

# Elementi di Informatica

ultimo aggiornamento 20 settembre 2011

versione 1.2

I contenuti sono rilasciati sotto la  
Licenza Creative Commons  
[Attribuzione - Non commerciale 2.5](https://creativecommons.org/licenses/by-nc/2.5/)

## Indice generale

|  |    |
|--|----|
| 1 L'informazione.....  | 4  |
| 1.1 Codifica dell'informazione.....                                | 4  |
| 1.1.1 Codifica dei numeri.....                                     | 4  |
| 1.1.1.1 Conversione da binario a decimale.....                     | 4  |
| 1.1.1.2 Conversione da decimale a binario.....                     | 4  |
| 1.1.1.3 Complemento a due.....                                     | 4  |
| 1.1.1.2 Virgola Mobile (o floating point).....                     | 5  |
| 1.1.2 Codifica dei testi.....                                      | 5  |
| Codice ASCII.....  | 5  |
| Codice UNICODE.....  | 6  |
| Formattazione.....   | 6  |
| 1.1.3 Conversione analogico/digitale e digitale analogico.....     | 6  |
| Trasduttore.....   | 6  |
| Campionamento e legge di Nyquist.....                              | 6  |
| Quantizzazione.....  | 7  |
| Modulazione.....   | 7  |
| Modem fonici, ISDN e ADSL.....                                     | 7  |
| Multiplazione.....   | 7  |
| 1.1.4 Sistemi di compressione.....                                 | 8  |
| Compressione senza perdite (loseless).....                         | 8  |
| Compressione con perdita (lossy).....                              | 8  |
| Codifica di compressione Run-length.....                           | 8  |
| Codifica di Huffman.....   | 8  |
| 1.1.5 Codifica della voce e dei suoni.....                         | 8  |
| Differenze tra la codifica della voce e quella della musica.....   | 8  |
| Algoritmo MP3.....   | 9  |
| Codifica MIDI.....   | 9  |
| 1.1.6 Codifica delle immagini.....                                 | 9  |
| 1.1.6.1 Formati grafici per le immagini fisse.....                 | 9  |
| 1.1.6.1.1 I formati più utilizzati per le immagini raster.....     | 9  |
| 1.1.6.2 Formati grafici per le immagini in movimento.....          | 10 |
| 1.1.6.3 Rappresentazione dei colori.....                           | 10 |
| 1.2 Misurare l'informazione.....                                   | 11 |
| 2 Il software.....   | 13 |
| 2.1 Architettura.....  | 13 |
| 2.2 Linguaggi di programmazione.....                               | 13 |
| 2.2.1 Traduttori.....  | 13 |
| 2.2.1.1 Interpreti e Compilatori.....                              | 13 |
| 2.2.2 Linguaggi algoritmici, di interrogazione e di marcatura..... | 13 |
| 2.2.3 GUI languages.....   | 14 |
| 2.3 Sistema operativo.....   | 14 |
| Funzioni di un sistema operativo.....                              | 14 |
| Sistemi monotask e multitask.....                                  | 14 |
| Lo scheduler nei sistemi multitask.....                            | 14 |
| Formattazione fisica e logica.....                                 | 15 |

|   |    |
|---|----|
| Aree di sistema: BOOT, FAT, Direttorio.....   | 15 |
| Avvio del Sistema Operativo.....  | 15 |
| 3 Architettura dell'elaboratore.....  | 16 |
| 3.1 I tipi di elaboratore.....  | 16 |
| 3.2 Struttura dell'elaboratore.....   | 16 |
| 3.2.1 CPU.....  | 16 |
| 3.2.2 MEMORIA.....  | 17 |
| 3.2.2.1 Memoria centrale.....   | 17 |
| 3.2.2.1.1 RAM (Random Access Memory).....   | 17 |
| 3.2.2.1.2 ROM (Read Only Memory).....   | 18 |
| 3.2.2.2 Memoria video.....  | 18 |
| 3.2.2.3 Cache.....  | 18 |
| 3.2.2.4 Memoria di massa.....   | 18 |
| 3.2.2.4.1 Dischi magnetici e zoned bit recording.....                                 | 18 |
| 3.2.2.4.2 Dischi ottici.....  | 19 |
| 3.2.2.4.3 Hard-disk e architettura RAID.....  | 19 |
| 3.2.3 I/O e Bus.....  | 20 |
| 3.2.3.1 Interni.....  | 20 |
| 3.2.3.2 Esterni.....  | 21 |
| 3.3 Architettura del personal computer.....   | 21 |
| Scheda madre (motherboard o mainboard).....   | 21 |
| Clock di sistema.....   | 21 |
| Chipset.....  | 21 |
| 5 Reti di Calcolatori.....  | 22 |
| 5.1 Tipologia di collegamento.....  | 22 |
| 5.2 Reti personali PAN (Personal Area Network).....                                   | 22 |
| 5.3 Reti locali LAN (Local Area Network).....   | 22 |
| 5.3.1 Ethernet.....   | 22 |
| 5.3.2 WLAN.....   | 23 |
| 5.4 Reti geografiche WAN (Wide Area Network) e/o MAN (Metropolitan Area Network)..... | 23 |
| 5.4.1 Rete Internet e Reti Intranet.....  | 23 |
| 5.4.1.1 Introduzione.....   | 23 |
| 5.4.1.2 VPN (Virtual Private Network).....  | 24 |
| 5.4.1.3 Indirizzo IP.....   | 24 |
| 5.4.1.4 Indirizzo URL.....  | 25 |
| 5.4.1.5 Pacchetti dati.....   | 25 |
| 5.4.1.6 Protocolli.....   | 25 |
| 5.4.3 Rete client-server ed il P2P.....   | 26 |
| 5.4.4 Rete di accesso e Rete di transito.....   | 26 |
| 7 Sicurezza.....  | 27 |
| 7.1 Le minacce.....   | 27 |
| 7.2 La sicurezza dei dati.....  | 27 |
| 7.2.1 Crittografia.....   | 27 |
| 7.3 Protocolli di sicurezza.....  | 28 |

# 1 L'informazione

## 1.1 Codifica dell'informazione

La scelta di utilizzare nei calcolatori il linguaggio binario è stata dettata dal fatto che la tecnologia utilizzata per la loro costruzione è basata sull'uso di transistor, i quali sono componenti elettronici assimilabili a degli interruttori che possono assumere due stati fondamentali: acceso o spento. Al primo stato (accesso), vi è il passaggio di corrente elettrica, ed è a questo stato che è stato associato il valore 1, al secondo stato (spento), non vi è invece il passaggio della corrente elettrica e, di conseguenza, vi è stato associato il valore 0.

### 1.1.1 Codifica dei numeri

- dati N bit si hanno  $2^N$  combinazioni binarie differenti
- per codificare K oggetti occorrono N bit  $\rightarrow N = \text{int}(\log_2 K)$

**Metodo per calcolare il logaritmo in base 2 con una comune calcolatrice**

$\log_x y = \log_k y / \log_k x$  (dove k è una qualunque base)

Esempio:  $\log_2 3 = \log_e 3 / \log_e 2$  (si può scegliere indifferentemente anche la base 10)

- sistema di numerazione posizionale in base r: a ciascuna cifra è associato un peso che dipende solo dalla posizione della cifra nella sequenza

#### 1.1.1.1 Conversione da binario a decimale

$$\begin{array}{c}
 1 \ 0 \ 1 \ 1 \\
 1 * 2^3 + 0 * 2^2 + 1 * 2^1 + 1 * 2^0 \\
 8 + 0 + 2 + 1 = 11_{(10)}
 \end{array}$$

#### 1.1.1.2 Conversione da decimale a binario

$$\begin{array}{l}
 25:2=12 \text{ (resto 1)} \rightarrow \text{bit meno significativo} \\
 12:2=6 \text{ (resto 0)} \\
 6:2=3 \text{ (resto 0)} \\
 3:2=1 \text{ (resto 1)} \\
 1:2=0 \text{ (resto 1)} \rightarrow \text{bit più significativo} \\
 25_{(10)} = 11001_{(2)}
 \end{array}$$

#### 1.1.1.3 Complemento a due

Il metodo del “complemento a due” viene utilizzato nei calcolatori per rappresentare i numeri negativi (ovvero i numeri interi relativi). Questo metodo permette di ottimizzare le risorse potendo fare a meno sia dei circuiti di riconoscimento del segno, sia dei circuiti di sottrazione, in quanto si

ha necessità soltanto di quelli di addizione. Inoltre, si ottiene un altro vantaggio, che è quello che per lo 0 (zero) si ha un'unica rappresentazione e non due come nella “Rappresentazione in modulo e segno”<sup>1</sup> dove lo zero assume i valori -0 e +0.

