezzo fisico, es. l'immagine fotografica e il disegno, la poesia e il racconto, ecc.

Anche se multimediale, un libro può essere letto solo sequenzialmente (pagina dopo pagina, dalla prima all'ultima). L'utente deve tenere conto mentalmente dei legami logici che esistono tra le varie parti del libro con il rischio di dimenticare e perdere il significato del tutto (es. ora l'autore della dispensa vi parla di sistema DNS e voi non ricordate più che cosa è, perché la definizione era scritta molte pagine indietro).
Grazie all'avvento dei calcolatori è stato possibile creare dei documenti organizzati in modo completamente differente. *Un ipertesto è un documento che può essere letto in modo non lineare*, cioè il lettore può leggerlo saltando da un'informazione all'altra in modo assolutamente libero. Se progettato bene, un ipertesto consente una chiave di lettura

personale del documento (aiutando la memorizzazione dei legami logici), e apre al lettore la possibilità di scoprire nuovi legami tra le varie informazioni contenute.

Un ipertesto si basa quindi su un'organizzazione reticolare dell'informazione, ed è costituito da un insieme di unità informative (i nodi) e da un insieme di collegamenti (detti nel gergo tecnico *link*) che da un nodo permettono di passare ad uno a più nodi. Se le informazioni che sono collegate tra loro nella rete non sono solo documenti testuali, ma in generale informazioni veicolate da media differenti (testi, immagini, suoni, video), l'ipertesto diventa multimediale, e viene definito *ipermedia*. Una idea intuitiva di cosa sia un ipertesto multimediale può essere ricavata dalla figura seguente.

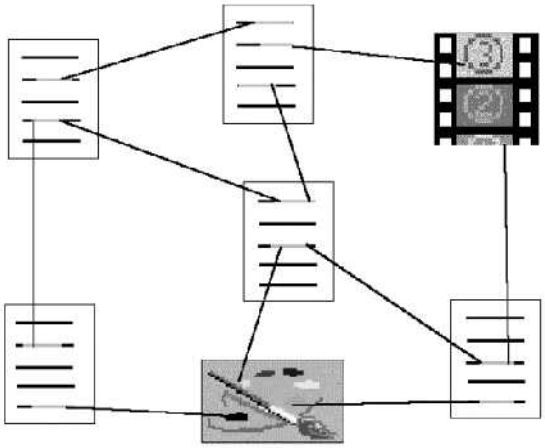

Figura 39: Schema logico di ipertesto multimediale

I documenti, l'immagine e il filmato sono i nodi dell'ipertesto, mentre le linee rappresentano i collegamenti (*link*) tra i vari nodi: il documento in alto, ad esempio, contiene tre link, da dove è possibile saltare ad altri documenti o alla sequenza video. Il lettore, dunque, non è vincolato dalla sequenza lineare dei contenuti di un certo documento, ma può muoversi da una unità testuale ad un'altra (o ad un blocco di informazioni veicolato da un altro medium) costruendosi ogni volta un proprio percorso di lettura. Naturalmente i vari collegamenti devono essere collocati in punti in cui il riferimento ad altre informazioni sia *semanticamente rilevante*: per un approfondimento, per riferimento tematico, per contiguità analogica. In caso contrario si rischia di rendere inconsistente l'intera base informativa, o di far smarrire il lettore in peregrinazioni prive di senso.
Dal punto di vista della implementazione concreta, un ipertesto digitale si presenta come un documento elettronico in cui alcune porzioni di testo o immagini presenti sullo schermo, evidenziate attraverso artifici grafici (icone, colore, tipo e stile del carattere), rappresentano i diversi collegamenti disponibili nella pagina. Questi funzionano come dei *pulsanti* che attivano il collegamento e consentono di passare, sullo schermo, al documento di destinazione. Il pulsante viene 'premuto' attraverso un dispositivo di input, generalmente il mouse o una combinazioni di tasti, o un tocco su uno schermo *touch-screen*.
L'incontro tra ipertesto, multimedialità e interattività rappresenta dunque la nuova frontiera delle tecnologie comunicative. Il problema della comprensione teorica e del pieno sfruttamento delle enormi potenzialità di tali strumenti, specialmente in campo didattico, pedagogico e divulgativo (così come in quello dell'intrattenimento e del gioco), è naturalmente ancora in gran parte aperto: si tratta di un settore nel quale vi sono state negli ultimi anni - ed è legittimo aspettarsi negli anni a venire - innovazioni di notevole portata.

World Wide Web (WWW) è una di queste innovazioni (non è l'unica esistente, ma certamente la più nota). WWW si tratta infatti di un sistema ipermediale (un ipertesto di ipertesti); con la particolarità che i diversi nodi della rete ipertestuale sono distribuiti sui vari host collegati tra loro tramite la rete Internet. Attivando un singolo link si può dunque passare da un documento ad un altro documento che si trova archiviato fisicamente su un altro computer della rete.

Il funzionamento di World Wide Web non differisce molto da quello delle altre applicazioni Internet. Anche in questo caso il sistema si basa su una interazione tra un client ed un server. Il protocollo di comunicazione che i due moduli utilizzano per interagire si chiama *HyperText Transfer Protocol* (*HTTP*). L'unica - ma importante - differenza specifica è la presenza di un formato speciale in cui debbono essere memorizzati i documenti inseriti su Web, denominato *HyperText Markup Language* (*HTML*).

I client Web sono gli strumenti di interfaccia tra l'utente ed il sistema; le funzioni principali che svolgono sono:

- ricevere i comandi dell'utente
- richiedere ai server i documenti
- interpretare il formato e presentarlo all'utente.

Nel gergo telematico questi programmi vengono chiamati anche *browser*, dall'inglese *to browse*, sfogliare, poiché essi permettono appunto di scorrere i documenti. In occasione del rilascio delle specifiche ufficiali dell'HTML 4.0, il W3C (World Wide Web Consortium, http://www.w3c.org) ha promosso il termine, più generico, *user agent* al posto di browser, per indicare però qualunque programma in grado di connettersi al web e scaricare contenuto via HTTP, senza avere necessariamente le caratteristiche di interfaccia

Nel momento in cui l'utente attiva un collegamento - agendo su un link o specificando esplicitamente l'indirizzo di un documento - il client invia una richiesta ('request') ad un determinato server con l'indicazione del file che deve ricevere.

