

informatica

FILIPPO SAPUPPO

DISPENSA CREATA TRAMITE WIKIPEDIA

Sommario

Informatica.....	5
Aspetti tipici.....	5
Etimologia e significato.....	5
Storia.....	6
Descrizione e ambiti	6
Hardware	7
Software	7
Sistema informatico.....	8
Applicazione	8
Computer.....	8
Etimologia e traduzione del termine	9
Cenni storici	10
Descrizione	10
Hardware	12
Scheda madre	13
Northbridge	13
Southbridge	13
Memoria secondaria.....	13
Supporti rimovibili	13
Dispositivi di Input.....	14
Tipi di computer.....	14
Personal computer.....	17
Utilizzo	17
Tipi	17
Accezione e significato del termine	17
Storia.....	19
Caratteristiche generali.....	20
Componenti	21
Architettura (hardware e software).....	22
Software	22
Storia del software.....	23
Aumento del codice e potenziamento dell'hardware	23
Classificazione.....	24
Versione	25
Licenze d'utilizzo e distribuzione	25
File	25
Descrizione	25
Tipi di file	26

Operazioni sui file	27
Trasferimento di file.....	27
Recupero dei file.....	28
Reti di computer	29
Rete / Mainframe	29
Tipi di reti.....	30
Internet.....	32
Descrizione	32
Struttura fisica	33
Accesso	33
Funzionamento logico	34
Storia di Internet.....	35
Nascita del World Wide Web (1991)	37
Protocollo di rete	39
Indirizzo IP	39
I servizi correlati.....	42
Server.....	42
Descrizione	42
Caratteristiche	43
Funzionamento.....	43
Architettura client-server o N-tier	43
Gestione.....	44
Servizi e tipologie di server.....	44
Client.....	45
Descrizione	45
Transmission Control Protocol	45
Descrizione	45
Caratteristiche principali.....	46
World Wide Web	47
Caratteristiche generali.....	47
Descrizione	48
I contenuti.....	48
I servizi.....	49
Implementazione.....	49
Funzionamento.....	50
Accessibilità e usabilità.....	50
Storia.....	51
Dal web statico ai web service.....	51
Dal web statico al web semantico	52

Tim Berners-Lee.....	53
Biografia.....	53
World Wide Web Consortium.....	54
Storia.....	54
Membri.....	55
Obiettivi del W3C.....	55
Le raccomandazioni.....	55
Browser.....	56
Descrizione.....	56
Storia.....	56
HTML.....	57
Descrizione.....	58
Accesso via Internet.....	59
Elementi sintattici.....	59
Struttura di un documento HTML.....	59
Pagina web.....	60
Descrizione.....	60
Tipologie.....	60
URL.....	61
Grafica.....	61
Creare una pagina web.....	62
Standard per pagine web.....	62
Sito web.....	62
Descrizione.....	63
Accesso alle pagine web.....	63
Tipologie di siti web.....	63
Categorie.....	64
Progettazione e realizzazione.....	65
I linguaggi.....	65
Accessibilità dei siti web.....	66
Usabilità del web.....	66
Motore di ricerca.....	68
Motori di ricerca per il web.....	69
Le fasi.....	69
Catalogazione.....	69
Risposta.....	69
Risultati sponsorizzati.....	70
Raffinazione della ricerca.....	70
Motori di ricerca più utilizzati.....	70

Ipertesto	71
Collegamenti ipertestuali	72
Storia.....	72
Database	73
Gestione delle informazioni.....	74
Struttura o modello logico.....	74
Linguaggi per basi di dati.....	75
Server di database	75

Informatica

L'**informatica** (termine che deriva dal francese *informatique*, contrazione di *informazione automatica*) è la scienza che ha per oggetto lo studio dei fondamenti teorici dell'informazione, della sua computazione a livello logico e delle tecniche pratiche per la loro implementazione e applicazione in sistemi elettronici automatizzati detti quindi *sistemi informatici*.



Una rappresentazione artistica della macchina di Turing

Aspetti tipici

In altri termini è frequentemente descritta come lo studio sistematico dei processi algoritmici che descrivono e trasformano l'informazione ed è quindi punto di incontro di almeno due discipline autonome: il progetto, la realizzazione e lo sviluppo di macchine rappresentatrici ed elaboratrici di numeri (dunque l'elettronica) e i metodi di risoluzione algoritmica di problemi dati (algoritmica), che sfruttino a pieno le capacità di processamento offerte dalle macchine elaboratrici stesse per l'ottenimento di determinati risultati in output a partire da determinati dati in input. La domanda principale che sostiene l'informatica è dunque: "*Come si può automatizzare efficientemente un determinato processo?*"

L'informatica, assieme all'elettronica e alle telecomunicazioni unificate insieme sotto la denominazione Information and Communication Technology, rappresenta quella disciplina e allo stesso tempo quel settore economico che ha dato vita e sviluppo alla terza rivoluzione industriale attraverso quella che è comunemente nota come rivoluzione informatica.

Etimologia e significato

« L'informatica non riguarda i computer più di quanto l'astronomia riguardi i telescopi. »

(Edsger Wybe Dijkstra)

L'etimologia italiana della parola "informatica" proviene dal francese, dalla compressione di *inform(ation électronique ou autom)atique*, e sicuramente Philippe Dreyfus, che per primo utilizza nel 1962 il termine *informatique* (informatica) voleva intendere il trattamento automatico dell'informazione mediante calcolatore (naturale o artificiale). Da notare il fatto che in lingua inglese non esista l'equivalente di "informatica" ma venga utilizzato il termine "computer science" che, ad ogni modo, presuppone l'esistenza della figura dello scienziato e del ricercatore interessato all'approfondimento della conoscenza della tecnologia dell'elaborazione.

È importante anche notare il differente significato di origine tra queste tre lingue nel denominare lo strumento base dell'informatica:

- *elaboratore* (o *calcolatore*), in italiano, per le sue svariate capacità di elaborazione;
- *ordinateur*, in francese, a sottolineare le sue capacità di organizzare i dati (oggi le informazioni);
- *computer*, in inglese, letteralmente calcolatore, in diretta discendenza delle calcolatrici, prima meccaniche, poi elettromeccaniche, poi elettroniche.

È importante tenere a mente che, a dispetto di quanto molto comunemente si ritiene, quella dei computer **non** è affatto classificabile come intelligenza. L'elaboratore non fa altro che eseguire istruzioni preventivamente "impartitegli" da un essere umano. Ciononostante, il computer è diventato insostituibile nei campi più disparati della vita e della scienza.

Una branca specifica dell'informatica, l'intelligenza artificiale (AI), si occupa di creare tecniche, algoritmi e programmi atti a simulare processi di pensiero e ragionamento. Queste tecniche non sono meno algoritmiche e deterministiche nei loro esiti di quelle usate in altri settori dell'informatica, e tuttavia hanno il potenziale di catturare conoscenza e farne uso per fornire risposte spesso di qualità superiore a quella ottenibile mediante l'uso di esperti umani.

Storia

Oggi "informatica" è un termine di uso comune ma dai contorni ben poco definiti.

Come scienza si accompagna, si integra, o è di supporto a tutte le discipline scientifiche e non; come tecnologia pervade pressoché qualunque "mezzo" o "strumento" di utilizzo comune e quotidiano, tanto che (quasi) tutti siamo in qualche modo utenti di servizi informatici.

La *valenza* dell'informatica in termini socio-economici ha scalato in pochi anni la *piramide di Anthony*, passando da *operativa* (in sostituzione o a supporto di compiti semplici e ripetitivi), a *tattica* (a supporto della pianificazione o gestione di breve termine), a *strategica*.

A questo proposito è importante ricordare che l'informatica è diventata talmente strategica nello sviluppo economico e sociale delle popolazioni che il non poterla utilizzare, uno status battezzato con il termine *digital divide*, è un problema di interesse planetario.

Descrizione e ambiti

Informatica pratica e informatica teorica

Esistono frange di persone che confondono l'informatica con aree vocazionali che tipicamente riguardano l'utilizzo di software applicativo e che comprendono il semplice utilizzo di programmi per l'ufficio, il navigare sul web o il gaming. L'informatica invece vede editor di testo, browser e videogame come semplici strumenti di lavoro o svago. Quello che interessa, nell'informatica pura, non è tanto saper usare i cosiddetti applicativi per come essi si presentano, quanto piuttosto capirne, tramite ad esempio l'analisi del sorgente, la struttura ed eventualmente saperla migliorare con l'uso di algoritmi più efficienti sotto diversi criteri (uso di memoria, numero di istruzioni, parallelismo, ...). Nelle università esistono dei dipartimenti e dei corsi di Informatica. Un informatico dovrebbe sempre avere un interesse genuino per i fondamenti teorici dell'informatica; che poi, per professione o per passione, spesso faccia lo sviluppatore di software è possibile ma, potendo sfruttare le proprie capacità di *problem solving* in diversi ambiti, non è scontato. In ogni caso l'informatica, almeno nella sua parte applicativa, è una disciplina fortemente orientata al problem solving.

I fondamenti teorici della disciplina discendono quindi direttamente dalla matematica (matematica discreta), a cui l'informatica è strettamente legata, ed è anche per tale motivo che di recente lo studio dell'informatica ha assunto rilevanza multidisciplinare nel cercare di chiarire o giustificare processi e sistemi complessi del mondo reale, quali ad esempio la capacità del cervello umano di generare pensieri a partire da interazioni molecolari (studi facenti capo alla bioinformatica).

Dallo studio dei linguaggi formali e degli automi provengono quelle conoscenze che permettono la progettazione di linguaggi di programmazione e di compilatori le cui finalità sono quelle di poter fornire gli strumenti essenziali per sviluppare il software di base, come il sistema operativo, e quello applicativo (editor di testo, browser, etc.) e, di conseguenza, di rendere il computer universalmente accessibile alle persone.

Basata sull'informatica è anche l'implementazione logica dei sistemi di telecomunicazioni almeno per ciò che riguarda l'implementazione dei protocolli di rete attraverso software per le telecomunicazioni.

Hardware

Nell'informatica pionieristica degli anni sessanta questo termine inglese, letteralmente "*ferramenta*", ben si prestava a indicare le macchine utilizzate.

Anche con la sostituzione delle valvole termoioniche in favore dei transistor e poi dei primi circuiti integrati MOS, tali macchine erano composte da telai e pannelli metallici robusti tutti rigorosamente assemblati mediante bullonature vistose, per contenere i preziosissimi e delicatissimi circuiti elettronici che erano il cuore degli elaboratori e delle prime periferiche di base.

Oggi, quando risulta difficile ritenere *ferramenta* un mouse o una webcam, il termine è rimasto più che altro per distinguere tutto ciò che è *macchina*, *strumento*, dai programmi (il software) per far funzionare la macchina o lo strumento. In pratica l'hardware è tutto ciò che è palpabile e visibile con gli occhi, come un hard disk, un monitor, lo stesso case del computer.

Software

È importante distinguere il software di base (oggi chiamato sistema operativo) dal software applicativo (comunemente detto programma o applicazione):

- Il software di base serve per rendere operativo l'elaboratore;
- Il software applicativo serve per implementare nuove funzioni e/o rendere operative parti dell'elaboratore.

Una visione semplicistica può essere quella secondo cui l'elaboratore usa:

- il software di base per sé stesso, come servizio interno;
- il software applicativo per fornire un servizio alle persone che lo utilizzano.

Nondimeno, molte funzioni del software di base offrono valore aggiunto anche all'utente finale (ad esempio, il file system consente all'utente di memorizzare e poi riutilizzare secondo necessità il proprio lavoro). Perciò, all'interno del software di base è possibile ulteriormente distinguere le funzionalità a valore aggiunto per l'utente e quelle che sono meramente di servizio per garantire il funzionamento della macchina.¹

Il software di base, oggi detto sistema operativo, nasce con l'evoluzione tecnologica. Mentre i primi elaboratori contenevano tutte le istruzioni necessarie al proprio funzionamento, ivi incluse quelle di interazione con l'hardware (spesso complesse e non portabili), l'aumento della memoria e della potenza di calcolo ha portato all'evoluzione di strati software interposti tra il programmatore e l'hardware (all'epoca, tutti gli utenti erano anche programmatori e non esisteva ancora la nozione di software *general purpose*).