Per la procedura di conversione possiamo sfruttare questo metodo:

Rappresentazione di un numero decimale negativo (p.e. -6) in un numero binario con N bit (p.e. N=8):

1. si converte in binario il suo valore assoluto:  $6 \rightarrow 00000110$
2. si complementa a uno invertendo sostanzialmente i bit:  $11111001$
3. si somma 1:  $11111010$

### 1.1.1.2 Virgola Mobile (o floating point)

Nel mondo dei calcolatori, per rappresentare i numeri con virgola, si adotta un sistema noto come “floating point” che evita sostanzialmente di dover trovare un modo per specificare dove si trova la virgola. In pratica ogni numero (N) viene suddiviso in tre elementi: segno (+/-), mantissa (M) ed esponente (E):  $N = +/- M * 2^E$ .

Il segno verrà codificato dal primo bit più a sinistra (0=positivo, 1=negativo), seguiranno i bit per la codifica della mantissa e poi quelli per l'esponente. La base sarà ovviamente 2 e non vi è quindi bisogno di specificarla.

Tra i vantaggi principali vi è quello di poter rappresentare dei valori molto grandi o molto piccoli con un numero di bit relativamente basso; tra gli svantaggi, invece, vi è quello dell'approssimazione dovuta alla limitazione del numero dei bit per rappresentare la mantissa, e quello della possibilità di incorrere nel problema di *overflow* (ovvero: i bit assegnati non sono sufficienti a rappresentare il risultato di un certo calcolo), a tal proposito vi sono nei calcolatori dei circuiti dedicati al rilevamento di questo problema in modo tale che venga segnalato all'operatore che sta eseguendo un determinato programma o al programma stesso.

### 1.1.2 Codifica dei testi

Come per i numeri, anche per i caratteri è possibile utilizzare la notazione binaria. Lo si fa semplicemente associando ad ogni singolo carattere una ben determinata sequenza di bit. Naturalmente, al fine di poter scambiare l'informazione tra più utenti e/o macchine, è necessario che queste associazioni siano condivise attraverso degli opportuni standard. Tra i più noti, possiamo citare: il Codice ASCII e il Codice UNICODE.

#### Codice ASCII

Il Codice ASCII è uno standard internazionale utilizzato per la codifica testuale. Il Codice è basato su sequenze di 8 bit (un byte) a cui si associano univocamente singoli caratteri che possono essere: lettere dell'alfabeto (minuscole e maiuscole), segni di interpunzione, caratteri numerici e altri caratteri speciali. L'utilizzo del byte per la codifica dei caratteri, fa sì che attraverso il Codice ASCII sia possibile descrivere  $2^8 = 256$  diversi caratteri.

---

<sup>1</sup> Questo è il modo più semplice per rappresentare e distinguere numeri positivi e negativi: al numero binario vero e proprio viene anteposto un bit che, per convenzione, assume il valore 0 se il numero è positivo ed assume il valore 1 se il numero è negativo. Il difetto è quello di avere due modi per scrivere il numero zero. (fonte Wikipedia)

## **Codice UNICODE**

Questo codice, utilizzando 2 byte per la codifica, permette di avere  $2^{16}$  (più di 65000) diverse combinazioni; ciò ha permesso di poter codificare anche i caratteri di altri alfabeti oltre quello occidentale, come per esempio i caratteri ideografici cinesi o l'alfabeto cirillico.

## **Formattazione**

La codifica ASCII o UNICODE permette la rappresentazione dei caratteri, ma non contiene quelle informazioni utili alla formattazione dei testi (corsivo, grassetto, sottolineato, allineamento, tipo di carattere, dimensione, ecc.). La formattazione è una tecnica specifica, legata al programma di videoscrittura che si utilizza, che non fa altro che aggiungere le informazioni più sopra elencate attraverso l'aggiunta di ulteriori bit a quelli dei codici ASCII e UNICODE. Tra i programmi di videoscrittura più utilizzati possiamo citare il Word di Microsoft e il Write di OpenOffice. La limitazione di questi programmi è dovuta al fatto che il formato dei loro file non rappresenta uno standard riconosciuto universalmente. Tali limitazioni vengono superate dal formato PDF (Portable Document Format) che oggi rappresenta un vero e proprio standard, il PDF permette di formattare i testi in maniera indipendente dal sistema operativo utilizzato e, più in generale, dalle diverse caratteristiche software e hardware dei vari sistemi. Ciò permette quindi la libera circolazione e fruizione dei documenti creati.

## **1.1.3 Conversione analogico/digitale e digitale analogico**

### **Trasduttore**

I fenomeni fisici quali per esempio il suono, la luce, la temperatura, ecc., non possono essere “letti” direttamente dai nostri calcolatori senza l'ausilio di particolari attrezzature/circuiti quali i “trasduttori”. Questi dispositivi non fanno altro che tradurre le variazioni del fenomeno fisico in variazioni elettriche (tensione, corrente, potenza) che possono a loro volta essere elaborate dai calcolatori.

### **Campionamento e legge di Nyquist**

Il passaggio da valori analogici a valori digitali avviene attraverso la tecnica del campionamento. In pratica si fanno delle misure della grandezza fisica che si vuole codificare ad intervalli di tempo regolari e predeterminati; a questi valori vengono quindi associati dei numeri binari (quantizzazione). La sequenza di questi numeri (campioni) rappresenterà l'andamento del fenomeno fisico che si sta analizzando.

Si è detto che i vari campioni vengono presi ad intervalli di tempo predeterminati. Esiste una teoria (legge di Nyquist) che definisce il valore ideale per questi intervalli di tempo: tale valore viene definito “frequenza di campionamento” ed è pari al doppio della Banda del segnale, dove per Banda s'intende l'insieme delle frequenze che caratterizzano il segnale e per frequenza s'intende il numero degli eventi caratterizzanti il fenomeno che in una data unità di tempo si ripetono identici o quasi identici.

## Quantizzazione

Nella quantizzazione i valori vengono approssimati. La precisione della misura dipende dal numero di bit che utilizziamo. Il numero di bit ideale si basa sul rapporto segnale/rumore.

## Modulazione

Gli studi nel campo delle telecomunicazioni hanno mostrato che i segnali di tipo sinusoidale decadono (perdono potenza) in maniera minore rispetto ad altri tipi di segnali. Il segnale sinusoidale è stato quindi scelto come “portante” (segnale principale che trasporta l'informazione) nelle trasmissioni wireless. Con il termine **modulazione** si indica la strategia utilizzata per codificare un segnale digitale binario (composto quindi da soli 0 e 1), in un segnale analogico trasportato a sua volta da una frequenza sinusoidale portante. La strategia consiste nel far corrispondere agli stati binari diversi valori di differenti grandezze analogiche quali:

- la potenza del segnale: avremo quindi la modulazione digitale di ampiezza (ASK);
- la frequenza del segnale: avremo quindi la modulazione digitale di frequenza (FSK);
- la fase del segnale: avremo quindi la modulazione digitale di fase (PSK).

## Modem fonici, ISDN e ADSL

I **modem** sono gli apparati che si occupano della modulazione e della demodulazione del segnale. Il nome stesso di questo apparato sta per MODulation and DEModulation.

Occorre innanzitutto considerare che il modem fonico è stato concepito per trasportare il segnale vocale che è abbastanza limitato come banda di frequenze udibili, inoltre i collegamenti sono realizzati attraverso fili in rame (doppino telefonico). Le velocità che si ottenevano con una normale linea telefonica erano dell'ordine di 56 Kbps che è un valore assolutamente non idoneo per la trasmissione di notevoli quantità di dati. Si è quindi passati alla linea **ISDN** dove la trasmissione delle informazioni avveniva soltanto in digitale senza la necessità quindi della conversione in analogico; nonostante questo le velocità di trasmissione rimanevano ugualmente basse e quindi si è fatto un ulteriore passo arrivando all'attuale tecnologia ADSL.

L'**ADSL** (Asymmetric Digital Subscriber Line) è un sistema che sfrutta al massimo le potenzialità del doppino telefonico e che permette di avere due canali separati: uno di discesa dalla centrale telefonica verso il Router-modem Adsl (downstream) e l'altro che fa il percorso inverso (upstream). Il termine asimmetrico sta ad indicare che il sistema favorisce il canale downstream rispetto a quello in upstream in quanto normalmente sono più i dati che vengono ricevuti dall'utente finale che quelli che si ha necessità di trasmettere.

## Multiplicazione

Con il termine **multiplicazione**, invece, si intende la strategia utilizzata per gestire più connessioni contemporanee evitando i problemi di interferenza. Tra le tecniche più diffuse abbiamo:

- FDM (Frequency Division Multiplexing): dove ad ogni utente viene assegnata una frequenza di trasmissione e una di ricezione;
- TDM (Time Division Multiplexing): dove ogni utente “parla” per un determinato periodo di tempo;
- WDM o DWDM: utilizzate nel campo delle fibre ottiche e che permettono di attribuire a ciascuna componente ottica un canale di comunicazione.

## 1.1.4 Sistemi di compressione

### ***Compressione senza perdite (loseless)***

Nella compressione senza perdite, pur ottenendo dei file con una dimensione significativamente minore rispetto a quella dei file originali, non si ha nessuna perdita di dati. Per questo motivo questa tecnologia viene utilizzata quando si ha a che fare con dei testi, con dei programmi o con dei database dove ogni minima informazione riveste un carattere significativo e assolutamente necessario.