Il server Web, o più precisamente server HTTP, per contro si occupa della gestione, del reperimento e del recapito dei singoli documenti richiesti dai client. Naturalmente esso è in grado di servire più richieste contemporaneamente. Ma un server può svolgere anche altre funzioni. Una tipica mansione dei server HTTP è l'interazione con altri programmi, interazione che permette di modificare documenti in modo dinamico (ad esempio, la pagina web di un giornale è aggiornata da uno speciale programma che inserisce gli articoli mano a mano, durante la giornata, arrivano le notizie in redazione).

Servizi moderni VOIP, IPTV, ecc.

Che il web non coincida con Internet sta diventando sempre più chiaro dato che sopra i protocolli di Internet si stanno oggi lanciando molti nuovi servizi che sfruttano la banda larga disponibile grazie alle connessioni ADSL e UMTS per riuscire ad ottenere, utilizzando quindi le stesse modalità di distribuzione dei pacchetti di informazione agli indirizzi degli host visti finora, servizi interattivi audio/video come la televisione on demand (IPTV) la telefonia (VOIP, Voice over IP) o varie forme di informazione/intrattenimento (giochi multiutente, sistemi informativi geografici, mappe online, ecc.).
Skype è un client per la telefonia via Internet ormai largamente diffuso e tutti i principali fornitori di servizi ADSL offrono tra i loro servizi anche il broadcasting televisivo.
Si presuppone ormai una prossima convergenza su Internet di tutti i canali di

comunicazione, informazione ed intrattenimento, cosa che secondo alcuni critici potrebbe causare seri problemi alla funzionalità della rete.

Evoluzioni, tendenze: il web2.0

Per quanto riguarda il web, anch'esso, comunque sta evolvendo. Le pagine web sono diventate più dinamiche ed interattive mediante tecnologie “standard” quali l'HTML dinamico (che include programmi “eseguiti” sul client e sono stati aggiunti ai contenuti dei siti web maggiori contenuti multimediali mediante protocolli (spesso purtroppo proprietari) per distribuire contenuti audio e video (streaming), animazioni 2D e 3D, ecc.
Da un punto di vista “sociale” sono emerse nuovi servizi e tendenze dominanti sul web, legati in particolar modo a siti che permettono a utenti non specialisti di inserire contenuti multimediali personalizzati (Web log o “blog” per il testo, “You Tube” per il video) o creare comunità online (social network).
Alcuni siti si avvalgono sistemi di creazione di contenuto collaborativi per permettere a comunità di utenti di creare progetti comuni (emblematico il caso di Wikipedia, un'enciclopedia online creata da centinaia di migliaia di autori indipendenti).
Spesso questi siti o i nuovi servizi in rete vengono largamente pubblicizzati o propagandati dai media e vanno a costituire mode destinate a rapidi ridimensionamenti (come nel caso dei blog o dello stesso Second Life, una sorta di videogioco online più discusso sui giornali che effettivamente utilizzato).
A questo insieme di tendenze ci si riferisce spesso con il termine “Web 2.0”, anche se questo non rispecchia necessariamente un cambio improvviso di tecnologie o protocolli, ma più che altro una serie di mutamenti progressivi che hanno modificato sia la *tecnologia* che l'*uso* dei siti Internet.

7 Glossario

Nel seguito trovate spiegazioni di alcuni termini e sigle usati da chi si occupa di calcolatori, reti, internet ed all'informatica in genere e che stanno ormai passando al linguaggio comune, anche se non sempre in modo appropriato

Accessibilità
"La capacità dei sistemi informatici, nelle forme e nei limiti consentiti dalle conoscenze tecnologiche, di erogare servizi e fornire informazioni fruibili, senza discriminazioni, anche da parte di coloro che a causa di disabilità necessitano di tecnologie assistive o configurazioni particolari" (Legge 4 del 9 gennaio 2004)
ADSL (Asymmetric Digital Subscriber Line)
ADSL è una tecnologia che consente di utilizzare un normale doppino telefonico per la trasmissione dati utilizzando tutta la banda passante disponibile. Le tecnologie xDSL
AGP (Accelerated Graphics Port)
Connessione dedicata per l'inserimento di schede grafiche, consente di connettere la scheda direttamente alla memoria fisica, più volte migliorata, è stata però recentemente sostituito dal PCI express.
AJAX (Asynchronous JavaScript + XML)
Tecnologia Web utile alla realizzazione di applicazioni interattive utilizzando pagine che possono essere modificate parzialmente sulla base di connessioni al server senza essere completamente ricaricate.
Alfanumerico
Si dice di un dato che contenga sia caratteri alfabetici (A/Z) che numerici (0/9).
Algoritmo
Sequenza ordinata di operazioni da far eseguire e che portano alla soluzione desiderata da un insieme di dati di partenza
ALU
Arithmetic and Logic Unit (Unità aritmetico-logica): parte della CPU del calcolatore (integrata nel microprocessore) che si occupa dei calcoli matematici e delle operazioni logiche sui numeri binari.
Analogico
Dato misurato dalla variazione continua di una quantità fisica. Si contrappone a digitale (vedi).
API (Interfaccia di Programmazione di Applicazione)
Insieme di procedure disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per un determinato compito.
ASCII
Standard di codifica binaria dei caratteri sui computer. Si pronuncia "aschi". Il codice Ascii originale utilizza gruppi di 7 bit per codificare caratteri alfanumerici, punteggiatura, simboli grafici. 32 codici vengono utilizzati per controllo periferiche. Esistono versioni estese
Atom
Formato di interscambio di contenuti basato su XML, può essere utilizzato per la sottoscrizione a servizi web (blog, giornali)
Avatar
Rappresentazione digitale tridimensionale dell'utente in ambienti di realtà virtuale
Backup
Copia di sicurezza di un disco rigido, di una parte del disco o di un insieme di file.
Banner

Spazio pubblicitario su una pagina Web

Beta

E' il termine che si usa per indicare lo stadio di prova di un software. Le persone incaricate del collaudo dei programmi si chiamano beta tester

BIOS

Basic Input/Outout System. È l'insieme dei programmi di sistema che presiedono a tutte le funzionalità di base di un computer, come la comunicazione tra la CPU e il disco rigido o lettore floppy e la scheda grafica. Viene richiamato dal microprocessore per inizializzare il pc all'accensione e caricare il sistema operativo. E' memorizzato in una EPROM

Bit

Abbreviazione di binary digit ossia cifra binaria, è l'unità di misura del sistema di numerazione binario.

Bitmap

Letteralmente "mappa dei bit", cioè ricostruzione per punti di un'immagine sullo schermo o da una stampante. È strettamente legata alla risoluzione del monitor o della stampante. Lo sono i disegni realizzati con un software di grafica per punti

Blog

Abbreviazione di web log, indica una sorta di diario online in cui i creatori possono inserire testi e immagini mediante un Content Management System ed interagire in modo standard con i lettori.