Oggi, se si escludono i sistemi dedicati a particolari servizi critici come l'avionica, la chirurgia assistita, il controllo di impianti a ciclo continuo, mezzi in movimento, ecc. il rapporto tra software di base e software applicativo è sempre superiore a uno.

Nel caso dei personal computer domestici tale rapporto sale spesso tranquillamente al di sopra di 10 a 1.

¹Si potrebbe argomentare che tutte le funzioni del SO offrono valore aggiunto all'utente, ma molte di queste operano in background e sono trasparenti rispetto all'utilizzatore del sistema, mentre altre sono parte integrante della.

Sistema informatico

Hardware e software formano un *sistema informatico*; formano uno strumento *utile per fare qualcosa*.

Questa generica definizione, apparentemente banale, in realtà denota il fatto che oggi con un sistema informatico ci si fa di tutto, e che ogni giorno ci si fa qualcosa di più.

A parte il classico personal computer o il *server* di rete pensiamo ad esempio al *telefono cellulare*, alla *fotocamera digitale*, alla *playstation*, al *cruscotto auto con il navigatore satellitare*, al *monitoraggio in sala di rianimazione*, ecc.

Sono tutti sistemi informatici, che ci forniscono servizi specifici. Pensiamo ad un aereo moderno al suo interno possiamo trovare non uno, ma molti sistemi informatici, ciascuno con un preciso compito.

Internet nel suo insieme è un sistema informatico, formato a sua volta da una rete di sistemi informatici che lavorano per un obiettivo comune: permettere a chiunque di connettersi e scambiare informazioni con chiunque, in qualsiasi parte del globo.

Applicazione

Del sistema informatico si è detto che è uno strumento *utile per fare qualcosa*. Questo qualcosa è *l'applicazione*.

Il termine applicazione informatica è nato quando il computer è uscito dalle mani degli scienziati e dalle stanze degli istituti di ricerca ed è entrato nel resto del mondo. Naturalmente il computer era utile anche prima (e lo è di certo ancor di più oggi, in quegli ambienti), ma come ad esempio in ingegneria e così in informatica, si distingue l'attività della ricerca pura da quella applicata.

Le prime applicazioni pratiche si ebbero, negli anni sessanta e inizio settanta, nelle grandi aziende, e in generale nelle grandi organizzazioni pubbliche o private, laddove soluzioni informatiche abbastanza semplici permettevano significativi risparmi di tempo nelle operazioni quotidiane e di routine. Basti ricordare Banca d'Italia, Alitalia, Eni, Montedison, Enel.

Con gli anni, e con uno sviluppo sempre più veloce delle capacità di elaborazione in parallelo all'abbassamento dei costi, l'informatica ha pervaso qualsiasi settore, fino alla vita quotidiana e all'intrattenimento personale.

Computer



Blue Gene, macchina di grandi dimensioni, con potenze di calcolo dell'ordine del PetaFLOPS, in un centro di ricerca (*High Energy Accelerator Research Organization, KEK*)

Un **computer** (in italiano anche **calcolatore** o **elaboratore**) è una macchina automatizzata in grado di eseguire calcoli matematici complessi e, eventualmente, altri tipi di elaborazioni. Il computer nasce infatti come macchina calcolatrice evoluta e, a partire dalla seconda metà del XX secolo, evolve in macchina in grado di eseguire le elaborazioni più varie restituendo un certo output a partire da istruzioni impartite in input dall'esterno.

Nel corso della storia, l'implementazione tecnologica di questa macchina si è modificata profondamente sia nei meccanismi di funzionamento (meccanici, elettromeccanici ed elettronici), che nelle modalità di rappresentazione dell'informazione (analogica e digitale) che in altre caratteristiche (architettura interna, programmabilità, ecc.). Al giorno d'oggi, ci si riferisce comunemente al computer come ad un dispositivo elettronico e digitale, programmabile a scopo generico costruito secondo la cosiddetta *architettura di von Neumann* ed il modello teorico-computazionale della cosiddetta macchina di Turing. Sebbene i computer programmabili a scopo generico siano oggi i più diffusi esistono in specifici ambiti di applicazione modelli di computer dedicati (automazione industriale, domotica, ecc.).

In questa forma e al pari della televisione, esso rappresenta il mezzo tecnologico simbolo che più ha modificato le abitudini umane dal secondo dopoguerra ad oggi: la sua invenzione ha contribuito alla nascita e allo sviluppo dell'informatica moderna, che ha segnato l'avvento della cosiddetta *terza rivoluzione industriale* e della società dell'informazione.

Etimologia e traduzione del termine

Il termine *computer* è il nome d'agente del verbo inglese *to compute*, derivato dal latino *computare* attraverso il francese *computer*². L'etimo latino è composto da *com* = *cum* (insieme) e *putare* (tagliare, rendere netto - da cui l'odierno *potare*) e significa propriamente: "confrontare per trarre la somma netta"³. In inglese, il termine indicava originariamente un essere umano⁴, incaricato di eseguire dei calcoli. Il primo utilizzo nel senso moderno è attestato nel 1897⁵⁶ ma bisognerà attendere la metà degli anni cinquanta perché questa accezione diventi di uso comune (si notino, a questo proposito, i diversi acronimi dei computer ASCC ed ENIAC).

Come per gran parte della terminologia informatica, l'utilizzo dell'originale inglese prevale sulla traduzione italiana (*calcolatore* o *elaboratore*)⁷, sebbene in altre lingue europee accada diversamente (francese: *ordinateur*, spagnolo: *computadora* o *ordenador*). Tale tendenza è spesso biasimata in una diatriba sull'esterofilia della lingua italiana recente⁸⁹, ma le proposte alternative - come il *computiere*¹⁰ del professor Arrigo Castellani, accademico della Crusca e fondatore degli Studi Linguistici Italiani - non hanno ancora vasta applicazione.

²Online Etymology Dictionary

³Ottorino Pianigiani, *Vocabolario Etimologico*, Polaris, 1993 [0]

⁴J. Copeland, *The Modern History of Computing*

⁵Oxford English Dictionary, s.v. "Computer"

⁶M. Campbell-Kelly e W. Aspray, *Computer: a history of the information machine*, Westview Press, Boulder, 2004, p.3

⁷"Anch'io gioco con le "net - parole"" - Intervista a Francesco Sabatini, Presidente dell'Accademia della Crusca [1]

⁸Corriere della Sera.it - Scioglilingua - Forum

⁹Corriere della Sera.it - Scioglilingua

¹⁰Arrigo Castellani *Morbus anglicus*, Studi linguistici italiani , n. 13, pp. 137-153

Cenni storici

Il *computer* è la versione più evoluta di una serie di strumenti di calcolo inventati sin dall'antichità: l'abaco, la macchina di Anticitera, i bastoncini di Nepero. Gli esemplari di macchine calcolatrici più famosi sono forse le macchine di Pascal (1645) e Leibniz (1672), ma va ricordata anche la macchina calcolatrice di Wilhelm Schickard, del 1623, della quale sono rimasti soltanto i progetti. Il passaggio da macchina calcolatrice a vero e proprio computer (nel senso di dispositivo programmabile) si deve a Charles Babbage. La sua Macchina analitica, progettata nel 1833 ma mai realizzata, è il primo computer della storia. Si trattava di una colossale macchina a ingranaggi, alimentata a vapore e dotata di input, output, unità di memoria, di unità di calcolo decimale con registro di accumulo dei dati e di un sistema di collegamento tra le varie parti. Contrariamente a quanto si potrebbe pensare, la Macchina analitica era interamente digitale¹¹.

Nel corso dei secoli seguenti il computer passerà attraverso vari stadi: il computer analogico (ne è un esempio l'analizzatore differenziale di Vannevar Bush del 1927), i computer digitali meccanici ed elettromeccanici (la Serie Z di Konrad Zuse, la macchina di Stibitz e l'ASCC di Howard Aiken) ed infine quelli digitali ed elettronici (l'ABC di John V. Atanasoff e Clifford E. Berry, l'ENIAC di John Presper Eckert e John William Mauchly, il Colossus britannico.). Nel corso del XX secolo, inoltre, importanti progressi nel campo dell'elettronica - come il transistor e il circuito integrato - e dell'informatica hanno contribuito all'evoluzione del computer nella sua forma attuale passando da dispositivo elettronico presente solo in aziende e centri di ricerca a dispositivo ad uso comune e consumo di massa per gli utenti comuni.

Descrizione

Struttura logica

La struttura logica del computer attuale può ricondursi alla presenza di almeno cinque elementi fondamentali o sottosistemi:

- Unità centrale di elaborazione (o CPU);
- Memoria;
- Schede elettroniche di espansione;
- Dispositivi di input/output;
- Canale di comunicazione dei dati (o Bus).

Questo schema venne proposto per la prima volta dal matematico John von Neumann all'interno di uno scritto informale del 1945 noto come *First draft of a report on the EDVAC* [2]. L'opera nasce dalla partecipazione di von Neumann al progetto ENIAC e raccoglie le idee provenienti da vari membri del team di sviluppo su come migliorare la funzionalità del computer nascente.

Va inoltre ricordato che von Neumann era stato profondamente influenzato da Alan Turing¹², il quale aveva proposto nel 1937¹³ un modello di calcolo - passato alla storia come Macchina di Turing - che ben si prestava a descrivere le operazioni eseguite da un computer, pur non essendo stato concepito per quello scopo (Turing si stava occupando in quella sede del problema della computabilità, non della realizzazione di un calcolatore). Il funzionamento della Macchina di Turing suggerì a von Neumann l'idea che la memoria dovesse contenere non solo i risultati delle operazioni svolte dal computer, ma anche le istruzioni di programmazione.

¹¹G. Dalakov, "The Analytical Engine of Charles Babbage"

¹²Alan Turing Scrapbook - Who invented the computer?

¹³A.M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", Proceedings of the London Mathematical Society, 1937, 2(42), pp. 230-265.

Premesso il contributo degli altri progettisti dell'ENIAC e quello di Alan Turing, a von Neumann va il merito di aver approfondito, arricchito e messo a sistema gli spunti raccolti: motivo per il quale la struttura logica sopra indicata è oggi nota come architettura di von Neumann.

Applicazioni pratiche

In sostanza un computer è in grado di eseguire operazioni logiche come ad esempio calcoli numerici e storicamente è stato proprio il calcolo la prima forma di applicazione del computer ovvero il problema che ha spinto all'invenzione e alla realizzazione del computer stesso. Dalla nascita della struttura più elementare in seguito si sono sviluppati e succeduti molti tipi di computer, costruiti per svariati compiti, professionali e non, che, pur affinando sempre più il calcolo numerico fino ad arrivare al moderno calcolo scientifico ad elevate prestazioni grazie ai supercalcolatori, vanno ben oltre offrendo applicazioni diffusissime all'utente comune come la videoscrittura in sostituzione delle vecchie macchine da scrivere e in generale applicazioni di office automation, creazione e gestione di database, la progettazione assistita al calcolatore, la computer grafica, la simulazione, sistemi elettronici di controllo, applicazioni di diagnostica medica, i videogiochi ecc.