### ***Compressione con perdita (lossy)***

Nella compressione con perdita, oltre alla riduzione dei file, si ha la perdita di una parte dell'informazione. Naturalmente le strategie utilizzate fanno sì che l'informazione persa sia la meno significativa possibile. Il fatto comunque che si perda l'informazione porta a limitare l'uso di questa tecnologia a quegli ambiti dove è ammissibile appunto una perdita di informazione, per esempio: quando si ha a che fare con dei file audio, oppure con delle immagini, o magari dei video.

### ***Codifica di compressione Run-length***

Tra le varie codifiche di compressione dell'informazione vi è la codifica Run-length che si basa sul principio che dove vi è una ripetizione di valori, è possibile sostituirla con il simbolo del valore ripetuto e il numero delle ripetizioni.

### ***Codifica di Huffman***

Questa codifica si basa sul principio che in un qualsiasi contesto informativo vi sono dei simboli che si ripetono con una frequenza maggiore: per esempio in un testo alcune vocali come la "a" e la "e" sono molto più frequenti delle consonanti "x", "y" o "z". La codifica assegnerà così dei codici con un numero minore di bit ai caratteri con più alta frequenza, mentre assegnerà dei codici formati da più bit per i caratteri meno frequenti ottenendo, di conseguenza, una riduzione della dimensione complessiva dei file. La codifica di Huffman viene spesso utilizzata insieme ad altre codifiche come la ZIP o RAR.

## 1.1.5 Codifica della voce e dei suoni

### ***Differenze tra la codifica della voce e quella della musica***

La differenza fondamentale risiede nel fatto che nelle comunicazioni telefoniche ciò che si deve trasmettere è fondamentalmente la voce umana la cui Banda frequenziale si estende da 300 a 4000 Hz; mentre in un CD audio, dovendo trasmettere della musica, si ha la necessità di trasmettere una Banda frequenziale significativamente maggiore: circa tre volte rispetto alla sola voce. Inoltre, volendo riprodurre l'effetto stereo, le informazioni da trasmettere si raddoppiano ulteriormente. Ciò comporta l'adozione di codifiche di compressione più sofisticate per il trasporto della musica.

## **Algoritmo MP3**

L'MP3 fa parte della famiglia degli algoritmi di compressione con perdita. Nonostante questo la qualità di riproduzione dei file audio trattati con questo algoritmo è più che soddisfacente per un gradevole ascolto. Per capire l'efficacia dell'MP3 basta pensare che può arrivare a comprimere un file audio originale fino a circa dieci volte.

Questa codifica si basa sulle caratteristiche della percezione uditiva umana: gli studi hanno mostrato come l'orecchio in realtà non percepisce alcune frequenze nelle seguenti condizioni:

1. quando vi sono frequenze che ricadono fuori dallo spettro audio udibile;
2. quando l'ampiezza del segnale non è sufficiente;
3. quando vi è un mascheramento di alcune frequenze causato da frequenze adiacenti con ampiezza molto maggiore.

Pertanto l'algoritmo elimina tutti questi segnali, fundamentalmente non necessari, senza che per chi ascolta sia possibile apprezzare un degrado qualitativo del segnale in uscita.

A questa tecnica, inoltre, vengono affiancate anche altre tecniche come quella di Huffman.

## **Codifica MIDI**

Nella codifica MIDI i suoni vengono rappresentati come sequenze di note (si pensi allo spartito musicale). Vi è quindi una vera e propria rappresentazione simbolica dell'audio molto efficace ed efficiente in termini di memoria necessaria.

## **1.1.6 Codifica delle immagini**

### **1.1.6.1 Formati grafici per le immagini fisse**

Il termine inglese **raster** può essere tradotto in italiano con la parola “trama” o anche “griglia”. Questa tecnologia suddivide ogni immagine appunto in una griglia, ed ogni cella/elemento di questa griglia, il pixel (picture element), viene descritto fundamentalmente attraverso due informazioni: il colore e l'intensità di questo. Per questo motivo le immagini raster vengono anche chiamate bitmap, e sono definite pixel oriented.

Quando la nostra immagine raster deve essere visualizzata in un monitor, occorre fare un'ulteriore approssimazione in quanto anche i monitor sono fundamentalmente delle griglie, ma la loro “densità”, o per meglio dire, la loro definizione, può non corrispondere alla definizione della nostra immagine. Questa non corrispondenza la ritroviamo anche nella definizione degli elementi base che compongono l'immagine: per le immagini raster parliamo di pixel, per i monitor parliamo di dot (punto).

#### **1.1.6.1.1 I formati più utilizzati per le immagini raster**

Tra i formati più utilizzati per le immagini raster possiamo distinguere tra formati con perdita e senza perdita di informazione.

Senza perdita:

- **BMP** (BitMaP): è un formato privo di compressione, ciò significa che all'interno del file vi è una corrispondenza biunivoca tra pixel e sequenza di bit. Ottimo se si ha necessità di alta definizione, ma occupa molte risorse in termini di memoria e non è, di conseguenza, adatto all'utilizzo all'interno di reti di computer.

- **GIF** (Graphic Interchange Format): permette la codifica di soli 256 colori (8 bit per pixel). Tale limite permette comunque un notevole risparmio di memoria e questo lo rende particolarmente adatto per un utilizzo all'interno delle reti di computer (Internet/Intranet).
- **PNG** (Portable Network Graphics ): simile al GIF ma più “evoluto”: permette l'uso della trasparenza e può trattare molti più colori.

Con perdita:

- **TIFF** (Tagged Image Format File): è un formato con perdita e viene utilizzato soprattutto nel mondo della stampa, dell'editoria ed è anche utilizzato nel mondo degli scanner per l'acquisizione di documenti in formato digitale.
- **JPEG** (Joint Photographic Experts Group): nato nel mondo della fotografia permette di ottenere immagini compresse salvaguardando comunque la qualità. Anche questo formato è molto usato nel mondo Internet in quanto permette di ridurre la dimensione dei file rispetto al formato BMP, ma, allo stesso tempo, permette di codificare più colori rispetto al formato GIF.

### **1.1.6.2 Formati grafici per le immagini in movimento**

Nelle immagini di tipo **vettoriale**, l'immagine non è codificata pixel per pixel, ma invece viene convertita (modellata) in singoli elementi geometrici. Ciò significa che non è necessaria la descrizione completa dell'intera griglia (rappresentazione raster o bitmap), ma è sufficiente codificare ogni singolo elemento in base alle sue coordinate spaziali. In questo modo, oltre ad ottenere un significativo risparmio di spazio in quanto vi è meno informazione da memorizzare, le immagini vettoriali, essendo ottenute fondamentalmente da calcoli geometrici, risultano più veloci da creare e modificare ingrandendole, rimpicciolendole o deformandole senza perdere definizione. Queste caratteristiche le rendono particolarmente utili nel campo della progettazione e nel campo della grafica animata.

Per contro, la conversione dell'immagine originale in singoli elementi geometrici, introduce un'approssimazione che ne degrada sensibilmente la qualità e questo rende la grafica vettoriale non adatta in particolare nel campo della fotografia.

### **1.1.6.3 Rappresentazione dei colori**

Per la rappresentazione dei colori si utilizzano soprattutto due diversi modelli:

- il modello **RGB** (Red, Green, Blue): in sostanza ogni componente colore di base viene rappresentata con un certo numero di bit: possiamo andare dal singolo byte, e in questo caso potremo rappresentare per ogni colore sino a 256 ( $2^8$ ) diverse sfumature, sino oggi ad arrivare a 24 bit, oltre 16 milioni ( $2^{24}$ ) di differenti sfumature. La composizione per addizione (questo modello viene definito, appunto, “additivo”) delle tre componenti colore, infine, ci permette di rappresentare la quasi totalità dello spettro dei colori percepibile dall'occhio umano. La composizione dei tre colori base, alla loro massima intensità, produce il bianco (tutta la luce viene riflessa);
- il modello **CMY** (Cyan, Magenta, Yellow): rispetto all'RGB, questo modello si basa su tre componenti di colore diverse che, mescolate, “filtrano” la luce ottenendo così gli altri colori. La composizione dei tre colori base, alla loro massima intensità, produce il nero o, per meglio dire, un marrone molto scuro (tutta la luce viene assorbita). Le differenze tra i mo-

delli RGB e CMY fanno sì che non vi sia una esatta corrispondenza tra una stessa immagine codificata secondo l'uno o l'altro modello. Siccome dalle sole tre componenti CMY, come si è già detto, non si riesce ad ottenere un nero pieno, si utilizza più spesso un modello chiamato CMYK dove la componente del nero è gestita separatamente, questo, oltre a migliorare la resa estetica delle parti nere (per esempio i testi), permette una maggiore efficienza nella gestione degli inchiostri delle stampanti.