Blu-ray Disc

Supporto di memorizzazione ottico di recentissima generazione, sviluppato da Sony e che grazie all'utilizzo di un laser a luce blu, riesce a contenere fino a 54 GB di dati. Sarà il supporto standard per il video ad alta definizione del futuro dato che il progetto rivale HD-DVD è stato abbandonato all'inizio del 2008.

Bluetooth

Specifica industriale per reti personali senza fili, sviluppata da Ericsson e utilizzata per connessioni a corto raggio (Es. pc-palmari, telefoni, fotocamere, giochi)

Boot (bootstrap)

Insieme deii processi che vengono eseguiti dal computer durante la fase di avvio, dall'accensione fino al completato caricamento del sistema operativo.

Bridge

Apparecchiatura per l'interconnessione tra segmenti di rete che agisce a livello dati (data link), riconoscendo l'indirizzo del destinatario e consentendo l'instradamento dello stesso al corretto segmento.

Broadcasting

Trasmissione di un messaggio da un sistema trasmittente ad numero di riceventi non definito a priori.

Browser web

Programma che consente agli utenti di visualizzare e interagire con testi, immagini e altre informazioni contenute in una pagina web di un sito

Buffer

Memoria RAM dedicata ad immagazzinare dati per il successivo utilizzo. Viene utilizzata quando bisogna aspettare l'esecuzione di una periferica più lenta (es: stampante), per immagazzinare i risultati di un comando od una elaborazione, per conservare copia di dati che potrebbero servire di nuovo, oppure per non appesantire la memoria RAM di sistema. Per questo ultimo scopo sono dotate di buffer le schede grafiche, il buffer delle stampanti serve a contenere i caratteri prossimi ad essere stampati, quello del processore serve a lavorare più velocemente

Bus

Connessione che consente ai dati di transitare fra diversi componenti (chip, schede, periferiche...). Può essere di diverse tipologie (PCI, ISA, MCA...) ed i componenti collegati devono essere compatibili. Ciascun bus ha un'ampiezza, in termini di bit, per la trasmissione dei dati attraverso il computer, che rappresenta il numero di bit che possono essere inviati contemporaneamente da un componente all'altro

Byte

Unità di informazione consistente in otto bit. Le memorie sono in genere strutturate in multipli del byte. I multipli del byte indicati con i termini Kilo, Mega, Giga, ecc indicano moltiplicazioni successive per fattori 1024 (la potenza di 2 che approssima più da vicino 1000).

C

Linguaggio di programmazione molto diffuso creato da Brian Kernighan e Dennis Ritchie. E' stato utilizzato per lo sviluppo del sistema operativo Unix. **C++**

Linguaggio di programmazione derivato del C cui aggiunge il paradigma ad oggetti

Cache

Memoria aggiuntiva utilizzata nei moderni calcolatori che ricopia in dispositivi a più rapido accesso informazione contenuta in memorie più “distanti”. La cache del PC ricopia parte del contenuto della memoria principale. Si parla di cache anche per una applicazione di rete (es.web browser), in tal caso ricopia sul PC locale informazioni che stanno sul server per recuperarle senza collegarsi in caso di nuova richiesta.

CAD

Computer Aided Design: tipo di software usato in architettura od in ingegneria per disegnare e creare progetti.

Card

Sinonimo di scheda.

CD-R

Compact Disk (dispositivi di memorizzazione ottica) sui quali è possibile scrivere con un apposito masterizzatore. La registrazione di dati può essere effettuata in una singola sessione, o in più riprese (multisessione).

Cd Rom

Compact Disk-read only memory, CD che si può solo leggere, non scrivere.

CD-RW

CD-ROM sui quali è possibile scrivere con un apposito masterizzatore. La registrazione di dati può essere effettuata in più riprese (multisessione) ed i dati registrati sono modificabili (riscrivibili circa 1000 volte)

CGA

Colour Graphics Adapter: tipo di scheda grafica non più utilizzata. Consentiva la visualizzazione di 320 per 200 pixel in 4 colori oppure di 640 per 200 pixel a 2 colori.

Chiave USB

Memoria di massa portatile di dimensioni contenute collegabile al computer mediante la porta USB. E' ormai il principale strumento per il salvataggio su strumento rimuovibile di file avendo soppiantato i floppy disk.

Chip

Circuito integrato, componente elettronico realizzato su un substrato di materiali semiconduttori e contenente da pochi a decine di milioni di parti elementari come transistor, diodi, condensatori e resistori.

Chipset

Insieme di circuiti integrati della scheda madre che gestiscono lo scambio di dati fra CPU e memoria e con il bus. Ogni chipset è specificamente progettato per una famiglia di processori e un tipo di memoria.

Client
Componente (hardware o software) che accede ai servizi o alle risorse di un altro, detto server. Se la componente è software (es. programma di posta elettronica, browser web) si parla di client software.
Clock
Segnale periodico generato da un oscillatore ed utilizzato per sincronizzare l'attività di un dispositivo elettronico digitale.
CMOS
Complementary MOS: tecnologia di costruzione di chip utilizzando transistor.
CODEC
Programma o un dispositivo che si occupa di codificare e/o decodificare digitalmente un segnale (tipicamente audio o video) perché possa essere salvato su un supporto di memorizzazione o richiamato per la sua lettura.
Codice sorgente
Un programma prima di essere stato compilato. Il codice sorgente può essere letto e interpretato, nonché corretto o modificato, mentre il programma compilato è pressoché incomprensibile (infatti il codice sorgente può contenere dei commenti che non verranno compilati) e immodificabile
CompactFlash
Tipologia di scheda di memoria flash, con capacità variabile da 16 Mb a 64 Gb
Compilazione
Traduzione di un codice sorgente in codice eseguibile effettuata con un programma compilatore.
Compressione
Operazione che riduce le dimensioni di un file per minimizzare il tempo di trasmissione o l'occupazione di memoria.
CMS (Content Management System)
Insieme di programmi che si installa su un server web e permettono la gestione automatizzata di siti web tramite interfacce semplici. Esistono vari tipi di CMS specializzati per siti specifici (es. didattica, commercio, notizie, blog, wiki)
Coprocessore
Processore che, occupandosi di particolari operazioni, solleva il processore (CPU) da funzioni specializzate (es. calcoli matematici complessi, la visualizzazione sul monitor...).
CPU (Central Processing Unit)
L'unità centrale di elaborazione del calcolatore, in genera implementata fisicamente dal microprocessore. Contiene l'Unità aritmetico logica, l'Unità di controllo e registri di memoria.
Database
Archivio di dati digitali, strutturato in modo tale da consentire la gestione dei dati stessi (l'inserimento, la ricerca, la cancellazione ed il loro aggiornamento) da parte di opportune applicazioni
dB (Decibel)
Un'unità di misura utilizzata in elettronica per quantificare l'intensità di un rumore o di un segnale
DBMS (DataBase Management System)
Sistema software per la gestione di database
DDR (Double Data rate)
Particolare tipo di SDRAM (vedi) che usa entrambi i fronti del clock per trasmettere informazioni.
DMA

Direct Memory Access: metodo di comunicazione dei dati fra disco o scheda, che vede i dati arrivare direttamente alla memoria di sistema, senza passare attraverso la CPU. Il chip DMA entra in azione quando una grossa quantità di dati deve transitare dal disco rigido alla memoria del computer, in modo tale da evitare il passaggio attraverso la CPU, che rallenterebbe l'operazione.