Successo dei computer

Le chiavi di successo del computer sono dunque la grande *potenza di elaborazione* e la notevole capacità di *memorizzazione* che, in qualità di macchine, estendono le rispettive capacità umane di calcolo, processamento e memorizzazione consentendo una diminuzione sensibile dei tempi richiesti per la risoluzione di un problema dato o spesso la sua effettiva risoluzione, altrimenti non possibile. Ad essa si aggiunge la grande *usabilità* grazie al suo utilizzo in molti campi dell'elaborazione dati (vedi i sistemi embedded) e, sul fronte del consumo di massa, allo sviluppo di innumerevoli applicazioni a favore di una vasta gamma di utenti e, nella maggior parte dei casi, anche la *semplicità* di utilizzo da parte dell'utente stesso grazie all'implementazione di sistemi di interfacciamento utente/macchina molto spesso di tipo *user-friendly*.

Hardware, software, programmazione

A differenza della mente umana, intesa come attività del cervello, che è in grado di affrontare e risolvere problemi nuovi a mezzo di facoltà cognitive come intuizione e ragionamento, il computer, in quanto macchina, pur partendo ugualmente da dei dati in input, è invece un semplice *esecutore* di ordini, compiti o istruzioni impartite dall'esterno per risolvere determinati problemi d'interesse attraverso un algoritmo di risoluzione del problema stesso in modo tale che a partire da determinati input produca determinati output attraverso elaborazione degli stessi. Esso nasce dunque per eseguire *programmi* o applicazioni: un computer senza un programma da eseguire è inutile.

Il computer, in quanto esecutore, ha dunque bisogno di essere *istruito* o *programmato* dall'esterno per mano competente di un *programmatore*: questi comunica/interagisce con la macchina attraverso i linguaggi di programmazione ad alto, medio o basso livello di astrazione (linguaggio macchina), che si differenziano tra loro secondo vari paradigmi di programmazione.

Queste parti immateriali che consentono l'utilizzo di *programmi* vengono comunemente chiamate software in contrapposizione all'hardware che è invece la parte fisica (elettronica), nel senso di materiale e tangibile, degli elaboratori e che consente l'esecuzione dei programmi stessi. Tutto ciò che si può ottenere con l'utilizzo accoppiato di strumenti informatici o risorse *hardware* e *software* costituisce un'*applicazione* informatica.

Ovviamente quando si lavora su un computer scompare agli occhi dell'utilizzatore il funzionamento hardware interno della macchina interfacciandosi con essa unicamente tramite il linguaggio di programmazione ad alto livello oppure grazie all'interfaccia utente del software stesso garantendo così quella grande flessibilità di utilizzo anche a chi non ne conosce i principi fisico-elettronici di funzionamento e che costituisce, assieme alla potenza di elaborazione, a quella di memorizzazione e alla riprogrammabilità, la chiave di successo dei computer stessi. Questo processo di astrazione a partire dal livello fisico è comune a tutta l'informatica.

Il programma o software di base di gran lunga più importante per un computer (se previsto nell'architettura generale) senza il quale la macchina non potrebbe funzionare è il sistema operativo, che si occupa di gestire la macchina, le sue risorse hardware e le altre applicazioni che vi sono eseguite¹⁴. Fornisce così all'utente un'interfaccia (grafica o testuale) con la macchina gestendo o coordinando in maniera efficiente le operazioni di Input/Output comandando le rispettive periferiche di sistema, gestendo le operazioni di esecuzione dei programmi applicativi assegnandovi di volta in volta le risorse hardware necessarie al processamento (scheduling) ed infine gestendo l'archiviazione e l'accesso ai dati memorizzati attraverso il cosiddetto file system.

Dal punto di vista hardware il computer inoltre è un sistema elettronico digitale (programmabile) che elabora cioè tutti i dati in formato digitale o numerico ovvero come sequenze di 0 e 1 corrispondenti a due livelli di tensione (alto e basso) corrispondenti a loro volta ai due stati fisici di interruttore aperto e chiuso. La logica elettronica interna di funzionamento è dunque basata sull'Algebra di Boole. Il linguaggio dell'hardware è propriamente un linguaggio macchina ovvero sequenze di bit cui si associano per semplice codifica i linguaggi a più alto livello. Ovviamente l'immissione dei dati in input alla macchina elaboratrice è di tipo alfa-numerico decimale per cui è necessaria un'operazione di traduzione o codifica delle informazioni in binario; viceversa per ottenere una visualizzazione in output dei risultati dell'elaborazione è necessaria una traduzione o codifica inversa da binario ad alfa-numerico decimale. Tale codice è generalmente il codice ASCII.

A larghe linee la progettazione di sistemi di elaborazione a livello hardware ricade in massima parte nel campo dell'ingegneria informatica (la quale a sua volta afferisce conoscenze dall'ingegneria elettronica), mentre lo sviluppo software è affidato in massima parte al campo dell'informatica pura. Pur tuttavia i confini tra le due discipline non sono rigidi bensì spesso assai sfumati.

Hardware

In ingegneria elettronica e informatica con il termine **hardware** si indica la parte fisica di un computer, ovvero tutte quelle parti elettroniche, meccaniche, magnetiche, ottiche che ne consentono il funzionamento (dette anche **strumentario**).

Più in generale il termine si riferisce a qualsiasi **componente fisico** di una periferica o di una apparecchiatura elettronica.



L'hardware di un PC con raffreddamento a liquido

¹⁴Rich Didday, *Home computers: 2E10 Questions & answers, Vol II :Software*, Dilithium press, 1977 Forest Grove, Oregon

L'etimologia del vocabolo nasce dalla fusione di due termini della lingua inglese, *hard* (duro, pesante) e *ware* (manufatto, oggetto), in contrapposizione con il software, la parte logica (e perciò *soft*, "morbida, leggera") che compone il personal computer costituendone insieme all'hardware le cosiddette applicazioni.

Scheda madre

La Scheda madre è il componente principale di un PC; occupa molto spazio nel case ed è dotata di circuiti integrati che hanno il compito di collegare fra loro gli altri componenti del sistema. Alloggiati sulla scheda madre è possibile trovare componenti quali la CPU, la memoria RAM, l'insieme di unità di memoria ROM che contengono il BIOS, le schede grafiche e di rete. Tutti i componenti collegati alla *motherboard* sono divisi tra il northbridge e il southbridge.

Northbridge

Il northbridge è un circuito che permette il collegamento ad alta velocità tra i componenti critici per le prestazioni del computer:

- La CPU, che ha il compito principale di effettuare i calcoli che permettono al computer di funzionare; ci si riferisce spesso come al *cervello* del computer;
- La memoria RAM (*random-access memory*, memoria ad accesso casuale), che contiene le parti essenziali del sistema operativo in uso per una maggiore velocità di accesso da parte dei processi, e che costituisce la memoria primaria;
- Gli slot di espansione principali, come PCI, PCI Express e AGP, che ospitano schede e processori grafici o fisici ad alte prestazioni;
- Nelle schede madri più recenti sono presenti slot SATA e/o ATA per il collegamento di unità disco.

Southbridge

Il southbridge gestisce le interfacce a bassa velocità, quali:

- le porte seriali e parallele;
- le porte USB;
- l'interfaccia Ethernet;
- le interfacce per la tastiera e il mouse.

Memoria secondaria

La memoria secondaria o di massa è l'insieme di supporti che permette ai dati e alle informazioni di persistere nel sistema anche dopo lo spegnimento del computer:

- L'Hard disk, o *disco rigido*, che utilizza un supporto magnetico.
- Il Disco a stato solido, che immagazzina i dati in formato digitale, come le memorie di tipo RAM.

Supporti rimovibili

I supporti rimovibili sono unità di memoria non strettamente legate ad una singola macchina:

- CD e DVD, dischi che utilizzano sistemi ottici.
- Dischi floppy, ormai obsoleti.
- Nastri magnetici.
- Memorie flash, un tipo di memoria a stato solido.

Dispositivi di Input

Le periferiche di input consentono di acquisire dati dall'esterno. Le principali sono:

- Tastiera (standard input): acquisisce caratteri, numeri e comandi grazie alla pressione di tasti; i caratteri corrispondenti al tasto premuto sono memorizzati nel buffer di input o, in caso di digitazione troppo veloce nel buffer di tastiera.
- Dispositivi di puntamento: consentono di trasmettere dati di posizione spaziale al computer; sono dispositivi di puntamento il mouse e il touch pad.
- Scanner: grazie a una tecnologia simile a quella della fotocopiatrice, acquisisce immagini; può essere associato a un programma OCR (Optical Character Recognition), in grado di convertire l'immagine in un file di testo.
- Microfono: acquisisce e digitalizza suoni; esistono programmi di riconoscimento vocale che trasformano i suoni acquisiti in un file di testo.

Tipi di computer

I computer possono essere idealmente divisi in alcune categorie molto generali, a seconda delle loro caratteristiche salienti, dell'uso che in generale se ne fa, del software e dei sistemi operativi che fanno girare e dell'epoca in cui sono comparse. Le definizioni nel tempo sono molto cambiate e i confini non sono mai così netti.

Una prima classificazione in base all'uso distingue tra computer general purpose ovvero riprogrammabili per svariate applicazioni dall'utente e computer special purpose ovvero sistemi embedded dedicati ad una sola applicazione specifica come ad esempio i microcontrollori.

Un'altra classificazione è basata sull'accesso condiviso o meno alle risorse hardware: un computer general purpose può essere monoutente oppure multiutente sfruttando il cosiddetto timesharing delle risorse e con diversi sistemi operativi. Un computer monoutente può essere monotasking oppure multitasking ovvero può eseguire più processi in contemporanea (in pratica tutti i computer moderni). Ovviamente un computer multiutente è anche multitasking.

Una classificazione basata sulle dimensioni, sull'evoluzione storica e sulla destinazione d'uso è invece quella che segue. In linea di massima le dimensioni dei sistemi di elaborazione sono progressivamente diminuite nel tempo in virtù del miglioramento delle tecniche di elaborazione digitale ovvero all'alta capacità di integrazione dei transistor all'interno dei processori ed il parallelo aumento della capacità di processamento. Computer di grandi dimensioni rimangono ora solo nei grandi centri di calcolo.

Mainframe

Negli anni 40 i computer occupavano stanze intere, l'energia richiesta per il funzionamento ed il raffreddamento era elevata e, naturalmente, erano costosissimi; per questo motivo si tendeva a sfruttarli il più possibile e, quindi, l'utilizzo era suddiviso generalmente fra un numero di utenti piuttosto grande. Spesso la comunicazione con la macchina sia in input che in output avveniva direttamente in binario tramite l'uso di schede perforate.

Oggigiorno questi sistemi vengono utilizzati nelle grandi aziende, nelle banche, in società di assicurazioni, nella pubblica amministrazione ed in tutte quelle strutture che hanno bisogno di una potenza di elaborazione proporzionata al grande numero di utenti che devono fruirne. Dal punto di vista "estetico" non sono molto più grandi dei normali PC ma sono più potenti, però sono molto onerosi sia in termini di investimento iniziale che di costi di gestione.

Occorre sottolineare che ormai anche le realtà di piccola, se non micro, dimensione che necessitano di un'infrastruttura di rete seppur minimale¹⁵, dispongono almeno di un computer server: i prezzi delle macchine di fascia bassa sono ormai paragonabili a quelli di una buona workstation. Il termine mainframe, in questi casi (che costituiscono la stragrande maggioranza), è "esagerato" e lo si può attualmente abbinare solo ai sistemi più potenti e prestazionali tipo i blade server¹⁶.

Minicomputer

In un secondo tempo, negli anni sessanta, in particolare da Digital e da HP, vennero introdotti elaboratori dal costo abbastanza ridotto da poter essere comprati anche da piccole aziende o da singoli dipartimenti di ricerca e di dimensioni paragonabili ad un armadio. Questo permise un utilizzo più flessibile e quindi le prime sperimentazioni in campo informatico. Per distinguerli dai mainframe venne coniato il termine **minicomputer**.

Microcomputer

All'inizio degli anni settanta l'introduzione del primo microprocessore, l'Intel 4004, rese disponibili computer dal prezzo abbastanza ridotto da poter essere acquistati anche da una singola persona. La prima generazione di questi dispositivi era destinata soprattutto agli appassionati, perché di difficile utilizzo. I personal computer possono essere considerati microcomputer.