Queste caratteristiche fanno sì che i due modelli vengano utilizzati in contesti differenti a seconda della fonte di illuminazione e quindi di come viene generato il colore: nei monitor, per esempio, la luce viene generata dall'interno attraverso l'uso di emettitori di radiazioni con lunghezze d'onda nello spettro del rosso, del verde e del blu, è quindi la somma di queste componenti che produce i vari colori, tutto questo fa sì che sia indicato utilizzare il modello additivo RGB; nella stampa, al contrario, la luce viene dall'esterno, e viene riflessa dal foglio su cui vengono stesi i pigmenti delle tre componenti ciano, magenta e giallo che fungono da filtri (sottraggono luce), per questo motivo si utilizza il modello sottrattivo CMY.

## 1.2 Misurare l'informazione

Lessico di base:

- **Bit**: entità elementare di rappresentazione (binary digit)
- **Byte**: 8 Bit (entità fondamentale per la rappresentazione dell'informazione). Dal Bit meno significativo 0 al più significativo 7
- **Word**: 4 Byte (usato tipicamente per rappresentare i numeri interi)
- **Double Word**: 8 Byte (usato tipicamente per rappresentare i numeri reali)
- **Algoritmo**: sequenza finita (non ambigua) di passi per il raggiungimento di uno scopo (per risolvere un ben definito problema computazionale)
- **Multimedialità**: Multi = molti, Media = mezzo

I vantaggi offerti dall'informatica nel trattamento dell'informazione:

- **Digitalizzazione** (qualità nel tempo)
- **Compressione** (economia)
- **Trasmissione** (efficienza)

Tre momenti dell'informatica:

- **Elaborazione**
  - MIPS (milioni di istruzioni al secondo)
  - MFLOPS (milioni di istruzioni aritmetiche su numeri reali al secondo)
- **Memorizzazione**
  - RAM
  - Disco rigido
  - CD

Unità di misura (multipli binari):

|       |   |          |                  |
|-------|---|----------|------------------|
| chilo | K | $2^{10}$ | ≈ un migliaio    |
| mega  | M | $2^{20}$ | ≈ un milione     |
| giga  | G | $2^{30}$ | ≈ un miliardo    |
| tera  | T | $2^{40}$ | ≈ mille miliardi |

( $2^{10} = 1024 \approx$  un migliaio)

- **Trasporto**

- Modem
- Fibra ottica
- Internet

Unità di misura (Bit al secondo: bps):

Kbps =  $10^3$  bps

Mbps =  $10^6$  bps

## 2 Il software

### 2.1 Architettura

Possiamo fare una prima suddivisione in:

- **software applicativo:** programmi concepiti per risolvere o permettere di compiere determinate attività (p.e.: scrivere testi, elaborare delle immagini grafiche, gestire database, ecc.). Questi software si “appoggiano” ai software di sistema;
- **software di sistema:** quei programmi che gestiscono l'hardware o le risorse di sistema. Possono essere assimilati a “macchine virtuali” che permettono ai software applicativi di sfruttare le risorse della macchina senza occuparsi dei dettagli tecnico costruttivi. Tra questi software abbiamo il Sistema Operativo, i traduttori, i DBMS che sono sistemi di gestione dei database.

### 2.2 Linguaggi di programmazione

#### 2.2.1 Traduttori

I programmi traduttori permettono di trasformare le istruzioni dei linguaggi di programmazione di “alto livello” contenute in quello che viene denominato file sorgente, e cioè quel file contenente tutte quelle istruzioni facilmente comprensibili all'uomo, in istruzioni scritte in “linguaggio macchina”, ovvero in serie di bit (0/1).

##### 2.2.1.1 Interpreti e Compilatori

I programmi traduttori si suddividono in due tipologie: interpreti e compilatori. Gli interpreti traducono in linguaggio macchina le varie istruzioni, scritte nel file sorgente, che di volta in volta si trovano in esecuzione; mentre i compilatori traducono l'intero file sorgente in linguaggio macchina producendo quello che viene chiamato “file eseguibile” e che ha per estensione “.exe”. La compilazione permette una maggiore velocità di esecuzione dei programmi, e una maggiore protezione del codice in quanto evitano la diffusione dei file sorgenti contenenti le istruzioni nei linguaggi di alto livello.

#### 2.2.2 Linguaggi algoritmici, di interrogazione e di marcatura

Tra i linguaggi di programmazione possiamo identificare tre principali categorie:

1. i linguaggi **algoritmici:** come, p.e., il Java, il C, il C++, il Visual Basic, il Python. Questi linguaggi permettono la realizzazione di una molteplicità di programmi, tra questi possiamo citare quelli ormai di uso comune come quelli di videoscrittura, di elaborazione delle immagini e dei video, e gli stessi Browser;

2. quelli di **interrogazione**: quei programmi che lavorano sui database inserendo dati, estraendoli, modificandoli, cancellandoli o elaborandoli;
3. quelli di **marcatura**: permettono la formattazione delle pagine, sia di quelle utilizzate sul proprio elaboratore, sia di quelle utilizzate attraverso Internet. Tra queste ultime possiamo fare un'ulteriore suddivisione tra pagine statiche e dinamiche.

### **2.2.3 GUI languages**

L'acronimo GUI sta per Graphics User Interface, e raccoglie tutti quei linguaggi che permettono la realizzazione di programmi o pagine che attraverso l'uso di finestre, immagini, e vari componenti multimediali, favoriscono una fruizione dei contenuti facilitata, particolarmente confortevole, e ricca di stimoli.

## **2.3 Sistema operativo**

Il sistema operativo è un software di sistema che si occupa di gestire tutte quelle interazioni tra i programmi e le molteplici componenti hardware e software di un computer.

### **Funzioni di un sistema operativo**

Tra le principali funzioni di un sistema operativo vi sono quelle per la **gestione dei dati** in memoria (memorizzazione, recupero, modifica), la **gestione degli utenti**, quella per l'utilizzo in **multitasking**

Il sistema operativo evita inoltre a chi scrive i programmi di dover conoscere nel dettaglio il funzionamento di ogni possibile dispositivo, processo, o circuito, offrendo una sorta di **interfaccia** virtuale che permette di programmare ad un livello di astrazione molto elevato.

### **Sistemi monotask e multitask**

La maggior parte dei computer, soprattutto quelli di uso personale, sono dotati di un'unica CPU. Tra i compiti di un sistema operativo vi è quello di gestire le risorse tra i vari programmi/utenti. Lo svolgimento di questo compito avviene generalmente seguendo due diverse strategie: la prima viene identificata con il termine monotask, la seconda con multitask. Nei sistemi monotask le risorse sono assegnate singolarmente ad ogni utente/programma per il tempo necessario all'esecuzione completa di un determinato processo. Nei sistemi multitask, invece, le risorse vengono suddivise temporaneamente (time sharing) tra i vari utenti/programmi secondo strategie che mirano a massimizzare l'efficienza dell'intero sistema.

### **Lo schedatore nei sistemi multitask**

Nel contesto di un sistema operativo multitasking, lo schedatore è il sistema che si occupa di distribuire nel tempo le risorse tra più utenti/programmi. Le strategie di assegnazione sono concepite in modo tale da ottimizzare le risorse di calcolo disponibili in base ai compiti da svolgere: vi sono infatti alcuni processi che richiedono più risorse e tempi di esecuzione ristretti, altri che possono invece essere gestiti con minori risorse e senza particolari limitazioni temporali. Per esempio lo schedatore può dare priorità a un programma che gestisce una video chat, che è bene non subisca in-

terruzioni, e far lavorare un antivirus in background assegnandogli di conseguenza una bassa priorità.

### **Formattazione fisica e logica**

Con il termine **formattazione fisica** di un disco si intende la sua strutturazione in settori e tracce che costituiscono le unità elementari di memorizzazione delle informazioni. Tipicamente un hard disk viene formattato all'inizio e poi raramente questa procedura (che elimina comunque tutti i dati memorizzati) viene ripetuta.

Dopo la formattazione fisica, segue la **formattazione logica** che permette al sistema operativo, attraverso il file system di gestire i contenuti (directory e files) della memoria. In sostanza questo tipo di formattazione predispose delle aree di sistema in specifiche porzioni del disco ed è legata strettamente al sistema operativo utilizzato.

### **Aree di sistema: BOOT, FAT, Direttorio**

Vi sono tre aree di sistema disposte tipicamente nelle tracce esterne:

1. **BOOT**: generalmente occupa un solo settore e viene utilizzato dalla CPU per avviare il sistema operativo.
2. **FAT**: sono delle tabelle di allocazione dei file dove è memorizzato per ogni file l'indirizzo del primo settore. I file infatti sono generalmente memorizzati in settori non consecutivi. Questo fatto rallenta la lettura di un file in quanto la testina dell'hard disk deve saltare da un settore all'altro. Per ovviare a tale inconveniente vi è la procedura di deframmentazione che cerca di allineare i vari settori di ogni file.
3. **Direttorio**: è assimilabile ad una tabella dove sono riportate tutte le informazioni dei dati presenti in memoria sotto forma di file: nome del file; estensione, dimensione, data dell'ultima modifica.