DirectX

Collezione di per la computer grafica 2D e 3D, sviluppato da Microsoft, costituisce una delle due interfacce standard per la programmazione degli effetti grafici utilizzando le caratteristiche dell'hardware della scheda grafica, cioè per inviare alla scheda i comandi per la generazione delle immagini.

Disco RAM

Parte della RAM che viene utilizzata e considerata dal computer come se fosse un disco rigido

Disco rigido

Unità di memoria di massa che costituisce in genere il principale archivio non volatile del computer contenente il sistema operativo, il software ed i dati degli utenti.

DivX

Tecnologia per la distribuzione di contenuti viedo compressi, basata su un codec che implementa le specifiche Mpeg-4. Permette di ridurre le dimensioni di un film in DVD in modo da salvarlo su comuni CD mantenendo una buona qualità video. Ha riscosso molto successo per la disponibilità di programmi freeware per la codifica/decodifica, anche se si tratta di software non libero. Esistono soluzioni alternative come 3ivx, OpenDivx, Xvid.

DNS (Domain Name System)

Servizio utilizzato per la risoluzione di nomi di host su internet in indirizzi numerici IP e viceversa. Il servizio è realizzato in maniera distribuita su una gerarchia di server.

DOS

Disk Operating System: sistema operativo memorizzato sulla memoria di massa del calcolatore, cioè insieme di programmi per la gestione della macchina e l'interazione con l'utente.

dpi

Dots Per Inch: misura, espressa in punti per pollice, della risoluzione grafica di una periferica (monitor, stampante, scanner...) o di una immagine. Più la misura è grande, migliore è la resa grafica. Il monitor ha 72 dpi, le stampanti laser hanno un minimo di 300 dpi, quelle ad alta definizione arrivano fino a 1200 dpi

DRAM

Dynamic RAM: tipo di RAM che immagazzina i bit utilizzando la carica di condensatori. LA carica in ciascun condensatore determina se il bit è 1 o 0. Se il condensatore perde la carica, l'informazione è perduta: nel funzionamento la ricarica avviene periodicamente. (operazione di refresh).

DSP

Digital Signal Processor: processore posto sulla scheda audio con funzioni specializzate per la resa dell'audio in tempo reale per l'audio posizionale. Processori DSP vengono usati anche in telefoni cellulari, TV digitali, riproduttori MP3 ed altre apparecchiature audio.

Dual core (doppio nucleo)

Architettura che unisce due processori indipendenti e le rispettive Cache in un singolo package. Questo tipo di architettura consente di aumentare la potenza di calcolo senza aumentare la frequenza di lavoro e il conseguente calore dissipato.

DVD (Digital Versatile Disk)

Supporto di memorizzazione di tipo ottico. Esistono vari formati per la registrazione su DVD: DVD-R, DVD+R (non riscrivibili), DVD-RW, DVD+RW, DVD-RAM riscrivibili. Le

versioni normali consentono di memorizzare 4.7 Gb, esistono però anche i formati dual layer che raddoppiano lo spazio a disposizione.

EDO

Enhanced Data Output: tipo do RAM che permette un accesso più veloce ai dati utilizzando un buffer dati doppio. Riduce di circa il 10% il tempo di accesso rispetto alle DRAM standard.

EEPROM

Elettronically Erasable Programmable Read-Only Memory: ROM di tipo programmabile dall'utente ove le operazioni di scrittura, cancellazione e riscrittura hanno luogo elettricamente. È quindi possibile modificarne il contenuto anche quando già installata sulla scheda e sul personal, senza alcuna necessità di estrazione o di intervento hardware. È presente in molti dispositivi e schede.

EGA

Enhanced Graphics Adapter: interfaccia grafica per monitor, ora in disuso, che consente la visualizzazione di 16 colori con una risoluzione di 320 per 200 pixel o di 640 per 200 pixel.

EIDE

Enhanced IDE: bus per la connessione di periferiche, come hard disk, più veloce del precedente IDE

E-mail

Servizio di Internet che permette di mandare messggi con eventuali allegati in modo asincrono ad altri utenti

Ethernet

Protocollo per le reti locali, concepito da Xerox Corporation. Ethernet utilizza un algoritmo di accesso multiplo alla rete detto CSMA/CD (Carrier Sense Multiple Access with Collision Detection).

FDD

Floppy Disk Drive: il lettore di dischetti floppy.

Fibre ottiche

Filamenti cavi di materiale vetroso utilizzati come guide luminose che consentono di trasferire dati a grandissima velocità. Vengono utilizzate nelle telecomunicazioni e per la fornitura di accesso alla rete con velocità al Tbit/s.

File

Un file è un insieme definito di informazioni e dati conservati e trattati come una sola unità logica.

File Allocation Table: tipo file system utilizzato nelle prime versiondi di Microsoft DOS e Windows,.

File system

La parte del sistema operativo che si occupa della gestione dei file, formattando opportunamente le unità di memoria di massa, registrando e leggendo i file.

Flash ROM

Tipo di ROM, permanente riscrivibile (EEPROM) organizzata a blocchi, ovvero un circuito a semiconduttori sul quale è possibile immagazzinare dati binari mantenendoli anche in assenza di alimentazione.

Forum

Sito internet che realizza una bacheca per l'inserimento di messaggi degli utenti

Free Software Foundation (FSF),

Fondazione creata da Richard Stallman che si occupa di eliminare le restrizioni sulla copia, sulla redistribuzione, sulla comprensione e sulla modifica dei programmi per computer.

Freeware

Software distribuito gratuitamente. Non confondere con il software “libero”

GIF (Graphics Interchange Format**)**

Diffuso formato di codifica dei file contenenti immagini (grafica e fotografie).