Home computer

La seconda generazione di microcomputer, che prende il nome popolare di **home computer**, fece il suo ingresso nel mercato nella seconda metà degli anni settanta e divenne comune nel corso degli anni ottanta, per estinguersi entro i primi anni novanta con l'ascesa dei personal computer.

Gli home computer, macchine a costo contenuto e di utilizzo prevalentemente domestico, contribuirono largamente a diffondere a livello popolare l'uso del computer e all'alfabetizzazione informatica di vasti strati di popolazione (specie giovanile) nei paesi sviluppati.

Basati su processori a 8 bit e costruttivamente molto semplici, erano dotati di interfacce esclusivamente testuali e come memorie di massa sfruttavano, almeno inizialmente, supporti analogici come le cassette audio (che fra l'altro, come è noto, tendono a smagnetizzarsi con l'uso). Erano utilizzati prevalentemente come console per videogiochi, oppure per i primi approcci con la programmazione.

È curioso notare che alcuni di questi computer avevano (fra l'altro) dei notevoli bug nei calcoli in virgola mobile, e quasi mai avevano sistemi di protezione della memoria, per cui si potevano fare dei danni semplicemente dando dei comandi di modifica della memoria passando degli indirizzi "opportuni" (POKE).

Con oltre dieci milioni di macchine vendute, il più rappresentativo computer di questa categoria fu il Commodore 64. Anche lo ZX Spectrum della Sinclair ebbe buona diffusione.

¹⁵Basti pensare che, anche solo per 2-3 utenti di una rete aziendale o analogica che richieda semplici servizi centralizzati e/o un software applicativo con database condiviso, ha comunque necessità di una macchina server seppur "entry level"

¹⁶L'avvento della virtualizzazione nonché la diffusione in corso della logica cloud (disporre all'esterno di macchine server e acquistare le funzioni connesse come un normale servizio) svincolerà sempre di più le organizzazioni da considerazioni sulla tipologia di "mainframe" da adottare all'interno

Personal computer

Per **Personal Computer** (PC) si intende un microcomputer economico destinato, prevalentemente, a un utilizzo personale da parte di un singolo individuo. Si distingue da un Home computer principalmente perché si prestano - grazie alle maggiori risorse hardware e software a disposizione - a utilizzi maggiormente produttivi rispetto a questi ultimi, destinati invece a un utilizzo ludico o didattico. A differenza degli Home Computer, i PC odierni sono sempre più espandibili, sempre più spesso hanno più processori per supportare il multitasking, inoltre si prestano alla multimedialità, tutte cose che al tempo degli Home Computer erano considerate quasi "fantascienza"; si pensi che nella maggior parte dei casi questi nemmeno erano dotati di disco rigido e visualizzavano al massimo 16 colori "standard" (i mezzitoni si ottenevano con artifici grafici).

Dato che la definizione di Personal Computer si consolidò definitivamente con la diffusione dei computer PC IBM, oggi per Personal Computer (PC) spesso si intende un computer da essi derivato, ma inizialmente il termine andrebbe riferito a macchine con sistemi operativi e tipi di microprocessori del tutto diversi, quali l'Olivetti P6060 e le prime versioni dell'Apple (Apple I e soprattutto l'Apple II).

Altre categorie

- I supercomputer: elevatissima capacità di elaborazione (spesso grazie ad un calcolo parallelo) e distinti dai mainframe perché solitamente destinati ad una singola applicazione come previsioni meteorologiche, climatiche, calcolo scientifico ad elevate prestazioni e simulazioni varie in generale.
- Le workstation: computer general purpose monoutente dotati di più elevate risorse di elaborazione e costi maggiori rispetto ai normali personal computer e destinati ad uso e compiti professionali.
- I microcontroller: elaboratori completi totalmente contenuti in singoli circuiti integrati e dedicati a specifiche applicazioni (specific purpose¹⁷) in sistemi embedded. Appartengono a quest'ultima categoria anche i dispositivi interni di commutazione in una rete di calcolatori.
- I sistemi barebone: personal computer preassemblati, costituiti di solito da case e scheda madre, pronti per ulteriori personalizzazioni da parte di rivenditori o utenti finali.
- I computer portatile o notebook computer trasportabili che possono entrare in una valigetta o essere appoggiati sulle ginocchia (laptop). Nella famiglia dei portatili tipologie tecnico-commerciali distinte dai notebook, propriamente detti, sono rappresentate dai netbook e dagli ultrabook.
- I computer palmari: computer di ridotte dimensioni, tali da essere portati sul palmo di una mano.
- I tablet computer computer grandi come una tavoletta e dotati di uno schermo tattile (Es. iPad della Apple, Galaxy Tab della Samsung)
- Gli smartphone hanno ormai caratteristiche molto simili (o per meglio dire sovrapponibili, c'è una crescente tendenza ad eliminare le "barriere" tradizionali nell'informatica) ai computer palmari: oltre agli apparati elettronici di telecomunicazioni per la ricetrasmisione possiedono infatti anche un nucleo elettronico di elaborazione ed una memoria dati con tanto di sistema operativo *ad hoc* e svariate applicazioni oltre ovviamente ai consueti dispositivi di input-output.
- Le console di videogiochi: rappresentano attualmente i computer domestici special purpose dalle capacità di processamento più elevate per far fronte alle più svariate esigenze di elaborazione, nonché anche il settore trainante per lo sviluppo di nuovi più evoluti processori.

¹⁷Esempi: nel cronotermostato di casa, in un impianto home video, in un forno a microonde, per non parlare delle moderne automobili o veicoli in generale.

Anche se non compresi nell'elenco, gli attuali cellulari, appena sopra il livello base, potrebbero anch'essi essere considerati computer dato che comprendono, in forma semplificata e minimale, le classiche funzioni e applicazioni da smartphone.

Si noti come queste categorie abbiano una valenza più che altro convenzionale, dal momento che le prestazioni di una "workstation" di cinque-dieci anni fa potrebbero risultare nettamente inferiori a quelle un PC odierno. È significativa in tal senso la celebre (e profetica) legge di Moore sullo sviluppo dei sistemi di elaborazione.

Personal computer

Un **personal computer** (in italiano **calcolatore personale**), solitamente abbreviato in **PC**, è un qualsiasi computer di uso generico le cui dimensioni, prestazioni e prezzo di acquisto lo rendano adatto alle esigenze del singolo individuo.

Utilizzo

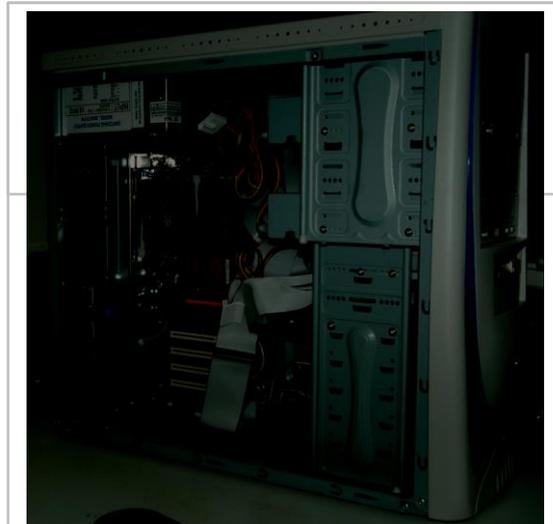
Un personal computer è pensato per essere utilizzato direttamente da un utente finale, senza l'intervento di un tecnico informatico specializzato. Questo aspetto è in contrasto con i modelli di batch processing o di time-sharing, che permettono a grandi e costosi sistemi mainframe di essere utilizzati da molti utenti, solitamente in contemporanea, o con i sistemi informatici per l'elaborazione di grandi basi di dati, che richiedono l'operato efficiente di una squadra di tecnici occupati a tempo pieno.

Tipi

Un personal computer può essere un *desktop PC* (in italiano: PC da scrivania o fisso), un *portatile*, un *netbook* (un computer portatile dalle dimensioni ancora più ridotte), un *tablet PC* (in italiano, PC a tavoletta) o un *palmare* e può essere utilizzato a casa, in ufficio o in viaggio.

Accezione e significato del termine

Personal computer dovrebbe essere, nel significato originale del termine, un qualsiasi computer che sia stato progettato per l'uso monoutente, cioè da parte di una sola persona per volta (in opposizione per esempio ai mainframe, a cui interi gruppi di persone accedono contemporaneamente attraverso terminali remoti).



Vista interna di un personal computer. In alto l'alimentatore, al centro il dissipatore della CPU. Ben visibile la scheda madre che alloggia la maggior parte delle interfacce.

Al giorno d'oggi il termine *personal computer* è invalso a indicare una specifica fascia del mercato dell'informatica, quella del computer a uso privato, casalingo, per uffici o per la piccola azienda. In alcuni casi, *personal computer* o *PC* è usato con un significato ancora più specifico, ristretto alla sola gamma delle macchine IBM compatibili. Infatti, nel gergo informatico più tecnico¹⁸, si contrappone *PC* a *Mac* intendendo, nel primo caso, un computer (IBM compatibile) con sistema operativo Windows e, nel secondo caso, un computer Apple sul quale quindi gira il sistema operativo Mac.

Il progressivo appiattimento tecnologico, inteso come avvicinamento delle caratteristiche prestazionali dei personal computer nei confronti dei server e dei mini computer, rendono spesso difficile la collocazione di un elaboratore elettronico all'interno delle canoniche classificazioni (super computer, mainframe, mini computer, server, personal computer).

Ciò che oggi differenzia un personal computer da un server o da un'avanzata console per videogiochi non è inerente alle caratteristiche fisiche, alle prestazioni o alla dotazione di interfacce, ma semplicemente alla dotazione di software e alla designazione d'uso.

L'architettura dei personal computer è ormai in grado di essere scalabile, ove necessario, e raggruppabile in *cluster* permettendo di raggiungere funzionalità e prestazioni comparabili a sistemi mainframe e supercomputer.

Con la stessa flessibilità, il personal computer può essere sostituito a specifiche tecnologie embedded progettate per applicazioni specifiche e tramite appositi software sostituirci la funzionalità.

I personal computer sono inoltre parte integrante dei più moderni impianti automatizzati a ogni livello nel settore industriale, medicale, utilities e servizi, sostituendo la maggior parte dell'elettronica sviluppata in modo verticale.