### **Avvio del Sistema Operativo**

All'avvio di un computer, dopo una prima fase di stabilizzazione delle tensioni interne, la CPU lancia un piccolo programma denominato BIOS<sup>2</sup> contenuto tipicamente in una memoria non volatile ROM. Questo programma verifica in prima battuta se vi è la presenza di un disco di avvio esterno (p.e. un floppy disk), in caso non vi fosse verifica la presenza nel disco rigido delle aree di sistema BOOT, FAT e Direttorio. Se non vi sono viene generato un segnale di errore, altrimenti il BIOS carica nella memoria centrale RAM il contenuto del BOOT che permette l'avvio del sistema operativo vero e proprio il quale, una volta avviato, si occuperà di configurare tutti quei servizi che ci permetteranno di gestire i device e i vari programmi installati nel computer.

---

2 In informatica, il Basic Input-Output System o BIOS è un insieme di routine software, generalmente scritte su ROM, FLASH o altra memoria non volatile, che **fornisce una serie di funzioni di base per l'accesso all'hardware e alle periferiche integrate nella scheda madre** da parte del sistema operativo e dei programmi. Nei computer IBM compatibili la ROM del BIOS contiene anche il POST, il primo programma che viene eseguito dopo l'accensione. (fonte Wikipedia)

## 3 Architettura dell'elaboratore

### 3.1 I tipi di elaboratore

I calcolatori elettronici sono suddivisibili in due grandi categorie:

- computer di uso generale (**general purpose**): non sono programmati a priori per svolgere compiti limitati e specifici, ma, al contrario, il loro punto di forza è che possono essere utilizzati per far funzionare programmi anche molto diversi tra loro (videoscrittura, elaborazione audio/video, ecc.). Questo tipo di macchine sono sia per uso personale (personal computer) che utilizzate in ambiti aziendali. Inoltre sono suddivisibili in quattro macro categorie:
  - i **personal computer** (ovvero i computer di uso personale) che, soprattutto con l'esplosione del fenomeno Internet, sono sempre più diffusi;
  - le **Workstation**, che hanno una potenza di calcolo significativamente superiore a quella dei personal computer, e che vengono utilizzati per svolgere compiti specifici;
  - i **Server**, anch'essi macchine particolarmente performanti che offrono servizi (si pensi a quelli di posta elettronica) a molti altri computer detti client;
  - i **Mainframe**, che sono una serie di computer estremamente potenti e sofisticati che vengono spesso utilizzati all'interno delle grandi aziende e organizzazioni di vario tipo per gestire e archiviare le informazioni più delicate e strategiche (contabilità, clienti, progetti, ecc.).
- computer dedicati a specifici compiti (**special purpose**): in questa categoria rientrano una vasta tipologia di apparati, tutti quelli concepiti per svolgere, con efficacia ed efficienza, compiti specifici: si va dai comuni lettori MP3, ai cellulari, agli apparati che gestiscono funzioni speciali nelle autovetture (ABS, ecc.), agli apparati che gestiscono i dati di navigazione degli aeroplani.

### 3.2 Struttura dell'elaboratore

Nonostante esistano molteplici modelli, e diverse tipologie di computer, possiamo identificare alcuni blocchi fondamentali che accomunano gli elaboratori elettronici:

- unità di input (tastiere, mouse, lettori digitali, sonde, ecc.);
- unità di output (monitor, casse audio, stampanti, ecc.);
- unità centrale di elaborazione (CPU o microprocessore)
- memoria

#### 3.2.1 CPU

L'unità centrale di elaborazione (CPU o microprocessore), esegue le istruzioni dei programmi, legge i dati dalla memoria, decodifica gli input, elabora le informazioni, scrive in memoria, produce

i relativi output e controlla il flusso dei dati coordinando le attività tra le varie componenti del computer.

All'interno di una CPU possiamo individuare tre blocchi principali:

1. **Unità di controllo** CU (Control Unit) che legge dalla memoria le istruzioni, se occorre legge anche i dati per l'istruzione letta, esegue l'istruzione e memorizza il risultato se c'è, scrivendolo in memoria o in un registro della CPU;
2. **Unità aritmetica** ALU (Arithmetic Logic Unit) che si occupa di eseguire le operazioni logiche e aritmetiche. È quindi una vera e propria unità di calcolo;
3. **Memoria locale** che contiene tutti quei dati su cui sono in corso delle operazioni. Tale memoria ha una capacità molto ridotta rispetto sia alla memoria centrale che, ancor più, rispetto alla memoria di massa, ma il suo scopo non è quello di memorizzare enormi quantità di dati, bensì quello di permettere un accesso e un trattamento veloce soltanto di quelli necessari per svolgere determinate operazioni. In pratica la CPU sposta di volta in volta i dati che le occorrono dalla memoria di massa, o da quella centrale, alla memoria locale organizzata in parti denominate **registri**.

Tra le caratteristiche distintive di una CPU abbiamo:

1. la **frequenza del clock** che regola il numero di operazioni nell'unità di tempo;
2. il **parallelismo dei dati** (sia per i calcoli che per la loro trasmissione lungo il Bus Dati) permette di incrementare l'efficienza dei calcolatori attraverso delle tecniche di calcolo parallelo come quelle definite con il termine **pipeline**<sup>3</sup>, o utilizzando più microprocessori come nel caso dell'integrazione su un unico chip di due microprocessori (**dual core**);
3. l'**indirizzamento** (ovvero la capacità di leggere le informazioni sulla memoria RAM attraverso l'Address Bus).

## 3.2.2 MEMORIA

La memoria registra l'informazione per un determinato periodo di tempo,

### 3.2.2.1 Memoria centrale

Con il termine “centrale” si vuole suggerire l'idea che questo tipo di memoria lavora in strettissimo rapporto funzionale con la CPU garantendo un efficace ed efficiente accesso ai dati. Le prestazioni di un microprocessore, infatti, dipendono molto anche dall'interlavoro con la sua memoria centrale, per questo è fondamentale che la velocità di accesso ai dati e di trasferimento sia molto elevata.

#### 3.2.2.1.1 RAM (Random Access Memory)

Memoria centrale di tipo “volatile” che lavora con alte prestazioni con l'unità di elaborazione. In questo tipo di memoria risiede sia il sistema operativo (una volta che è stato avviato), che i dati dei programmi in esecuzione.

La memoria RAM può essere di due tipi:

- SRAM, estremamente veloce ma molto costosa e limitata in termini di capacità;
- DRAM, più lenta ma meno costosa e con più capacità. Quest'ultima è oggi la più diffusa so-

---

<sup>3</sup> In Informatica la pipeline dati è una tecnologia utilizzata nell'architettura hardware dai microprocessori dei computer per incrementare il throughput, ovvero la quantità di istruzioni eseguite in una data quantità di tempo, parallelizzando i flussi di elaborazione di più istruzioni. (fonte Wikipedia)

prattutto nella categoria dei personal computer.

### 3.2.2.1.2 ROM (Read Only Memory)

Memoria centrale di tipo permanente. In questa memoria risiede il programma di avvio del Sistema Operativo (**boot program**). Viene comunemente indicata anche con il termine BIOS.

### 3.2.2.2 Memoria video

Contiene le informazioni dei singoli pixel che verranno visualizzati nel monitor.

### 3.2.2.3 Cache

La memoria cache è stata introdotta nei calcolatori per migliorare le loro performance. In particolare lo scopo è limitare i rallentamenti che si hanno quando il microprocessore preleva dei dati dalla RAM. In pratica la cache contiene tutti quei dati necessari per l'elaborazione corrente e, attraverso opportuni algoritmi, anche quei dati che molto probabilmente serviranno nell'immediato. Tutto questo fa sì che il microprocessore si trovi per molto del suo tempo a lavorare con questo tipo di memoria (spesso integrata addirittura all'interno dello stesso chip del microprocessore) che ha delle prestazioni notevolmente superiori a quelle di una comune RAM.

### 3.2.2.4 Memoria di massa

Le memorie di massa registrano in maniera permanente anche grandi quantità di dati. Possiamo distinguerle in tre differenti tipologie:

- **magnetiche**: hard-disk, floppy-disk, nastri magnetici;
- **ottiche**: CD o DVD;
- a **semiconduttore** (EEPROM): pendrive o flash memory (note anche con il nome di “chiavette USB”).

I parametri che caratterizzano una memoria di massa sono:

1. **capacità**: misurata in MB o GB;
2. **tempo di accesso**: cioè il tempo necessario per acquisire la prima informazione;
3. **velocità di trasferimento**: con quanta velocità si trasferiscono i dati verso la memoria dell'elaboratore.

#### 3.2.2.4.1 Dischi magnetici e zoned bit recording

Il principio di funzionamento delle memorie magnetiche si basa sulle proprietà del materiale ferromagnetico con cui sono realizzate e che è influenzabile da un campo magnetico esterno.