GNOME (GNU Network Object Model Environment)

Ambiente desktop sviluppato come software libero per le piattaforme GNU/Linux.

GNU

Progetto di sviluppo di software libero che ambisce a creare un sistema completo sdi programmi per coprire ogni necessità informatica. Fulcro del Progetto GNU è la licenza chiamata GNU General Public License (GNU GPL), che sancisce e protegge le libertà fondamentali dello sviluppo “free”.

Gruppo di continuità

Apparecchiatura che consente un'alimentazione elettrica senza interruzioni. Contiene un accumulatore che si sostituisce immediatamente alla linea elettrica qualora questa venga ad interrompersi.

Hard Disk

Disco rigido, memoria non volatile di tipo magnetico utilizzata come memoria di massa nella maggior parte dei personal computer.

Hardware

Termine che indica tutte le parti “fisiche” del calcolatore: CPU, schede, cavi, dispositivi elettronici, tastiera, mouse, ecc.

HDD

Hard Disk Drive: disco rigido interno al computer

Hz (Hertz)

Misura della frequenza, tipicamente misurata in cicli per Secondo.

Host

È un termine generico usato per indicare un calcolatore connesso in rete (a Internet) che può ospitare come server programmi utilizzati da più utenti.

HTML (Hyper Text Mark-Up Language)

Linguaggio di marcatura usato per descrivere i documenti ipertestuali disponibili nel World Wide Web. Consente di definire sezioni logiche di documento che sono interpretate opportunamente dal Browser

Hub

Dispositivo usato per estendere una rete in modo che possano essere attaccate altre workstation.

IEEE

Sigla di Institute of Electrical and Electronic Engineers, organismo che definisce standard e protocolli usati nell'elettronica e nelle telecomunicazioni.

Indirizzo IP

Numero (codificato da 32 bit per IPv4) che identifica univocamente un dispositivo all'interno di una rete informatica che utilizza lo standard IP (quindi ad esempio su Internet)

Interattività

Caratteristica di un sistema il cui comportamento varia al variare dell'e azioni dell'utente sullo stesso. Un programma in esecuzione su un computer è interattivo se l'utente può compiere delle azioni sui normali strumenti di input che ne possono modificare l'esecuzione.

Interfaccia

Dispositivo fisico o virtuale che permette la comunicazione fra due o più entità di tipo diverso

Interprete

Programma che traduce codice scritto in linguaggi ad alto livello in codice macchina e lo esegue passo passo durante la decodifica.

Ipertesto
Insieme di documenti collegati tra loro in modo non lineare mediante parole chiave o ancore che rimandano da un documento all'altro.

ISA
Industry Standard Architecture: bus utilizzato nei vecchi computer dei tipo PC, XT (bus a 8 bit) ed AT (bus a 16 bit) per il trasferimento dei dati fra scheda madre e schede utilizzate per il collegamento di periferiche.

ISO
Sigla di International Standard Organization, la più importante organizzazione a livello mondiale per la definizione di standard industriali e commerciali.

ISDN (Integrated Service Digital Network)
L'ISDN, rete numerica (digitale) integrata nei servizi, è un tipo di connessione telefonica. L'ISDN comporta l'installazione di una doppia linea telefonica (due numeri telefonici).Con una linea telefonica la velocità massima ottenibile è di 64Kbit/sec, mentre con ISDN si ottiene la velocità di 128Kbit/sec grazie all'utilizzo simultaneo delle due linee.

ISP (Internet Service Provider)
Organizzazione o struttura commerciale che fornisce agli utenti la connessione ad Internet e l'uso dei suoi servizi.

Java
Linguaggio di programmazione orientato agli oggetti derivato dal C++

Javascript
Linguaggio interpretato che viene utilizzato per la programmazione di eventi dinamici lato client sulle pagine web. I browser hanno incluso un interprete javascript ed attraverso le funzioni del linguaggio si può accedere alle varie parti del documento web e del browser attraverso una struttura standard detta Document Object Model). Non ha nulla a che fare con Java.

JPEG
Joint Photographic Expert Group: comitato che ha definito il principale standard per la compressione di immagini fotografiche o pittoriche. I file di immagini compresse JPEG sono identificati in genere dall'estensione .JPG.

Kb
Kilobyte 1 Kb è uguale a 1'024 byte.

Kbps (Migliaia Bit per Secondo)
Una misura della velocità di trasmissione dati.

Kernel
La parte principale del sistema operativo. Viene caricato in memoria subito dopo il BIOS e si occupa del trasferimento dei dati fra le componenti del sistema (disco fisso, RAM, CPU, schede, interfacce...) e della gestione della CPU. Riceve ed inoltra i comandi dell'utente tramite la shell.

KHz (KiloHertz)
Misura di frequenza pari a 1000 cicli al secondo.

LAN (Local Area Network)
Rete di computer limitata ad un'area circoscritta (un ufficio, un edificio).

LCD
Liquid Crystal Display: tipo di tecnologia utilizzata per i monitor dei computer, basato sulle proprietà ottiche di particolari sostanze denominate cristalli liquidi.

LED
Light-Emitting Diode: diodi che emettono luce usati per le spie e gli indicatori luminosi

presenti sulla maggior parte delle apparecchiature elettroniche (modem, monitor, tastiera...) sono led (in genere verdi, per indicare che l'apparecchio funziona, o rossi, per indicare un guasto od un problema).

Linguaggio di programmazione
Linguaggio formale, dotato di una sintassi e di una semantica ben definite, utilizzabile per l'implementazione di algoritmi per una macchina automatica (tipicamente, un computer).

Linux
Sistema operativo libero di tipo Unix (o unix-like) costituito dall'integrazione del kernel Linux con elementi del sistema GNU e di altro software sviluppato e distribuito con licenza GNU GPL o con altre licenze libere.

Malware
Software indesiderato installato malignamente sul PC attraverso la rete o scambi di files e che agisce contro l'interesse dell'utente

MAN (Metropolitan Area Network)
Rete di trasmissione dati che serve un'area che approssimativamente corrisponde a quella di una grande città. Le reti di questo tipo sono realizzate con tecniche innovative come, per esempio, il posizionamento di cavi a fibre ottiche.

Masterizzatore
Dispositivo in grado di creare o duplicare dischi ottici (CD, DVD) contenente dati o audio/video

Mbps (Milioni Bit per Secondo)
Una misura della velocità di trasmissione dati.

Mbyte
Unità di misura della capacità di memorizzazione dei dati.(1024x1024 bytes)

Memoria di massa
Dispositivo di memoria non volatile, che può essere interna al computer od anche essere costituita da dispositivi rimovibili.