¹⁸Sito Intel <http://www.intel.it/content/www/it/it/tech-tips-and-tricks/pc-vs-mac-the-big-debate.html>

Storia



Nel 1957 l'Olivetti, impegnata su impulso di Adriano Olivetti (stimolato da Enrico Fermi) nello sviluppo dei calcolatori elettronici, grazie al contributo dell'ingegnere italo-cinese Mario Tchou (morto nel 1961 l'anno dopo di Olivetti), lancia l'Elea 9003 (il nome evoca i filosofi eleatici), presentata alla fiera di Milano davanti al presidente Giovanni Gronchi. È tra i primissimi antenati del moderno computer. Nell'ottobre del 1965 venne presentato all'esposizione Bema Show di New York l'Olivetti Programma 101, antesignana del moderno computer. La P101, ideata dall'ingegnere italiano Pier Giorgio Perotto, era una macchina da calcolo per uso personale, che possedeva un set di istruzioni interne ben definito. Era realizzata con componenti discreti e output su nastro di carta. Agli inizi degli anni '70, le persone che accedevano alle istituzioni accademiche o di ricerca, avevano l'opportunità di utilizzare sistemi come il LINC che erano

monoutente, interattivi e si potevano utilizzare per prolungate sessioni di lavoro; comunque, questi sistemi erano ancora troppo costosi per essere posseduti da una singola persona. Negli anni 70 Hewlett Packard introdusse dei computer programmabili in BASIC che potevano essere posizionati sopra una normale scrivania (serie HP 9800), includendo una tastiera, un piccolo display ad una linea ed una stampante. Il Wang 2200 del 1973 aveva un monitor (CRT) di dimensioni standard con memorizzazione su nastro magnetico (musicassette). L'IBM 5100 del 1975 aveva un piccolo display CRT e poteva essere programmato in BASIC e APL. Questi erano comunque ancora computer costosi specializzati, pensati per un utilizzo scientifico o d'azienda. L'introduzione dei microprocessori, un singolo chip con tutta la circuiteria che prima occupava grossi case, portò alla proliferazione dei personal computer dopo il 1975. I primi personal computer — generalmente chiamati microcomputer — erano spesso venduti in forma di kit ed in volumi limitati, ed erano destinati in larga parte a hobbisti e tecnici. La programmazione era effettuata attraverso interruttori e l'output era fornito attraverso led presenti sul pannello frontale. Un uso pratico avrebbe richiesto periferiche come tastiere, terminali, disk drive e stampanti. Micral N fu il primo microcomputer commerciale, non venduto in kit, basato su un microprocessore, l'Intel 8008. Fu costruito a partire dal 1972 e furono vendute circa 90 000 unità. Il primo Wikipedia:Cita le fonti "personal computer" è stato il computer Sphere 1, creato in Bountiful, Utah nel 1975 da un pioniere dei computer Michael D. Wise (1949–2002). All'inizio, Sphere 1 era venduto come kit, ma più tardi fu venduto interamente assemblato, con inclusa una tastiera un tastierino numerico ed un monitor (fu il primo Computer all-in-one). Nel 1976 Steve Jobs e Steve Wozniak vendettero la scheda madre del computer Apple I, che conteneva circa 30 chip. Il primo personal computer di successo per il mercato di massa¹⁹ è stato il Commodore PET introdotto nel gennaio del 1977²⁰, che aveva un'impressionante somiglianza con lo Sphere 1 di due anni prima. Esso fu subito seguito dal TRS-80 (rilasciato il 3 agosto 1977) di Radio Shack e dal popolare Apple II (5 giugno 1977). I computer già assemblati venduti sul mercato di massa, permisero ad una più ampia categoria di persone l'utilizzo del computer, concentrandosi di più sullo sviluppo delle applicazioni software e di meno sullo sviluppo dell'hardware del processore.

¹⁹Total share: 30 years of personal computer market share figures

²⁰The Amazing Commodore PET History

Nel 1981 fece la sua comparsa nel mercato il primo di una serie di personal computer che divenne molto popolare: l'IBM 5150, meglio conosciuto come PC IBM. Il costo era ancora elevato (circa 3.000 dollari), la capacità di elaborazione bassa, la possibilità di gestire grosse moli di dati era legata all'acquisto di costosissimi dischi rigidi, o unità a nastro esterne. D'altra parte era una macchina solida e affidabile, che godeva di assistenza tecnica ed era espandibile tramite un bus interno (caratteristica che solo l'Apple II all'epoca possedeva). Grazie al suo successo, il PC IBM divenne lo standard *de facto* nell'industria del personal computer. Le industrie informatiche situate in Oriente (Taiwan, Singapore, etc.) si misero subito al lavoro per clonare (duplicare) questa macchina (operazione possibile perché IBM forniva assieme al PC anche gli schemi elettrici, ed il listato del sistema operativo era facilmente ottenibile, i componenti utilizzati, chip di memoria, processore, unità a disco erano "standard" e disponibili per tutti). In pochi anni il mondo fu invaso da enormi quantità di PC clonati, dalle prestazioni sempre più brucianti e dai costi sempre più bassi.

Nel frattempo il 24 gennaio 1984 Apple Computer presentò un'altra innovazione destinata a diffondersi su larga scala: il Macintosh, dotato di serie di interfaccia grafica e di mouse. Il Macintosh ottenne un buon successo di mercato, grazie al suo approccio *amichevole* (*user-friendly*) e alla facilità d'uso del suo sistema operativo, il Mac OS. L'interfaccia grafica (GUI) usava per la prima volta metafore facili da comprendere, quali il *cestino*, la *scrivania*, le *finestre*, gli *appunti* ecc. aprendo finalmente l'uso del computer anche a persone con limitate conoscenze informatiche.

In seguito al successo del Macintosh, molte di queste caratteristiche furono mutate dalla Microsoft nella creazione del proprio sistema operativo Windows, scatenando una battaglia anche legale durata oltre un decennio. Si diffondeva così il concetto di WIMP (*Windows Icons Mouse and Pointer*), da allora in poi bagaglio essenziale di tutti i personal computer, indipendentemente da quale fosse il loro sistema operativo.

Caratteristiche generali

Dal punto di vista hardware, i più comuni microprocessori utilizzati all'interno dei personal computer sono le CPU compatibili con l'architettura x86.

Una differenza fra i primi computer disponibili per autocostruzione e i personal computer veri e propri fu la disponibilità di software già pronto. Nonostante ciò, molti fra i primi utilizzatori di PC erano abituati a scrivere da sé il software. Al contrario, gli utenti di oggi hanno accesso a una larghissima scelta di software commerciale e non commerciale, che può essere fornito pronto per essere eseguito o pronto per essere compilato.

Sotto questo punto di vista le applicazioni software tipiche per personal computer sono quelle a uso di un utente comune e comprendono programmi di videoscrittura, fogli elettronici, database, navigatori web, programmi di posta elettronica, giochi e miriadi di altre applicazioni per la produttività personale e gli ambiti applicativi più disparati.

I personal computer possono essere connessi a una rete locale (LAN), sia tramite l'utilizzo di un cavo, sia con l'uso di connessioni radio (wireless).

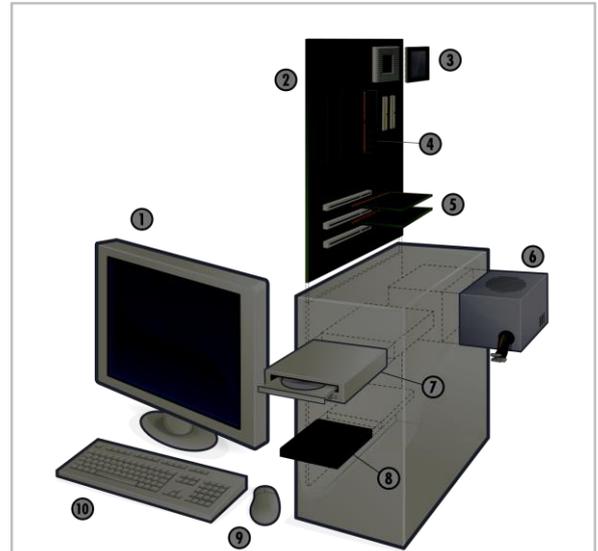
Gli attuali personal computer sono spesso predisposti per connessioni ad alta o bassa velocità a Internet, consentendo l'accesso al World Wide Web e a una vasta scelta di altre risorse.

Fin dagli anni ottanta del XX secolo, Microsoft e Intel hanno dominato buona parte del mercato dei personal computer con la piattaforma Wintel.

Componenti

- Unità centrale contenente lo spazio e gli agganci per ospitare:
 - Scheda madre, memoria e CPU,
 - Connettori,
 - Alimentatore,
 - Possibili periferiche di espansione,
 - Alloggi per le periferiche di base e i dispositivi per le memorie di massa (disco rigido, CD ecc.)
- Monitor/Display (o analogo unità di output esterna per la visualizzazione dei dati)
- Periferiche di input/output
 - Tastiera
 - Mouse
 - Interfacce seriali, parallele, USB
 - Diffusore acustico o Casse acustiche

Questa particolare struttura modulare rende il personal computer differente rispetto alle altre categorie di computer, e precisamente dai Mainframe (i cosiddetti computer "ad armadio") tipici dei grandi centri di calcolo come il PDP1 o il PDP 11 della Digital e dai microcomputer (computer ad uso hobbistico da assemblare, tipici degli albori dell'era informatica, che si programmavano tramite serie di microswitch, e avevano capacità di output limitate a batterie di led colorati) Altair.



Componenti fondamentali di un personal computer:

1. Monitor
2. Scheda madre
3. CPU (microprocessore)
4. RAM
5. Scheda di espansione
6. Alimentatore
7. Dispositivi ottici
8. Hard disk drive (HDD)
9. Mouse
10. Tastiera

Architettura (hardware e software)

Un computer, come premesso, non è altro che l'implementazione fisica, pratica, di una macchina di Turing, secondo l'architettura fisica ideata da Von Neumann. Per quanto riguarda l'architettura hardware sono diffuse le architetture dei processori basati su registri generali. A questo livello hardware e software viaggiano appaiati sullo stesso piano: le istruzioni elementari del processore (Instruction set) (salto, confronto, operazioni aritmetiche ecc.) sono definite in linguaggio macchina, ovvero codificate in sequenze numeriche di bit; i linguaggi di programmazione a medio ed alto livello, cioè più vicini o familiari al nostro linguaggio naturale, si ottengono come ulteriori codifiche (tabelle di associazioni tra i nomi delle istruzioni) a partire dal linguaggio macchina stesso. Così accade che un qualunque software applicativo abbia un suo codice oggetto o eseguibile espresso in linguaggio macchina ed un suo codice sorgente espresso in linguaggio ad alto livello. Una terza possibilità è l'assembly; questo particolare linguaggio è una "trascrizione" delle istruzioni base del processore in un linguaggio intelligibile (solitamente nel formato istruzione indirizzo, o istruzione valore, o altri - vedasi opcode). È possibile (per esempio) integrare l'assembly in programmi in C (ma questo approccio non è sempre consigliabile anzi oggi è quasi del tutto abbandonato perché si preferisce scrivere programmi adattabili a più piattaforme, vedi voce "Portabilità"; linguaggio macchina ed assembly non sono mai portabili mentre esistono compilatori C multipiattaforma). Con l'evolversi dei sistemi operativi (eccezion fatta per l'hacking) i linguaggi a basso livello sono usati sempre meno, invece all'epoca del DOS e soprattutto degli home computer questo approccio veniva seguito spesso.

Un processore munito di memoria, dispositivi di input/output e bus di collegamento tra periferiche è dunque un "elaboratore" secondo la classica architettura generale di Von Neumann.

All'architettura hardware si sovrappone l'architettura software che comprendente il sistema operativo cioè il sistema generale di gestione o coordinazione delle risorse hardware del computer e dei dati in esso immagazzinati. Le istruzioni base del sistema operativo fanno parte del set di istruzioni fondamentali del processore e, di conseguenza, sono definite anch'esse in linguaggio macchina e gestite dall'Unità di Controllo del processore. Le istruzioni del sistema operativo a livello utente, eventualmente interfacciate graficamente, sono anche qui ottenute tramite una codifica ad alto livello delle istruzioni base, espresse in linguaggio macchina, di gestione (pilotaggio) delle periferiche stesse, mascherando così ancora una volta il complesso funzionamento interno del computer.

L'insieme di architettura hardware (quindi il processore) e sistema operativo costituiscono la cosiddetta piattaforma informatica del sistema essendo questi due elementi strettamente correlati.

Software

Il **software** è l'informazione o le informazioni utilizzate da uno o più sistemi informatici e memorizzate su uno o più supporti informatici. Tali informazioni possono essere quindi rappresentate da uno o più programmi, oppure da uno o più dati, oppure da una combinazione delle due.

Il termine si contrappone tradizionalmente a hardware (la componente fisica di un sistema di calcolo) che rende possibile l'esecuzione del software. Nel tempo sono entrati nell'uso altri termini che descrivono elementi di un computer, come il firmware. Il suffisso *-ware* viene usato anche in altri termini che indicano particolari tipi di programmi: in funzione del ruolo che hanno in un sistema di calcolo (per esempio middleware); del tipo di licenza con cui sono distribuiti (freeware, shareware); e altro.