Possiamo immaginarle suddivise in piccolissime celle che possono essere polarizzate positivamente o negativamente attraverso un apposito sistema (tipicamente una testina). L'informazione binaria è quindi associata al segno della polarizzazione (p.e. 0 polo negativo, 1 polo positivo). Ecco allora che la scrittura non è altro che il processo di magnetizzazione delle varie celle, mentre il processo di lettura è la misura del segno di polarizzazione di queste.

I dischi magnetici sono suddivisi in tracce concentriche a loro volta suddivise in settori di 512 byte. La testina scorre lungo i settori per scrivere/leggere l'informazione. Occorre evidenziare che in questo modo i settori esterni del disco hanno un'area più ampia di quelli interni, di conseguenza, la velocità lineare della testina all'esterno risulta maggiore di quella all'interno. Ciò comporta, inoltre, che lo spazio occupato da un bit all'esterno è maggiore rispetto all'interno.

Questo inconveniente viene superato da una nuova tecnologia denominata **zoned bit recording** che suddivide il disco magnetico in diverse zone facendo in modo che il numero dei settori per ogni traccia non sia più costante ma vari in funzione della zona crescendo nelle zone più esterne. Ciò permette di aumentare la capacità dei supporti magnetici.

#### 3.2.2.4.2 Dischi ottici

I primi supporti ottici, CD audio, sono nati nel mondo della musica, ma visti i loro punti di forza (supporti non fissi, economicità, ridotte dimensioni, capacità di memorizzazione elevata, affidabilità e insensibilità alla interferenze elettromagnetiche), sono poi stati sfruttati nell'informatica per la memorizzazione dei dati generici. Tipicamente possiamo avere CD-ROM con informazioni già caricate dal fornitore, CD-R riscrivibili dall'utente una sola volta o CD-RW riscrivibili più volte.

A differenza dei dischi magnetici, nei dischi ottici non vi sono tracce concentriche ma vi è un'unica struttura a spirale simile a quella dei dischi in vinile suddivisa poi in tanti settori. Il principio di funzionamento si basa sul fatto che il raggio laser che colpisce la superficie del disco può venire più o meno riflesso in base alla configurazione che si è data alla superficie in un determinato punto. In pratica le caratteristiche chimico-fisiche di costruzione permettono nella fase di scrittura di incidere/modificare la superficie creando delle aree più o meno riflettenti. L'informazione binaria è quindi associata al grado di riflessione: per esempio si associa lo 0 a una bassa riflessione e l'1 ad un'alta riflessione. Nei CD riscrivibili è la presenza di più strati differenti a permettere la possibilità di incidere/modificare più volte l'informazione sul disco.

Rispetto al CD, nel DVD le incisioni sul disco sono più piccole e quindi vi è la possibilità di memorizzare più dati: generalmente si arriva a 4,7 GB di memoria, ma attraverso delle tecniche di doppio strato, o sfruttando entrambe le facce, si può raddoppiare la capacità di memorizzazione.

Sul mercato si trovano oggi anche gli HD-DVD che memorizzano ancora più dati; vi è poi anche un'evoluzione dei DVD attraverso la tecnologia BLU-Ray che sfrutta un raggio con una lunghezza d'onda più piccola rispetto a quella dell'infrarosso riuscendo in questa maniera a memorizzare ancora più dati.

#### 3.2.2.4.3 Hard-disk e architettura RAID<sup>4</sup>

Gli elementi che compongono un hard-disk sono:

- un piccolo motore elettrico che permette la rotazione a velocità costante del disco magnetico;
- il disco vero e proprio ricoperto da materiale ferromagnetico;
- una testina di lettura e scrittura che si muove a diverse velocità, guidata dalla CPU, sulle varie tracce;
- una memoria cache inserita nel controller del device per ottimizzare lo scambio dei dati.

Nelle macchine Server che hanno la necessità di immagazzinare una enorme quantità di dati, si

<sup>4</sup> RAID Redundant Array of Independent Disks, ovvero: insieme ridondante di dischi indipendenti. (fonte Wikipedia)

utilizza una particolare architettura per gli hard-disk denominata RAID.

In pratica questa architettura prevede l'utilizzo di più hard-disk che vengono comunque visti dalla CPU come un'unica memoria; sarà poi il controller dedicato a gestire il flusso delle informazioni tra i vari dischi.

Abbiamo diverse tipologie di architetture RAID:

- RAID 0: i dati vengono ripartiti equamente su più dischi senza informazioni di ridondanza o parità;
- RAID 1 (disk mirroring): prevede di avere due dischi e di memorizzare l'informazione in modo duplice su un disco e sull'altro. Così facendo si ha una ridondanza che permette in caso di errore o di guasto ad uno dei dischi di recuperare ugualmente l'informazione;
- RAID 4 e la sua evoluzione RAID 5: in questo caso l'informazione non viene duplicata, i dati sono memorizzati su più dischi, ma vi è un disco dedicato che memorizza la parità (algoritmo di rivelazione di errore) in modo tale che vengano rivelati eventuali errori e si operi una correzione.

### 3.2.3 I/O e Bus

I Bus sono canali di comunicazione ad altissima velocità attraverso i quali viaggiano i dati scambiati tra le varie componenti dei computer. I Bus sono anche suddivisi in base alle loro funzioni principali:

- **Address Bus**, che è il Bus dove “viaggiano” gli indirizzi di allocazione in memoria dei dati su cui dobbiamo svolgere operazioni di lettura/scrittura;
- **Data Bus**, ovvero il Bus dei dati dove viaggiano i dati veri e propri (ovvero le informazioni su cui lavorano i vari programmi);
- **Control Bus** dove viaggiano tutte quelle operazioni di servizio come per esempio se i dati presenti in un determinato momento sul Data Bus sono dati in lettura (p.e. i dati che dalla memoria esterna vanno alla CPU), oppure dati in scrittura (dalla CPU alla memoria o ai device esterni), ma anche le fondamentali informazioni di temporizzazione e sincronizzazione delle varie attività.

All'interno di un elaboratore, possiamo inoltre distinguere i bus in quattro tipologie:

- **Host bus** (o CPU bus o front side bus): è il bus più interno e collega il microprocessore con la ALU, il livello di memoria centrale e con l'eventuale cache di secondo livello.
- **AGP bus**: bus veloce per la connessione verso la scheda video. Trasporta una notevole quantità di informazione dovendo oggi gestire filmati in alta qualità.
- **PCI bus**: bus per le connessioni di rete (LAN).
- **ISA bus**: è più lento degli altri e permette la connessione verso alcune periferiche esterne come i lettori di dischi o le stampanti.

#### 3.2.3.1 Interni

I Bus interni collegano vari blocchi fisici e funzionali del computer come la CPU, la memoria cache, la memoria RAM; inoltre, attraverso dei connettori (detti slot), permettono la connessione di schede esterne (p.e. le schede di rete).

### **3.2.3.2 Esterni**

Al fine di permettere l'interlavoro dell'elaboratore con le varie periferiche esterne, vi è la necessità di un elemento che svolga la funzione di controllare il flusso delle informazioni di input/output. Questa funzione viene svolta da un elemento presente tipicamente già nella piastra madre e che è denominato controller. In realtà vi sono più controller specializzati a seconda del tipo di connessione.

I bus oggi più diffusi che collegano i controller ai vari dispositivi esterni sono di tre tipi:

- USB (Universal Serial Bus): oggi molto diffuso, collega molti tipi di device differenti;
- IEEE 1394 (FireWire): utilizzato specialmente per periferiche multimediali e nel trasferimento di video;
- SCSI e l'EIDE orientati al collegamento di memorie di massa (hard disk esterni).

## **3.3 Architettura del personal computer**

### **Scheda madre (motherboard o mainboard)**

Generalmente sulla piastra madre possiamo trovare i seguenti componenti: il microprocessore (CPU), le schede di memoria centrale (RAM/ROM) e una serie di connettori in cui, p.e., è possibile inserire le schede di rete ed eventuali ampliamenti di memoria o Bus di espansione.

### **Clock di sistema**

Con il termine clock si indica la temporizzazione dei processi che avvengono sia all'interno della CPU che tra i vari elementi dell'elaboratore. È bene evidenziare che per le operazioni di calcolo che avvengono all'interno della CPU si registrano le velocità maggiori: nei modelli più recenti dei personal computer si arriva a valori dell'ordine dei 3GHz.

Naturalmente nelle operazioni che coinvolgono elementi esterni alla CPU, come per esempio l'accesso alla memoria statica o le operazioni di Input/Output, i tempi di elaborazione dei processi sono notevolmente più bassi rispetto a quelli dei processi interni alla CPU.

### **Chipset**

I chipset sono un insieme di chip che permettono l'interlavoro tra la CPU e i vari elementi del calcolatore. Ne esistono di diversi tipi in relazione alla gerarchia dei dispositivi coinvolti. Il termine gerarchia sottintende i livelli sottostanti al microprocessore come quello della memoria o quello dei dispositivi periferici. I chipset si occupano del **sincronismo** del flusso delle informazioni che viaggiano a differenti velocità tra i vari livelli.