MHz (Mega-Hertz)
Abbreviato MHz. Un milione di cicli al secondo. La velocità di clock di un processore spesso è espressa in megahertz.

Microprocessore
Implementazione fisica della CPU in un singolo circuito integrato.

MIPS
Millions of Instructions Per Second: unità di misurazione della velocità di un computer, in milioni di operazioni al secondo. Viene utilizzata nelle prove (benchmark) per confrontare le prestazioni dei diversi modelli.

Mouse
Dispositivo di puntamento basato sulla riproduzione su un cursore dello schermo del movimento piano eseguito sul dispositivo stesso.

MP3
MPEG-4 Audio Layer III: tecnologia per la compressione/decompressione di file audio, che consente di mantenere un'ottima fedeltà e qualità anche riducendo il file audio di un fattore 10.

MPEG
Motion Picture Experts Group: comitato formato nel 1988 da membri ISO e IEC che stabilisce gli standard digitali per audio e video. MPEG-1 (ISO 11172) definisce lo standard usato per audio e video su CD. MPEG-2, definisce lo standard usato ad esempio per televisione digitale, terrestre e via satellite, e DVD. Il successivo MPEG-4, definisce nuovi standard utilizzati per la codifica audio video utilizzati in molti formati per file e streaming audio video su internet.

Multi core
CPU composta da due o più core ovvero da più "cuori" di processori fisici montati sullo stesso package, in modo da aumentare le prestazioni complessive sfruttando il parallelismo. Oggi le architetture dual core o quad core sono disponibili anche su PC di fascia bassa.
MultiMediaCard (**MMC**)
Uno dei primi standard di memory card basati su memoria Flash
Multimedialità
Presenza di diversi mezzi di comunicazione in uno stesso supporto o contesto informativo.
Multiplexer
Nella trasmissione dati è un dispositivo che ad un capo della connessione invia i segnali provenienti da diversi canali di trasmissione in un solo canale ad alta capacità di trasmissione. All'altro capo della connessione un altro multiplexer provvede ad invertire questo processo per rigenerare i singoli canali.
Multiprocessore
Architettura che utilizza più di un processore per il calcolo.
NTFS
File system di Windows NT.
OpenGL (Open Graphics Library)
Specifica che definisce una API (interfaccia per la programmazione) per più linguaggi e per più piattaforme per scrivere applicazioni che producono computer grafica 2D e 3D. Insieme alla rivale DirectX, costituisce una delle due interfacce standard per la programmazione degli effetti grafici utilizzando le caratteristiche dell'hardware della scheda grafica, cioè per inviare alla scheda i comandi per la generazione delle immagini.
Overclock
Aumento della frequenza della CPU o di un dispositivo sincronizzato da clock. Non è procedura consigliata in quanto aumenta la dissipazione di calore e invalide la garanzia del dispositivo.
Paginazione
Sistema di gestione della memoria del calcolatore riservata ai processi, che viene suddivisa ed allocata a blocchi.
Partizione
Suddivisione di un disco rigido (locale o sudi un server) in più unità logiche, in modo tale che appaia all'utilizzatore come due o più dischi distinti.
PCB
Process Control Block: in alcuni sitemi operativi salva lo stato dei processi in memoria
PCI
Peripheral Component Interconnect: bus per il trasferimento di dati fra la CPU e schede e periferiche, come la scheda grafica, quella audio, il modem Gli slot PCI sono di colore bianco. Ha rimpiazzeto i precedenti bus ISA, EISA ed MCA.
PCI Express
Evoluzione del bus di espansione PCI che sta anche prendendo il posto della vecchia interfaccia per schede grafiche, l'AGP.
PCM (Pulse Coded Modulation)
Un metodo usato per convertire un segnale analogico in dati digitali privi di rumore che possono essere salvati o elaborati da un computer. PCM usa una frequenza di 4 kHa 8 bit che permette di trasmettere 16 K di dati al secondo. PCM è spesso usato nelle applicazioni multimediali.
PCMCIA
Standard adottato da molti produttori di computer portatili per un connettore universale al

quale si possono accoppiare di ogni sorta: dischi rigidi, memorie e modem.

PDA

Personal Digital Assistant: sono computer moderni di tipo palmare, destinati ad usi specifici.

Pen drive

vedi Chiave USB

Pentium

Processore Intel a 32 bit introdotto negli anni novanta. Il nome pentium è ancora in uso per alcuni modelli di processore recenti, seppure molto differenti.

PHP

Linguaggio di scripting open source molto utilizzato per realizzare sistemi di creazione lato server di pagine web dinamiche

Pixel

Il singolo "puntino" che compone un'immagine sul monitor. Il numero dei pixel determina la risoluziione dello schermo, più il numero è alto e più l'immagine sarà ben definita e realistica.

Pipeline

Metodo che consente l'elaborazione di nuovi dati senza necessità di attendere che i dati precedentemente inviati termino la loro elaborazione. Con questo metodo i processori centrali, CPU, o quelli della scheda grafica, possono ricevere dati mentre ne elaborano altri, migliorando l'efficienza del sistema.

PNG

Portable Network Graphic: formato per immagini, utilizzabile sulle pagine web. Consente la compressione dell'immagine sia lossless ce lossy, l'uso di 16 milioni di colori, del canale Alfa e la registrazione dei valori Gamma.

Porta

Punto fisico (hardware) o logico (software), sul quale avviene una connessione, cioè il canale attraverso cui vengono trasferiti i dati da un dispositivo esterno e il PC. Si parla di porta anche per quanto riguarda le reti: il termine serve ad identificare il numero relativo ad una singola tra le molte possibili connessioni che un host può realizzare. In TCP una connessione è identificata oltre che dagli indirizzi degli host di sorgente e ricevitore, anche dalle “porte” attive che trasmettono e ricevono i dati a cui sono associati programmi in esecuzione.

Posta elettronica

Vedi E-mail .

PowerPC

Processore RISC a 32 bit prodotto da Motorola con la collaborazione di Apple ed IBM.

Processo

Programma in esecuzione su un calcolatore

Processore

Circuito integrato che svolge i compiti di calcolo fondamentali del computer

PS/2

Interfaccia di tastiere e mouse per il collegamento al computer. Usa connettori MiniDIN. È un’eredità dei computer PS/2 IBM, ultimamente rimpiazzata da USB.

QWERTY

Identificatore del tipo di tastiera (americana) costruito con le prime quattro lettere che appaiono nella prima fila in altro di tasti.