Etimologia

Il termine "software" è un prestito dell'inglese. Nella lingua inglese "software" nasce per imitazione del termine "hardware" e dalla composizione delle parole "soft" (in italiano "morbido", "tenero", "leggero")²¹ e "ware" (in italiano "merci", "articoli", "prodotti", "mercanzie").²²

La paternità della coniazione del termine "software" non è certa. L'americano Paul Niquette la rivendica sostenendo di aver coniato il termine "software" nel 1953.²³ Certa invece è la prima apparizione del termine "software" in una pubblicazione scritta: un articolo dell'*American Mathematical Monthly* scritto nel 1958 dallo statistico americano John Wilder Tukey.²⁴

Storia del software

Il termine *software* ha origine durante la seconda guerra mondiale. I tecnici dell'esercito inglese erano impegnati nella decrittazione dei codici tedeschi di *Enigma*, di cui già conoscevano la meccanica interna (detta *hardware*, *componente dura*, nel senso di ferraglia) grazie ai servizi segreti polacchi. La prima versione di *Enigma* sfruttava tre rotori per mescolare le lettere.

Dopo il 1941, ad *Enigma* venne aggiunto un rotore, e il team di criptanalisti inglesi, capitanati da Alan Turing, si dovette interessare non più alla sua struttura fisica, ma alle posizioni in cui venivano utilizzati i rotori della nuova *Enigma*.

Dato che queste istruzioni erano scritte su pagine solubili nell'acqua (per poter essere più facilmente distrutte, evitando in tal modo che cadessero nelle mani del nemico) furono chiamate *software* (*componente tenera*), in contrapposizione all'*hardware*.

Il senso moderno del termine deriva dalle istruzioni date ai computer, ed è stato utilizzato per la prima volta nel 1957 da John Wilder Tukey, noto statistico statunitense.

Dal 1950 l'analogia tra l'*hardware* ed il corpo umano e quella tra il *software* e la mente umana si è fatta molto forte, dal momento che Turing ha sostenuto che il progresso tecnologico sarebbe riuscito a creare, entro il 2000, delle macchine intelligenti (in grado cioè di "pensare" autonomamente) atte alla risoluzione dei problemi.

Aumento del codice e potenziamento dell'hardware

Alla storia dell'evoluzione del software è legato lo sviluppo dell'hardware. Come evidenziato dalla seconda legge di Moore, una minaccia alla velocità di elaborazione, oltre ai costi, proviene dal software. Infatti ciò che conta per un utente non è tanto la velocità di elaborazione del processore, quanto la velocità effettiva di elaborazione del codice, calcolata in base al tempo che occorre alla CPU per eseguire un'operazione (come la scrittura di un testo, la creazione di una cartella, ecc.).

Nathan Myhrvold, direttore dell'Advanced Technology Group della Microsoft, ha effettuato uno studio sui prodotti Microsoft calcolando le linee di codifica per le successive *release* dello stesso software:

- Basic: da 4.000 linee di codice nel 1975 a 500.000 nel 1995
- Word: da 27.000 linee di codice nel 1982 a 2.000.000 nel 2002

La continua aggiunta di nuove funzionalità al software esistente giustifica la costante richiesta di processori più veloci, memorie sempre più grandi e più ampie capacità di I/O (Input/Output).

Infatti, anche le altre tecnologie si sono evolute di pari passo:

- i dischi rigidi da 10 MB (1982) a 1 TB (2007);

²¹Confronta il lemma "soft" su «Il Sansoni - Inglese».

²²Confronta il lemma "ware" su «Il Sansoni - Inglese».

²³Paul Niquette, *Introduction: the software age*, «Sophisticated: The Magazine», 1995.

²⁴David Leonhardt, *John Tukey, 85, statistician; coined the word 'software'*, «New York Times», 28 luglio 2000.

- i modem analogici da 110 bit/sec a 56 kbit/sec.

Myhrvold traccia un parallelismo con la legge di Moore: "*abbiamo aumentato la dimensione e la complessità del software ancora più rapidamente di quanto non prevedeva la legge di Moore*", "*gli utenti del software hanno sempre consumato le maggiori capacità di elaborazione ad una velocità uguale o superiore a quella con cui i produttori di chip le mettevano a disposizione*" (Stewart Brand, 1995).

Classificazione

I software possono essere classificati in base a diverse loro caratteristiche:

- funzione (videoscrittura, foglio elettronico, database management system, grafica ecc.);
- grado di apertura della licenza (software libero o software proprietario);
- sistema operativo su cui possono essere utilizzati (Unix, Mac OS, Windows ecc.);
- da installare o portabile;
- tipo di interfaccia utente (testuale o grafica);
- *stand alone* (ovvero che possono girare completamente autonomi su sistemi isolati) oppure *network* (ovvero che funzionano in un ambito di rete). Il secondo caso può essere a sua volta diviso in software in versione client-server (sul client può essere installato magari solo un *agent* oppure l'intero software che però si appoggia ad un database installato su di un server), oppure in modalità *terminal-server* oppure cloud (applicazioni web)²⁵. Esempi classici di programma di rete, in ambito aziendale, è il sistema gestionale, la posta elettronica centralizzata (spesso associata ai moduli di pianificazione attività e risorse), il firewall della LAN.

Dal punto di vista gerarchico²⁶ i software possono essere divisi in quattro²⁷ categorie principali:

- firmware;
- software di base (che a sua volta si divide in sistemi operativi, compilatori e interpreti, librerie);
- driver;
- programmi applicativi cioè tutti quei software che vengono utilizzati nel quotidiano (*home*), dai programmi per l'ufficio, ai videogiochi, ai browser per navigare in internet, ai client di posta, ai player audio-video, alle applicazioni chat nonché i numerosissimi prodotti specifici in ambito professionale (*business*) per le più disparate esigenze/attività.

Con il termine *suite* si designa un software strutturato in diversi programmi/moduli, solitamente configurabili (nel senso di installabili, attivabili) separatamente ma facenti parte di un'unica soluzione. Microsoft Office o AVG AntiVirus o SAP ERP sono esempi di suite software.

²⁵Esempio semplice: Facebook è un'applicazione software web ("full web" in quanto non necessita di alcun agent installato a differenza, ad esempio, di Skype)

²⁶La distinzione è relativa al grado di "utilizzabilità" e "prossimità" rispetto all'utente.

²⁷Si potrebbe anche considerare una quinta categoria di software: gli "strumenti" (*tool* nel linguaggio informatico). Si tratta di programmi di gestione di una periferica (*device*), ad esempio il pannello di controllo della stampante, della scheda video, di un disco esterno, ecc. Si tratta di software di interconnessione tra il sistema operativo e periferica (e quindi dei driver relativi) che consentono di regolare (*setup*) la periferica e il suo funzionamento nei confronti del sistema operativo o di altri programmi. Esistono anche tool di gestione di software applicativi, ad esempio il pannello di controllo di un *anti-malware* oppure quello di supervisione delle applicazioni specifiche del produttore del computer (OEM). Infine, esistono tool per la gestione (installazione, modifica-sviluppo, aggiornamento, manutenzione, ecc) di firmware o sistemi operativi

Versione

Molte volte i software realizzati vengono rilasciati sotto forma di versioni e release successive, ciascuna identificata da un numero intero progressivo con in aggiunta uno o più numeri decimali che identificano la release: tipicamente l'ordine di rilascio segue un andamento progressivo della numerazione dove versione o release successive rappresentano evoluzioni delle precedenti con miglioramenti in termini di nuove caratteristiche e funzionalità aggiunte e/o bug corretti.

Licenze d'utilizzo e distribuzione

La licenza d'uso è un documento che accompagna il *software* e specifica i diritti e i doveri di chi lo riceve e di chi lo diffonde.

Tutte le licenze d'uso traggono il loro valore legale dalle norme sul diritto d'autore (il *copyright*).

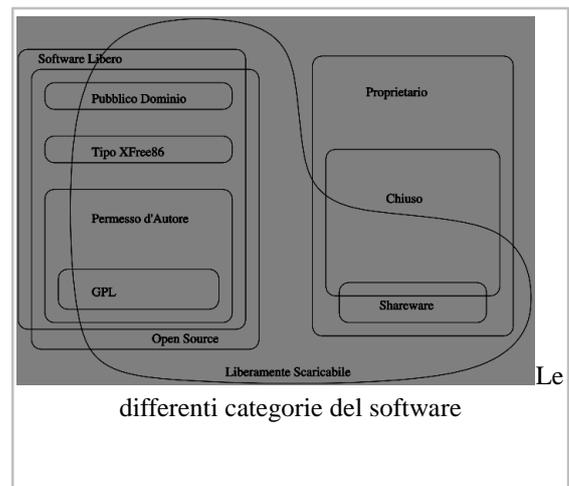
Esistono licenze libere, le licenze Open Source e licenze proprietarie. Nasce in seguito anche l'Open content che ha come scopo quello di trasferire le licenze su opere diverse dal software.

Le licenze di utilizzo e distribuzione del software libere ed Open Source sono numerose, ma quelle effettivamente diffuse sono poche. Per l'89% si tratta di GPL, LGPL e BSD (licenza storica di Unix, tornata in uso dall'avvento di Linux).

Alcune licenze libere:

- GNU-GPL (la licenza libera più diffusa)
- GNU-LGPL
- BSD
- Creative Commons

Ogni tipo di licenza differisce dagli altri per vari aspetti



File

In informatica il termine **file** (termine inglese per "archivio", ma comunemente detto anche "documento") indica un contenitore di informazioni/dati in formato digitale, tipicamente presenti su un supporto digitale di memorizzazione. Le informazioni scritte/codificate al suo interno sono leggibili solo attraverso software.

Descrizione

Tecnicamente, i dati codificati in un file sono organizzati come una sequenza (di byte), immagazzinati come un solo elemento su una memoria di massa attraverso il File system (sistema di archivi) esistente su quella data memoria di massa.

Ciascun file è identificato da un nome univoco, un'estensione ed un percorso (*path*) che ne individua posizione, contenitore, cartella o directory in uno spazio di nomi gerarchico all'interno del file system stesso.

Se dal punto di vista dell'utente un file è solitamente un singolo elemento, può succedere invece che fisicamente sia scritto o risieda su più zone diverse del supporto di memorizzazione che lo ospita: questo fenomeno è molto comune se il supporto di memorizzazione è un disco di memoria, mentre è molto raro su nastri magnetici. Uno dei compiti del sistema operativo è rendere trasparente alle applicazioni la reale suddivisione fisica del file e occuparsi di gestire il recupero delle informazioni in esso contenute (*lettura*) dai vari blocchi fisici del supporto di memorizzazione e la *scrittura*.

Il contenuto dei file è normalmente conforme ad un particolare formato, e per ciascun formato esistono una o più applicazioni che sono in grado di interpretarne e/o di modificarne il contenuto ("aprire" il file).

Alcuni sistemi operativi, come Microsoft Windows e l'MS-DOS, riconoscono il formato di un file in base all'estensione del loro nome; altri, come il Mac OS, da una serie di metadati salvati insieme al file; altri ancora, come Unix, identificano i tipi di file in base ai primi byte del loro contenuto, detti *magic number*.

Un tipo di file molto comune sono i file di testo ASCII. Un simile file è una sequenza di caratteri ASCII pura e semplice, tale da poter essere letta ed interpretata nello stesso modo da tutti i sistemi operativi. Il termine "file binario", invece, si riferisce solitamente a tutti i file che non sono di testo ASCII.

Il concetto di file è molto semplice ed elegante: in ultima analisi, un file è una sorgente (o un deposito) di informazioni, che si può leggere e scrivere; questa sorgente/deposito ha anche delle "proprietà" (nome, estensione, flag) che possono essere modificate. Questa definizione si adatta molto bene a molte periferiche ed interfacce hardware; i sistemi operativi Unix e derivati (Unix-like) hanno generalizzato il concetto di file tanto da farne una vera filosofia: in Unix *tutto è un file*, cioè può essere "aperto", "chiuso", "letto", "scritto" eccetera; questi "file speciali" possono essere delle comunicazioni fra processi, delle *pipe*, delle interfacce hardware o altro ancora: il sistema operativo si occuperà di gestire tutto nel modo appropriato lasciando ai programmi l'illusione di stare usando un normale file.