## 5 Reti di Calcolatori

### 5.1 Tipologia di collegamento

In base alle esigenze, ma anche in base a considerazioni di ordine economico, le reti di connessione possono essere realizzate attraverso collegamenti fisici (wireline), oppure attraverso collegamenti radio (wireless). Tra i primi rientrano i collegamenti realizzati attraverso i cavi in rame, quelli coassiali o, ultimamente, le fibre ottiche; tra i secondi quelli realizzati attraverso i ponti radio, le stazioni di telefonia mobile, il Wi Max, il Wi Fi, l'Hiperlan, ecc.

In base alla distanza dei collegamenti possiamo distinguere tra diverse reti.

### 5.2 Reti personali PAN (Personal Area Network)

Le reti PAN (Personal Area Network) sono reti locali di piccole dimensioni che permettono collegamenti nell'ordine di pochi metri.

Queste reti possono essere realizzate via cavo (USB, FireWire, ecc.), oppure si possono utilizzare soluzioni wireless. Tra le più note:

- **IrDA**: basata sull'uso di infrarossi;
- **Bluetooth**: spesso utilizzato per collegare device mobili (telefonini, computer palmari, ecc.) tra loro, con altri computer, o anche con stampanti o altri apparati;
- **ZigBee**: che ha il vantaggio, rispetto al Bluetooth, di utilizzare potenze più basse e quindi di essere utilizzato in contesti specifici quali quelli della domotica, in ambiente sanitario, automobilistico, nella sicurezza, nel settore dei giochi intelligenti, ecc.

### 5.3 Reti locali LAN (Local Area Network)

Una rete locale LAN interconnette diversi elaboratori tra loro secondo una struttura client-server, ed inoltre, oltre ai computer, all'interno della rete compaiono anche tutta una serie di macchine quali stampanti, plotter, ecc.

Tipicamente le LAN sono strutturate in base a due diverse strategie:

- **collegamento a Bus** dove i diversi calcolatori sono connessi appunto ad un unico Bus;
- **collegamento ad Anello (Ring)** dove i diversi calcolatori sono collegati luno all'altro in modo tale che l'informazione attraversa tutte le macchine collegate all'Anello.

#### 5.3.1 Ethernet

All'interno delle reti LAN il protocollo di comunicazione più diffuso è quello denominato Ethernet.

Le caratteristiche principali di questo protocollo sono:

- velocità di trasferimento di 10 Mbps;
- estensione massima della rete di circa 2,5 Km;
- dimensione del pacchetto di 512 bit (64 byte) sino ad un massimo di 1,518 byte.

Ogni pacchetto, inoltre, è composto da specifiche aree: 6 byte per l'indirizzo del destinatario, 6 byte per quello del mittente, i dati e, infine, alcuni byte di controllo per la verifica degli errori.

Il protocollo Ethernet è stato concepito inizialmente per una rete LAN di tipo a Bus e, di conseguenza, tutti i computer sono collegati ad un'unica risorsa Bus ed è necessario quindi che, per evitare conflitti, l'invio dei dati avvenga per un solo computer alla volta. La gestione della trasmissione dei dati è basata su una strategia denominata Medium Access Control (MAC). In pratica ogni computer, prima di inviare il proprio pacchetto, si mette in ascolto della rete e, se non vi è in corso una trasmissione di dati, invia il proprio pacchetto, dopodiché si rimette in ascolto e se vi è una collisione la trasmissione si blocca per un certo tempo, altrimenti prosegue con l'invio di un altro pacchetto.

Il limite principale di una LAN strutturata su un unico Bus è che nel caso di un guasto di una scheda di rete anche di una sola macchina, tutta la rete ne subisce gli effetti negativi sino al blocco totale.

Per cercare di risolvere questo inconveniente si è pensato di strutturare le LAN in sottogruppi ottenendo una migliore affidabilità e migliori prestazioni. Tale strategia utilizza degli apparati quali gli Hub, i Bridge o gli Switch che permettono appunto una gestione dinamica dei collegamenti.

### 5.3.2 WLAN

Una delle più recenti innovazioni per le reti LAN è quella che ha visto il passaggio da collegamenti realizzati soltanto con dei cavi (wireline), a collegamenti realizzati anche attraverso le onde radio (wireless). Quest'ultimo tipo di rete viene chiamata Wireless LAN (WLAN).

In sostanza ogni elemento della rete è dotato di un ricetrasmittitore che permette alla macchina di interconnettersi ad un punto di accesso (Access Point) il quale è a sua volta collegato via cavo alla LAN principale. Questo tipo di scelta tecnologica viene utilizzata soprattutto per creare delle sottoreti collegate per esempio alla LAN principale di un'Azienda o per creare dei punti di accesso ad Internet in luoghi pubblici (aeroporti, stazioni, Hotel, ecc.). Oggi, con la diffusione di Internet anche in ambito familiare, tale opzione viene utilizzata per creare delle piccole WLAN domestiche.

## 5.4 Reti geografiche WAN (Wide Area Network) e/o MAN (Metropolitan Area Network)

Quando abbiamo reti di computer che sono molto distanti tra loro. Internet, per esempio, può essere definita una rete WAN pubblica.

### 5.4.1 Rete Internet e Reti Intranet

#### 5.4.1.1 Introduzione

**Internet** è fondamentalmente una rete di computer collegati tra loro per lo scambio e la condivisione di informazioni. Il primo sviluppo lo si è avuto alla fine degli anni '60 per merito del Pentagono che intendeva collegare diverse università statunitensi. A metà degli anni '80 si svilupparono le prime applicazioni pubbliche e dagli anni '90 anche quelle commerciali.

Le reti **Intranet** sono in tutto e per tutto identiche alla rete Internet soltanto che hanno un perimetro circoscritto a una particolare organizzazione o azienda. Vi sono anche dei punti di uscita verso la

rete Internet, ma la maggior parte del traffico si svolge all'interno della rete Intranet, e interessa soprattutto le macchine e gli operatori di una specifica organizzazione.

Mandarsi lettere (email) in tempo reale, chiacchierare (chat) con qualcuno che sta a centinaia di chilometri lontano da noi, condividere testi, immagini, video, con amici e conoscenti senza limiti di spazio e di tempo, tutto questo a fatto sì che si parlasse di “villaggio globale” riferendosi a ciò che accade nel Web tutti i giorni. Tutto questo è stato possibile attraverso la diffusione dell'accesso in banda larga ad Internet e, soprattutto in questi ultimi tempi, con l'esplosione del fenomeno dei vari Social network.

Con la diffusione di Internet vi è stata la definizione di alcuni concetti chiave:

1. ogni macchina collegata in rete deve avere un suo specifico indirizzo (**Indirizzo IP**);
2. le risorse all'interno dei calcolatori devono poter essere raggiunte attraverso una denominazione standard (**URL**);
3. le informazioni (**pacchetti dati**) devono essere scambiate sulla base di ben specifici **protocolli** (p.e. **TCP/IP**).

#### 5.4.1.2 VPN (Virtual Private Network)

Con il termine **rete dedicata** si intende che due apparati sono collegati in maniera diretta tra loro, al contrario, con il termine **rete commutata** si intende invece una rete dove viaggiano le informazioni (pacchetti) di più utenti e dove le macchine sono collegate attraverso reti complesse suddivisibili in:

- reti a commutazione di circuito: come, per esempio, la rete telefonica dove il circuito viene stabilito a priori e viene mantenuto fin quando i due calcolatori si parlano;
- reti a commutazione di pacchetto: come, per esempio, le reti LAN, dove il canale non è pre-stabilito, ma i pacchetti vengono inviati su un'unica infrastruttura sino ad arrivare alla macchina destinataria.

Le VPN (Virtual Private Network) sono sottoreti private dedicate realizzate tipicamente all'interno di Intranet aziendali che offrono maggiori garanzie di sicurezza in quanto le macchine che ne fanno parte non sono “visibili” dal mondo esterno.

#### 5.4.1.3 Indirizzo IP

L'indirizzo IP può essere considerato come il nome identificativo di una determinata macchina e serve, appunto, a identificare univocamente quella macchina all'interno di una rete che ne prevede molte altre. L'indirizzo IP è composto da 32 bit, la prima parte rappresenta il Dominio ed è assegnata dall'ICANN (Internet Corporation for Assigned Names and Numbers), la seconda parte è assegnata dall'amministratore di Dominio (Internet Service Provider).

La seconda parte dell'indirizzo IP, quella assegnata dal Provider, può essere di due tipi:

1. statico: una macchina un indirizzo;
2. dinamico: l'utilizzo di indirizzi dinamici permette di sfruttare al meglio una risorsa comunque limitata quale è quella degli indirizzi IP. La possibilità di assegnare dinamicamente a più utenti degli indirizzi IP nasce dal fatto che non tutti sono collegati contemporaneamente e un'assegnazione statica rappresenterebbe uno spreco.