RAID

Redundant Array of Independent Disks: tecnologia che prevede l'uso di molti hard disk, visti dai computer in rete come uno solo, per consentire una gestione sicura dei dati, che

vengono duplicati in modo da consentire la continuità d'uso degli stessi anche in caso di guasti singoli o multipli

RAM

Random Access Memory: la memoria principale di lavoro del computer. La RAM è volatile quando si spegne il computer, perde tutto il suo contenuto. Esistono tipi diversi di moduli RAM, come SIMM (FPM ed EDO) e DIMM (SDRAM, DDR e SLDRAM).

Registro

“Casella” di memoria della CPU può contenere informazioni sul funzionamento, comandi o dati, sui quali vengono fatti i calcoli.

Repeater

Apparecchiatura utilizzata per interconnettere due reti mediante la ricezione, amplificazione e ritrasmissione dei segnali tra le reti interessate.

RIMM

Rambus In-line Memory Module: tipo di modulo RAM con Direct RDRAM.

RISC

Reduced Instruction Set Computer: generazione di processori che, avendo un numero ridotto di istruzioni da eseguire, funzionano più velocemente e con maggiore efficienza. Le istruzioni più complesse vengono eseguite componendo insieme alcune istruzioni semplici.

Risoluzione

Capacità, del monitor o di una stampante, di riprodurre immagini o testi. Viene misurata in punti per pollice, DPI.

RJ-11/RJ-45

Sile che identificano diffusi connettori telefonici. RJ-11 è un connettore con 4 o 6 pin usato per la maggior parte delle connessioni destinate all´uso vocale. RJ-45 è un connettore a 8 pin usato per la trasmissione dei dati su un cavo a doppino intrecciato telefonico.

ROM

Read Only Memory: memoria di sola lettura. Contiene le istruzioni fondamentali per il funzionamento del computer. I chip ROM mantengono le istruzioni ed i dati anche a computer spento.

Router Apparecchiatura per l'interconnessione tra segmenti di rete che agisce a livello di rete del modello ISO/OSI

RSS (Really Simple Syndication)

Formato per la distribuzione di contenuti Web è basato su XML. Notifica gli aggiornamenti di un sito in formato standard, in modo che un lettore RSS possa presentare in maniera omogenea notizie provenienti dalle fonti più diverse.

Scheda

Circuito stampato che si inserisce in uno slot della scheda madre o in connettori specifici e che è adibito a funzioni specializzate (es. collegamento in rete). La scheda può contenere processori, e dispositivi di memoria (ROM, RAM, ecc,)

Scheda madre

La parte principale del computer, ovverosia la scheda a circuiti stampati che comprende la CPU, la ROM, i coprocessori e la RAM oppure gli slot per inserire schede, coprocessori o moduli di memoria.

Script

Tipologia di programma per computer caratterizzato dalla semplicità e dall'uso di linguaggi interpretati

SCSI

Small Computer System Interface: standard per il collegamento di periferiche (dischi rigidi, lettori CD-ROM, scanner) ad un computer per un veloce scambio di dati, sviluppato dalla

Shugart Associates negli anni '70. Permette la connessione di più periferiche (fino a 7) ad un solo canale ed il loro funzionamento contemporaneo.

SDRAM

Synchronous DRAM: tipo di RAM utilizzata nelle DIMM per la memoria principale dei personal di tipo Pentium e successivi. Un segnale di clock temporizza e sincronizza le operazioni di scambio di dati con il processore, raggiungendo una velocità almeno tre volte maggiore delle SIMM con EDO RAM.

SDRAM-II

Tipo di RAM presente sui moduli DIMM che costituisce una evoluzione delle SDRAM. Chiamata anche DDR.

Seriale

Metodo di comunicazione per cui i singoli elementi di un'informazione sono trasmessi un bit alla volta.

Server

Componente informatica che fornisce servizi ad altre componenti (tipicamente chiamate client) attraverso una rete. Si parla di server sia per l'hardware che per il software.

Secure Digital (SD)

Il più diffuso formato di meory card basate su memorue Flash, che consentono una capacità fino a 16 Gb. Versioni miniaturizzate per telefonini e dispositivi portatili prendono il nome di miniSD e microSD.

Shareware

Software che si può provare gratuitamente e che in genere ha licenza d'uso limitata ad alcuni giorni dopo i quali viene proposto l'acquisto.

Shell

Programma che permette agli utenti di comunicare con il sistema (mediante interprete di linea si comando)e di avviare i programmi. È una delle componenti base di un sistema operativo

SIMD

Single Instruction Multiple Data: capacità di un processore di elaborare più dati con una singola istruzione. È possibile grazie alla presenza di più unità di calcolo nel processore.

SIMM

Single In-line Memory Module: moduli, piccole schede, rettangolari di memoria RAM per le memoria principale del computer. Rimpiazzati dai moduli DIMM.

Sincrona

Trasmissione di dati nella quale le informazioni di temporizzazione per lo scambio dei dati sono comprese nei dati stessi.

Sistema operativo

Insieme di programmi che gestiscono l'attività del calcolatore, i processi in esecuzione, la memoria e le periferiche, consentendo anche l'interazione da parte dell'utente.
Sono sistemi operativi MS/DOS, Unix, Windows '9x, Mac/OS, Linux, ecc.

Sito web

Insieme di pagine web o sistema di generazione dinamica di pagine web creato per consentire all'utente lo svolgimento di un determinato compito

Slot

Connettore presente sulla scheda madre del computer per inserire schede specializzate (come schede di rete, schede grafiche, modem su scheda) o di memoria RAM aggiuntiva.

Smart Card

Tipo di carta magnetica che utilizza un chip integrato nella carta stessa. Può registrare dati di qualsiasi tipo (nomi, indirizzi, numeri di telefono, importi, codici, password) che, con l'inserimento in un lettore, può rendere disponibili