Tipi di file

- File di testo;
- File audio (es. file musicale);
- File video (es. film);
- File multimediale (audio-video);
- File immagine;
- File di configurazione;
- File di log;
- File binario;
- File sorgente;
- File oggetto;
- File eseguibile.

Operazioni sui file

Poiché accedere ad un archivio informatico da una memoria di massa è una operazione piuttosto lenta, necessita di una conoscenza dettagliata dell'hardware del computer e crea un rischio di conflitto di risorse fra programmi in esecuzione, queste operazioni sono eseguite dal sistema operativo per conto dei programmi che le richiedono. Le operazioni più comuni, universalmente presenti in tutti i sistemi operativi, sono:

- **Apertura:** il programma segnala al sistema operativo che ha necessità di accedere a un certo file. Il sistema operativo controlla che il file esista e che non sia già usato o bloccato, da esso o da un altro programma; crea alcune strutture dati per gestire le operazioni successive e riserva una certa quantità di memoria RAM, detta buffer, per memorizzare i dati in transito da e per il file; restituisce al programma un simbolo (o un *handle* oppure un indirizzo di memoria fisica) a cui il programma farà riferimento nelle successive operazioni sul file; gli altri programmi possono accedere allo stesso file solo in modo limitato (solo lettura) o non possono accedervi affatto.
- **Lettura:** il programma richiede dei dati dal file. Il sistema operativo li legge, li memorizza nel buffer di cui comunica l'ubicazione al programma richiedente.
- **Scrittura:** Il programma vuole scrivere dei dati in un file. Per prima cosa memorizza i dati in questione in un buffer, di cui poi comunica l'ubicazione al sistema operativo.
- **Chiusura:** il programma comunica che non ha più bisogno del file. Tutte le strutture dati e il buffer allocato vengono rilasciate; il sistema operativo elimina il file dalla lista di quelli in uso; gli altri programmi hanno ora libero accesso al file.

Di per sé il sistema operativo è in grado di compiere le seguenti operazioni:

- **Riallocazione:** il sistema operativo (file system) è in grado di spostare o muovere il file da una directory all'altra tramite ad esempio copia e incolla o da terminale a riga di comando.
- **Rinominazione:** il sistema operativo è in grado di rinominare il file assegnandogli un nome diverso per via grafica o da riga di comando.
- **Eliminazione:** il sistema operativo è in grado di rimuovere, se richiesto dall'utente, il file dalla memoria di massa che lo contiene, andando però solo ad eliminare il collegamento logico-fisico del file in memoria, il cui spazio di memoria risulterà dunque non vuoto, ma *libero* nel senso di sovrascrivibile.

Un programma non può leggere o scrivere un file se prima non lo ha "aperto", e una volta finito di usarlo lo deve sempre "chiudere" per dare modo al sistema operativo di liberare le risorse occupate.

Trasferimento di file

Si possono utilizzare vari modi e tipologie di connessione diverse per trasferire file in una rete. Per esempio è possibile condividere e trasferire i file in una rete LAN tra due o più computer, utilizzare la posta elettronica, oppure usare dei servizi o programmi appositamente dedicati allo scopo (es. FTP).

Il trasferimento di file attraverso una rete può diventare problematico se non si dispone di una connessione sufficientemente veloce, in particolare se si tratta di file di grosse dimensioni. Per esempio per trasferire un file tramite Internet è opportuno contenerne la dimensione, altrimenti può essere necessario molto tempo per eseguire l'operazione. Quando le connessioni tra computer erano particolarmente lente, per contenere la dimensione di un file si usavano varie tecniche: per esempio era possibile comprimerlo o suddividere il singolo file in file più piccoli e spedirli singolarmente; queste operazioni potevano richiedere l'uso di programmi specifici. Con l'avvento della banda larga queste operazioni preliminari sui file stanno andando in disuso.

È possibile trasferire uno o più file tramite la comune e-mail. Per il trasferimento via e-mail va tenuto presente che normalmente i fornitori di servizi Internet limitano la dimensione di ogni singolo messaggio: per esempio Gmail limita la dimensione degli allegati in trasmissione e in ricezione a 25 Megabyte. Inoltre l'intera casella e-mail può avere dei limiti nella sua dimensione totale: per esempio Alice, nei contratti ADSL residenziali, limita a 3 Gigabyte la dimensione dell'intero contenuto di un account e-mail. È necessario tenere presente questi limiti quando si inviano file di grandi dimensioni, in modo da non rischiare che l'allegato non venga spedito, o di occupare interamente la casella di posta del destinatario.

In alternativa all'e-mail è possibile usare un programma di messaggistica istantanea, iscrivere il contatto e creare una cartella condivisa. Questi programmi, normalmente, non praticano limitazioni di banda legate all'estensione, alla dimensione e al tipo di file inviato.

Per il trasferimento di grandi quantità di file esistono anche dei servizi web di desktop remoto, che permettono di "parcheggiare" per alcuni mesi i propri file in un archivio privato, cui si accede con username e password (es. Dropbox) L'archivio può essere aperto alle persone desiderate dando loro username e password. Una variante di questi servizi sono quelli di upload remoto che creano una chiave di sessione comune a due utenti iscritti e collegati a Internet; caricato il file, il ricevente deve indicare un percorso dove salvarlo e autorizzarne il download.

Recupero dei file

È possibile recuperare un file cancellato definitivamente, utilizzando appositi programmi in grado di ripristinarli.

Quando si conferma la cancellazione di un *file*, per esempio svuotando il cestino di Windows, viene infatti cancellato (logicamente o fisicamente) il solo collegamento fisico esistente tra il nome del file ed il suo contenuto, per cui il file non è più accessibile né all'utente né ai programmi del computer. In assenza di altri collegamenti fisici, anche l'area della memoria di massa occupata dal contenuto dei file viene designata come disponibile per nuovi contenuti, tuttavia essa non viene cancellata o sovrascritta immediatamente, per cui i dati risultano ancora fisicamente disponibili sulla memoria con opportune tecniche fino all'effettiva sovrascrittura.

In particolare i programmi per il recupero dei file esaminano il contenuto di un *file system* cercando collegamenti fisici cancellati logicamente e controllando le aree occupate dai dati a cui essi fanno riferimento, ricopiandone il contenuto e, se possibile, ricostruendo le parti ormai sovrascritte.

Per prevenire invece il recupero di file cancellati (per esempio per questioni di privacy) esistono programmi che provvedono a sovrascrivere (anche più volte) il contenuto di un file prima di cancellarlo, o a sovrascrivere tutte le aree libere di un *file system*, in modo che non sia possibile il recupero dei dati.

Reti di computer

Più calcolatori possono essere collegati insieme in rete formando una rete di calcolatori sotto forma di sistemi distribuiti per la condivisione di dati e delle risorse software e/o hardware come nel caso del calcolo distribuito. Un esempio di rete di calcolatori sono le reti aziendali (Intranet) e la rete Internet. In particolare queste reti si suddividono in reti client-server in cui esistono calcolatori che erogano servizi (server) a calcolatori o utenti che ne fanno richiesta (client) (ad es. le Reti Windows e le Reti Linux), oppure reti peer to peer che offrono un'architettura logica paritaria dove ciascun calcolatore può svolgere sia funzioni di server che funzioni di client. Anche i sistemi di tipo mainframe formano una rete informatica tipicamente con risorse hardware centralizzate e condivise. I vantaggi di queste reti sono evidenti in termini di facilità ed efficienza di comunicazione e gestione delle informazioni all'interno delle organizzazioni o aziende private stesse. A livello logico tutte queste tipologie di reti sono realizzate e si differenziano tra loro grazie a diversi protocolli di comunicazione utilizzati e relativi software, che caratterizzano e rendono possibile quindi le funzionalità stesse di rete.

Rete / Mainframe

La costruzione di reti di calcolatori può essere fatta risalire alla necessità di condividere le risorse di calcolatori potenti e molto costosi (mainframe). La tecnologia delle reti, e in seguito l'emergere dei computer personali a basso costo, ha permesso rivoluzionari sviluppi nell'organizzazione delle risorse di calcolo.

Si possono indicare almeno tre punti di forza di una rete di calcolatori rispetto al mainframe tradizionale:

1. *fault tolerance* (resistenza ai guasti): il guasto di una macchina non blocca tutta la rete, ed è possibile sostituire il computer guasto facilmente (la componentistica costa poco e un'azienda può permettersi di tenere i pezzi di ricambio in magazzino);
2. *economicità*: come accennato sopra, hardware e software per computer costano meno di quelli per i mainframe;
3. *gradualità della crescita e flessibilità (scalabilità)*: l'aggiunta di nuove potenzialità a una rete già esistente e la sua espansione sono semplici e poco costose.

Tuttavia una rete mostra alcuni punti deboli rispetto a un mainframe:

1. *scarsa sicurezza*: un malintenzionato può avere accesso più facilmente ad una rete di computer che ad un mainframe: al limite gli basta poter accedere fisicamente ai cablaggi della rete. Inoltre, una volta che un *worm* abbia infettato un sistema della rete, questo si propaga rapidamente a tutti gli altri e l'opera di disinfezione è molto lunga, difficile e non offre certezze di essere completa;
2. *alti costi di manutenzione*: con il passare del tempo e degli aggiornamenti, e con l'aggiunta di nuove funzioni e servizi, la struttura di rete tende ad espandersi e a diventare sempre più complessa, e i computer che ne fanno parte sono sempre più eterogenei, rendendo la manutenzione sempre più costosa in termini di ore lavorative. Oltre un certo limite di grandezza della rete (circa 50 computer) diventa necessario eseguire gli aggiornamenti hardware e software su interi gruppi di computer invece che su singole macchine, vanificando in parte il vantaggio dei bassi costi dell'hardware.

Tipi di reti

Esiste una grande varietà di tecnologie di rete e di modelli organizzativi, che possono essere classificati secondo diversi aspetti:

Classificazione sulla base dell'estensione geografica

A seconda dell'estensione geografica, si distinguono diversi tipi di reti:

- si parla di *rete corporea* o BAN (*Body Area Network*) se la rete si estende intorno al corpo dell'utilizzatore con una estensione dell'ordine del metro
- si parla di *rete personale* o PAN (*Personal Area Network*) se la rete si estende intorno all'utilizzatore con una estensione di alcuni metri
- si parla di *rete locale* o LAN (*Local Area Network*) se la rete si estende all'interno di un edificio o di un comprensorio, con una estensione entro alcuni chilometri
- si parla di *rete universitaria* o CAN (*Campus Area Network*), intendendo la rete interna ad un campus universitario, o comunque ad un insieme di edifici adiacenti, separati tipicamente da terreno di proprietà dello stesso ente, che possono essere collegati con cavi propri senza far ricorso ai servizi di operatori di telecomunicazioni. Tale condizione facilita la realizzazione di una rete di interconnessione ad alte prestazioni ed a costi contenuti.
- si parla di *rete metropolitana* o MAN (*Metropolitan Area Network*) se la rete si estende all'interno di una città
- si parla di *rete geografica* o WAN (*Wide Area Network*) se la rete ricopre un'area geografica molto estesa e che intercorre nelle reti precedenti
- si parla di *rete globale* o GAN (*Global Area Network*) se la rete collega calcolatori dislocati in tutto il mondo, anche via satellite

Classificazione in base al canale trasmissivo

Reti Locali

Le reti locali vengono realizzate tipicamente utilizzando un sistema di cablaggio strutturato con cavi UTP in categoria 5 o superiore, che serve uno o più edifici utilizzati tipicamente da una stessa entità organizzativa, che realizza e gestisce la propria rete, eventualmente con la cooperazione di aziende specializzate.