La diffusione di Internet ha subito reso evidente la necessità di associare dei nomi più facilmente memorizzabili e di per se stessi anche informativi agli indirizzi IP. In pratica per molti indirizzi IP che fanno riferimento a particolari server/siti, si associano dei nomi come per esempio [www.tuttocitta.it](http://www.tuttocitta.it) che identificano meglio la risorsa. La conversione indirizzo IP/nome mnemonico viene fatta da particolari macchine denominate DNS (Domain Name System: sistema dei nomi a dominio) e l'assegnazione del nome di dominio viene gestita da un'autorità preposta che si occupa di definire anche il significato delle varie estensioni: .com (per siti commerciali), .edu (per siti a carattere educativo), ecc.

#### 5.4.1.4 Indirizzo URL

Per accedere alle varie risorse multimediali contenute in una specifica macchina, si utilizza un sistema di indirizzamento denominato URL (Uniform Resource Locator). Tale indirizzo è strutturato in questo modo: 1) la prima parte specifica il protocollo del servizio, per esempio HTTP per le pagine web o FTP per trasferire file; 2) la seconda parte riporta l'indirizzo mnemonico associato all'indirizzo IP; 3) la terza parte definisce la risorsa vera e propria nella classica forma adottata dal file system: directory/file.

#### 5.4.1.5 Pacchetti dati

Un pacchetto dati deve contenere una parte relativa alle informazioni del mittente, una per quella del destinatario, una per le informazioni di controllo e di sicurezza e, naturalmente, la parte relativa alla vera e propria informazione da trasmettere. L'informazione viene suddivisa in pacchetti per migliorare l'efficienza e l'efficacia della trasmissione: smistare piccoli pacchetti dati di più utenti è più efficiente che trasferire l'intera informazione un utente alla volta; inoltre, in caso di problemi sulla rete, risulta più facile recuperare e inviare di nuovo soltanto singoli pacchetti errorati, o andati persi, piuttosto che fare un intero nuovo invio di tutta l'informazione.

#### 5.4.1.6 Protocolli

Esistono molteplici protocolli in funzione del tipo di rete e degli scopi che si intendono ottenere. Quali la sicurezza, la velocità, la capacità, ecc.

In generale, comunque, un protocollo di rete non è altro che un insieme di regole circa le modalità di trasmissione dei dati in un particolare contesto di rete. Regole quali: la configurazione dei pacchetti, l'autenticazione, la rilevazione e la correzione degli errori.

Senza i protocolli non sarebbe di fatto possibile nessun tipo di trasmissione. L'anarchia nel mondo dei computer, come infondo in quello degli uomini, non funziona.

Internet può essere descritta su **quattro livelli di servizi** ciascuno con i propri protocolli (regole per lo scambio delle informazioni) e specifiche funzioni:

1. il **livello più alto** che contiene le informazioni che si vogliono scambiare (pagine web, file, posta elettronica, ecc.) che utilizza come relativi protocolli l'HTTP, l'FTP, l'SMTP, ecc.);
2. il **livello di trasmissione** di queste informazioni che utilizza i protocolli TCP e UDP e che serve per gestire il flusso dei vari pacchetti e le ritrasmissioni in caso di errori;
3. il **livello dei collegamenti** di rete di trasporto che utilizzano per esempio il protocollo IP e che si occupano del trasporto dei bit nella rete;
4. il **livello fisico** di collegamento (wireline/wireless) dove Internet non specifica nessun tipo

di protocollo lasciandone il compito ai progettisti e ai gestori delle reti.

### **5.4.3 Rete client-server ed il P2P**

In una rete client-server vi sono tipicamente delle macchine molto potenti (i server) che offrono varie tipologie di risorse e servizi a tutte le varie macchine della rete (i client, di norma meno potenti) che ne vogliono usufruire. In una rete P2P, invece, ogni computer collegato è, nello stesso momento, sia client che server, nel senso che come client può chiedere per esempio una certa risorsa o un particolare servizio ad un'altra macchina della rete e, contemporaneamente, come server, può offrire una risorsa o un servizio ad un altro computer della rete P2P che ne faccia richiesta. Purtroppo oggi questo tipo di reti sono note soprattutto per il traffico “pirata” di materiale audio e video.

### **5.4.4 Rete di accesso e Rete di transito**

Quando parliamo di una grande rete come Internet, ma anche di una Intranet, è bene suddividerla concettualmente in due parti ben distinte; una rete di accesso, e una rete di trasporto. La rete di accesso può essere vista come il primo collegamento che permette al singolo utente di collegare il proprio computer alla rete: può essere quindi il doppino in rame che arriva a casa, l'Adsl, la chiavetta in tecnologia UMTS, Wi-Fi o Wi-Max, ecc. La rete di trasporto è invece quella che l'utente “non vede”: ovvero tutto ciò che sta dietro a quel primo livello che abbiamo chiamato rete di accesso. È nella rete di trasporto che i nostri dati faranno il percorso più lungo ed è per questo tipo di rete che è stata usata la metafora delle “autostrade dell'informazione”.

## 7 Sicurezza

### 7.1 Le minacce

Le principali minacce per un computer collegato in rete sono:

- **hacking**: in generale rappresenta un accesso non autorizzato ad un computer. Può essere fatto a fini di studio per testare i sistemi di sicurezza, o per fini illegali;
- **virus**: programmi che passano da un computer ad un altro attraverso lo scambio dei file che avviene attraverso supporti di memoria quali dischetti, chiavette, ecc. I virus informatici hanno la specificità di duplicarsi proprio come i virus legati alle malattie;
- **worm** (verme): sono come i virus ma sfruttano la rete per propagarsi. Questo li rende particolarmente pericolosi perché in brevissimo tempo possono “infettare” molti computer;
- **spamming**: invio continuo o massiccio di informazioni non richieste o, comunque, non necessarie. Può essere fatto per superficialità o per mettere in crisi un sistema;
- **spyware**: si tratta di un software capace di raccogliere informazioni riguardanti l'attività online di un utente (siti visitati, acquisti eseguiti in rete, ecc.) senza che sia stato dato esplicitamente il consenso. Tali informazioni vengono poi inviate attraverso la rete Internet a chi potrà utilizzarle per trarne qualche profitto;
- **phishing**: si propaga attraverso la posta elettronica o “siti esca” i quali invitano a fornire informazioni personali o comunque riservate.

In generale possiamo dire che queste minacce sono assimilabili a moderni “cavalli di Troia”: sono programmi nascosti all'utente (inseriti spesso in altri programmi o documenti) che si attivano in un certo determinato momento, o a seguito di alcune azioni compiute dall'utente, e che possono compromettere il sistema danneggiando parti importanti, oppure, cosa forse ancora peggiore, acquisire informazioni personali o dati sensibili quali quelli legati ai conti correnti, le password, ecc.

### 7.2 La sicurezza dei dati

Tra le misure di sicurezza principali rientrano:

- stabilire dei sistemi di autenticazione con password complesse, e prevederne un periodico aggiornamento;
- la definizione per ciascun utente di precise limitazioni nell'uso di certi programmi o nell'accesso a file o dati;
- l'utilizzo di Antivirus sempre aggiornati;
- l'utilizzo di Firewall: software che viene posto tra Internet e la rete aziendale, o il proprio computer, al fine di filtrare tutti i pacchetti lasciando passare solo quelli che, secondo una determinata politica sono ritenuti sicuri (per esempio quando risultano provenire da fonti catalogate sicure).
- l'adozione di una precisa politica di backup che salvi periodicamente i dati più importanti.

#### 7.2.1 Crittografia

La crittografia è sostanzialmente quell'insieme di strumenti che permettono di rendere un canale

di comunicazione sicuro in modo tale che un soggetto terzo non possa decodificare facilmente i dati eventualmente intercettati (“spillati”).

Semplificando possiamo dire che attraverso la crittografia si altera la sequenza dei bit trasmessi secondo una precisa chiave di criptazione, soltanto chi è in possesso della chiave di decriptazione può ricostruire l'informazione originaria.

In questo ambito vi sono oggi due principali strategie:

1. **crittografia simmetrica a chiave segreta**: dove la chiave viene utilizzata sia per passare dai dati originali a quelli crittografati e viceversa. La chiave deve essere segreta e nota solo agli attori/macchine interessati dalla comunicazione.
2. **crittografia asimmetrica a chiave pubblica**: qui le chiavi sono due: con la prima il mittente cripta il messaggio e con la seconda il destinatario lo decripta. Ogni soggetto/macchina ha quindi due chiavi. Una chiave privata e un'altra chiave detta pubblica. La prima è ovviamente segreta, mentre la seconda va distribuita a tutti quelli interessati.

I contesti nei quali la crittografia è più utilizzata sono quelli del commercio elettronico.

### **7.3 Protocolli di sicurezza**

I protocolli di sicurezza sono quelli che permettono ai siti web di garantire degli accessi sicuri.

I protocolli crittografici che permettono nel mondo web di realizzare tutto questo sono una variante dell'Http e sono chiamati Https (Http + SSL Secure Security Layer). Il protocollo si basa su una chiave pubblica e una privata.