Smart Media
Tipologia di scheda di memoria Flash.
SMP
Symmetrical MultiProcessing: architettura che consente l'utilizzo di più di un processore sullo stesso computer. Utilizzata su server o stazioni grafiche.
Software
Programma o insieme di programmi per elaboratore elettronico
Software libero (free software)
Software rilasciato con licenza che permetta a chiunque di utilizzarlo e ne incoraggi lo studio, le modifiche e la redistribuzione
Spyware
Tipo di malware che registra ed invia in rete informazioni sul comportamento online dell'utente
SVGA
Super Virtual Graphics Array: standard più diffuso per i monitor. Consente la visualizzazione di 256 colori con una risoluzione di 1'024 per 768 punti.
Swapping
Procedura di trasferimento del contenuto di parte della memoria RAM sul disco rigido (che verrà ricaricata quando richiesto) necessaria se la memoria utilizzata dai processi è maggiore della memoria fisica. Implica un notevole rallentamento nelle prestazioni del calcolatore a causa dell'elevato tempo di accesso al disco rigido.
T1
Un circuito punto-a-punto a lunga distanza, che fornisce ventiquattro canali a 64 kilobit al secondo (Kbit/sec), offrendo un´ampiezza totale di banda di 1,544 megabit al secondo.
T3
Un servizio di comunicazioni punto-a-punto, a lunga distanza, che fornisce fino a ventotto canali T1.T3 può trasportare 672 canali a 64 kilobit al secondo per un´ampiezza totale di banda di 44,736 megabit al secondo ed è di solito disponibile su cavi in fibra ottica.
Terabyte
Terabyte: 1 Tb è uguale a 1'024 Gb.
TIFF
Tagged Image File Format: formato grafico per le immagini, in genere fotografiche, a colori. Utilizza la compressione LZW. I file sono identificati dall'estensione .TIF.
Transistor
Il transistor è un interruttore che viene comandato mediante un segnale elettrico invece della più familiare azione meccanica. In pratica, ai due terminali di uscita del transistor, viene applicato il segnale che si vuole trasmettere. Ad un terzo terminale (gate) si applica un segnale che abilita i poli di uscita del transistor. Il segnale in trasmissione è disponibile solo quando il segnale al gate abilita la porta d'uscita.
Touchpad
Dispositivo che sostituisce il mouse nei computer portatile e notebook. Consente di spostare il puntatore tramite lo spostamento e la pressione del dito su una superficie rettangolare sensibile
Twisted pair
Tipo di cavi per reti Ethernet 10Base-T.
TWAIN
Technology Without An Important Name: tipo di driver che consente, per gli scanner compatibili, l'esecuzione di riprese senza utilizzo di software specializzato, ma all'interno delle applicazioni grafiche.
U-SCSI/W-SCSI

UltraSCSI, WideSCSI: standard per il collegamento di periferiche al computer per un veloce scambio di dati, evoluzioni dello SCSI.

ULSI

Ultra Large Scale Integration: tipologia di chip con oltre 1'000'000 di transistor.

UPS

Uninterruptible Power Supply: gruppo di continuità. Serve a fornire una alimentazione elettrica continua e costante, filtrando sbalzi, cali, sovratensioni, brevi interruzioni, ed intervenendo in caso di mancanza di energia elettrica ad alimentare il computer per il tempo necessario a d arrestare in modo sicuro il sistema.

URL (Uniform Resource Locator)

Sequenza di caratteri che identifica univocamente l'indirizzo di una risorsa in Internet

Usabilità

L'efficacia, l'efficienza e la soddisfazione con le quali determinati utenti raggiungono determinati obiettivi in determinati contesti (definizione standard ISO). I progettisti di programmi e sistemi interattivi dovrebbero tener conto delle norme a supporto dell'usabilità per lo sviluppo dei loro sistemi.

USB

Universal Serial Bus: interfaccia per periferiche di tipo digitale, come telecamere, tastiere, mouse, scanner. Consente la trasmissione dei dati a velocità elevata.

UTP

Unshielded Twisted Pair: Il "doppino", cioè un cavo formato da due soli fili, senza schermatura, utilizzato nelle connessioni telefoniche o di rete.

UUENCODE UUDECODE

Procedimento per trasformare un file binario in ASCII e viceversa. Sono necessari per trasferire file non ASCII tramite posta elettronica.

Vettore

Gruppo di numeri che rappresentano i dati di posizione dei vertici, colore ed altri parametri per i triangoli che rappresentano le immagini di oggetti in 3D.

VGA

Virtual Graphics Array: standard grafico per i monitor che consente la visualizzazione di 256 colori con una risoluzione di 640 per 480 punti.

Virus informatico

Codice malware che modifica un file eseguibile causando un effetto indesiderato ed in grado di autoreplicarsi

VLSI

Very-Large Scale Integration: tipologia di chip che contengono fra 100'000 e 1'000'000 transistor.

VPN (Virtual Private Network)

È un sistema per la comunicazione sicura e cifrata tra due o più siti collegati a Internet che permette di realizzare una rete privata protetta sull'infrastruttura pubblica.

WAN (Wide Area Network)

Una WAN, letteralmente "Rete di Area Estesa", spesso è indicata col termine di area geografica; la WAN è una rete che connette degli elaboratori su grandi distanza, dell. ordine dei chilometri. Le prestazioni sono generalmente inferiori di qualche ordine di grandezza rispetto alle LAN, dalle decine di Kbit/s a qualche unità di Mbit/s.

Watt

Unità di misura di potenza che indica il consumo di un'apparecchiatura elettrica.

Wi-Fi (Wireless Fidelity)

Termine che indica dispositivi che possono collegarsi a reti locali senza fili (WLAN) basate sulle specifiche IEEE 802.11.

Wiki
Sito web che viene generato collaborativamente dagli utenti mediante l'uso di opportuni sistemi di gestione dei contenuti (content management systems). Il termine Wiki viene spesso usato anche per indicare i particolari CMS usati per gestire questi siti. Il più noto esempio è l'enciclopedia online Wikipedia
WiMAX (Worldwide Interoperability for Microwave Access)
Tecnologia di recente diffusione perrealizzare reti di telecomunicazioni a banda larga e senza fili. Dovrebbe garantire copertura di distanze molto ampie.
WIMP (Windows Mouse Icons Pointer)
Il paradigma di interazione tipico con le interfacce a finestra dei moderni sistemi operativi, basato sul movimento del puntatore sulle finestre e sui menu interattivi.
WLAN (wireless local area network)
Rete locale senza fili
World Wide Web
Servizio di Internet che permette di pubblicare contenuti multimediali (testi, immagini, audio, video, ipertesti) e servizi accessibili mediante opportuni programmi client (browser)
Worm
Tipo di malware consistente in un programma a sé stante in grado di creare potenziale danno al sistema e di autoreplicarsi
WORM
Write Once Read Many: tipo di disco ottico che consente di registrare i dati anche in diverse sessioni, ma ogni file rimane definitivamente archiviato e non modificabile. Un raggio laser debole serve per leggere i dati, mentre un raggio laser più potente viene utilizzato per scrivere.
XGA
eXtended Graphics Array: interfaccia grafica per monitor che consente la visualizzazione di 256 colori con una risoluzione di 1024 per 678 pixel.
XML
Metalinguaggio utilizzato per creare nuovi linguaggi, atti a descrivere documenti strutturati. L'attuale XHTML che descrive le pagine web è realizzato con XML