In molti casi, il cablaggio è complementato o sostituito da una copertura wireless.

Le LAN vengono realizzate soprattutto con la tecnologia ethernet, e supportano velocità di 10/100 Mbit/s, o anche 1 Gbit/s, su cavi in rame dalle caratteristiche adeguate (CAT5 o superiore), o su fibra ottica.

Reti pubbliche – Distribuzione

Le reti pubbliche sono gestite da operatori del settore, e offrono servizi di telecomunicazione a privati ed aziende in una logica di mercato.

Per poter offrire servizi al pubblico, è necessario disporre di una infrastruttura di distribuzione che raggiunga l'intera popolazione.

Per ragioni storiche, la gran parte delle reti pubbliche sono basate sul doppino telefonico (dette anche POTS, *Plain Old Telephone System*). Questa tecnologia era stata studiata per supportare il servizio di telefonia analogica, ma data la sua pervasività e gli alti investimenti che sarebbero necessari per sostituirla è stata adattata al trasporto di dati mediante diverse tecnologie:

- i modem per codificare segnali digitali sopra le comuni linee telefoniche analogiche. Il grande vantaggio di questa tecnologia è che non richiede modifiche alla rete distributiva esistente. Sono necessari due modem ai due capi di una connessione telefonica attiva per stabilire una connessione. Molti fornitori di servizio offrono un servizio di connettività Internet via modem mediante batterie di modem centralizzate. La velocità è limitata a circa 56 Kbit/s, con l'adozione di modem client e server che supportano la versione V92 dei protocolli di comunicazione per modem. Questo protocollo incorpora funzioni di compressione del flusso di bit trasmesso, quindi la velocità effettiva dipende dal fattore di compressione dei dati trasmessi.

- le reti ISDN trasmettendo dati e voce su due canali telefonici in tecnologia digitale. Mediante appositi adattatori, è possibile inviare direttamente dati digitali. La tecnologia ISDN è ormai molto diffusa nei paesi sviluppati. Usandola per la trasmissione di dati, arrivano ad una velocità massima di 128 Kbit/s, senza compressione, sfruttando in pratica due connessioni dial-up in parallelo, possibili solo con determinati provider. La velocità su un singolo canale è invece limitata a 64 Kbit/s. Ci sarebbe un terzo canale utilizzato per il segnale ma non per la comunicazione con una capacità di 16 Kbit/s (Esso non viene mai utilizzato per i dati).

Utilizzando modem analogici o ISDN, è possibile stabilire una connessione dati diretta tra due qualsiasi utenze della rete telefonica o ISDN rispettivamente.

- la tecnologia ADSL (Asymmetric Digital Subscriber Line) utilizza una porzione della banda trasmissiva disponibile sul doppino telefonico dalla sede dell'utente alla centrale telefonica più vicina per inviare dati digitali. È necessaria l'installazione di nuovi apparati di commutazione nelle centrali telefoniche, chiamati DSLAM, e l'utilizzo di filtri negli impianti telefonici domestici per separare le frequenze utilizzate per la trasmissione dati da quelle per la comunicazione vocale. La loro diffusione sul territorio è limitata dai costi, che la rendono conveniente solo nelle aree maggiormente sviluppate. Durante la connessione tramite ADSL è possibile continuare a utilizzare il telefono in quanto le frequenze della voce e dei dati non si sovrappongono. Questa tecnologia è inoltre chiamata *Asimmetric* in quanto le velocità di download e di upload non sono uguali: in Italia sono tipicamente pari a 4 Mbit/s in download e 512 Kbit/s in upload, ma per certi abbonamenti la velocità di download può arrivare anche a 12 Mbit/s, o anche 24 Mbit/s, usando tecnologie di punta come ADSL2+ e reti di distribuzione in fibra ottica di ottima qualità. Il doppino di rame presenta l'inconveniente di attenuare i segnali, e non permette il funzionamento di questa tecnologia per distanze superiori ai 5 km circa. In alcuni casi è anche possibile un'ulteriore riduzione della distanza massima dovuta a interferenze esterne che aumentano la probabilità d'errore. Un'altra limitazione importante è data dall'interferenza "interna", che si verifica quando molte utenze telefoniche sullo stesso cavo di distribuzione utilizzano il servizio ADSL. Questo fa sì che non si possa attivare il servizio ADSL su più di circa il 50% delle linee di un cavo di distribuzione.

ADSL è l'ultimo sviluppo sull'infrastruttura esistente di doppino telefonico.

Per superare queste velocità, l'infrastruttura di distribuzione basata sul doppino dovrà essere sostituita da supporti fisici più performanti.

Tra i candidati a sostituire il doppino per la distribuzione domestica dei servizi di telecomunicazioni, si possono citare:

- le fibre ottiche:
- le infrastrutture della TV via cavo (diffusa soprattutto negli USA)
- il trasporto di dati sulla rete elettrica (es. powerline) o nelle condutture del gas.
- le reti wireless
- le reti satellitari (che però sono tipicamente unidirezionali, dal satellite alla casa dell'utente, mentre il canale di ritorno deve essere realizzato con altre tecnologie, spesso su doppino telefonico).

Internet

Internet²⁸ (contrazione della locuzione inglese *interconnected networks*, ovvero "reti interconnesse") è una rete mondiale di reti di computer ad accesso pubblico, attualmente rappresentante il principale mezzo di comunicazione di massa, che offre all'utente una vasta serie di contenuti potenzialmente informativi e servizi. Si tratta di un'interconnessione globale tra reti informatiche di natura ed estensione diversa, resa possibile da una *suite* di protocolli di rete comune chiamata "TCP/IP" dal nome dei due protocolli principali, il TCP e l'IP, che costituiscono la "lingua" comune con cui i computer connessi ad Internet (gli host) sono interconnessi e comunicano tra loro indipendentemente dalla loro architettura hardware e software, garantendo l'interoperabilità tra sistemi e sottoreti fisiche diverse. L'avvento e la diffusione di Internet e dei suoi servizi hanno rappresentato una vera e propria rivoluzione tecnologica e sociologica dagli inizi degli anni 90 (assieme ad altre invenzioni quali i telefoni cellulari e il GPS) nonché uno dei motori dello sviluppo economico mondiale nell'ambito dell'Information and Communication Technology (ICT) e oltre.

Descrizione

Costituita da alcune centinaia di milioni di computer collegati tra loro con i più svariati mezzi trasmissivi, Internet è anche la più grande rete di computer attualmente esistente, motivo per cui è definita "rete delle reti", "interete", "rete globale" o "la Rete" per antonomasia collegando tra loro a livello globale reti LAN, MAN e WAN.

In quanto rete di telecomunicazione (una rete di computer è un sottotipo per comunicazione dati di una rete di telecomunicazioni) come diffusione è seconda solo alla Rete Telefonica Generale, anch'essa di diffusione mondiale e ad accesso pubblico, ma comprendente il pianeta in maniera ancor più 'capillare' di Internet. Quest'ultima condivide largamente la rete telefonica per l'accesso e il trasporto dei suoi utenti ed è destinata, in un futuro non troppo lontano con il miglioramento della tecnologia VoIP, a soppiantarla inglobandola in sé in quanto basata sulla più efficiente tecnica della commutazione di pacchetto.

Internet offre i più svariati servizi, i principali dei quali sono il World Wide Web e la posta elettronica, ed è utilizzata per le comunicazioni più disparate: private e pubbliche, lavorative e ricreative, scientifiche e commerciali. I suoi utenti, in costante crescita, nel 2008 hanno raggiunto quota 1,5 miliardi e, visto l'attuale ritmo di crescita, si prevede che saliranno a 2,2 miliardi nel 2013.²⁹

Le caratteristiche della Rete Internet possono essere descritte attraverso la sua struttura fisica (topologia, nodi, collegamenti trasmissivi e apparati di rete), e attraverso il suo funzionamento a livello logico-protocollare.

²⁸Il termine "Internet", essendo un nome proprio, dovrebbe essere scritto con l'iniziale maiuscola. Vista la sua ormai ampia diffusione nel linguaggio comune è però accettata anche la grafia con l'iniziale minuscola. Per evitare fraintendimenti con l'omonimo nome comune è comunque preferibile la grafia con l'iniziale maiuscola.

²⁹Raffaella Natale. (2009) *Internet: 2,2 miliardi di utenti nel 2013. Asia e Cina guideranno la crescita mentre l'Italia risulta ancora tra i Paesi poco connessi*. Riportato il 10 ottobre 2009.

Struttura fisica

In generale a livello fisico la rete Internet può essere vista come una complessa interconnessione di nodi con funzionalità di ritrasmissione, appoggiata a collegamenti trasmissivi di vario tipo, sia cablati che wireless (fibre ottiche, cavi coassiali, doppini telefonici, cavi elettrici in posa anche in strutture idrauliche, collegamenti sottomarini, collegamenti satellitari, collegamenti a radiofrequenza (WiFi) e su ponti radio) che consentono l'interconnessione da estremo a estremo (*end to end*) di un agente umano o automatico a un altro agente, praticamente qualsiasi tipo di computer o elaboratore elettronico oggi esistente.

Ogni dispositivo terminale connesso direttamente a Internet si chiama *nodo ospite*, in inglese *host o end system* (sistema finale o terminale utente), mentre la struttura che collega i vari *host* si chiama *link di comunicazione* passando attraverso i vari nodi interni di commutazione. Da qualche anno è ormai possibile collegarsi a Internet anche da dispositivi mobili come palmari, telefoni cellulari, tablet ecc... In breve dovrebbe essere possibile per ciascuno di questi dispositivi non solo «accedere» a Internet, ma anche «subire l'accesso» da parte di altri *host* Internet.

In quanto "Rete delle reti" Internet non possiede dunque una topologia ben definita, ma varia e soprattutto di natura capillare fino agli utenti, com'è tipico anche della rete telefonica. La "ragnatela" di collegamenti è composta, a livello fisico-infrastrutturale, da un'ossatura molto veloce e potente, nota come rete di trasporto, a cui si connettono, attraverso collegamenti di backhauling (*raccordo*), molteplici sottoreti a volte più deboli e lente e che costituiscono quindi la rispettiva rete di accesso, com'è tipico in generale anche della rete telefonica, della quale Internet condivide proprio l'infrastruttura di accesso per la connessione delle utenze private.

I collegamenti tra i vari nodi interni si appoggiano su criteri statistici di disponibilità (moltiplicazione statistica) e non su criteri totalmente deterministici, a causa della natura distribuita piuttosto che centralizzata dei processi in rete.

Molti nodi interni sono collegati tra loro in diversi modi e tramite diversi *path*. Questo tipo di interconnessione può essere compreso alla luce delle motivazioni che negli anni sessanta dettarono la nascita di Internet (allora denominata ARPANET): creare una rete di elaboratori decentrata che potesse resistere a un attacco nucleare da parte dell'Unione Sovietica. Una tale rete decentrata sarebbe sopravvissuta a molti attacchi visto che un attacco a un singolo elaboratore non avrebbe impedito il funzionamento generale, e i collegamenti ridondanti avrebbero sostituito quelli distrutti.

Le sottoreti componenti possono anche essere protette e, quindi, consentono l'accesso a Internet (e viceversa) solo in maniera condizionata. Si tratta delle Intranet e la protezione è tipicamente realizzata attraverso l'uso di un *firewall* (*muro tagliafuoco* in inglese).

La velocità di connessione o velocità di trasmissione in una comunicazione *end to end* tra due terminali è in ogni caso limitata dalle prestazioni più basse, in termini di velocità di trasferimento, della sottorete o del collegamento geografico attraversato, che fungono quindi da classico collo di bottiglia, e/o da eventuali situazioni di congestione interna della rete.

Accesso

Per l'accesso alla rete Internet e la fruizione dei suoi molteplici servizi, chiunque disponga di un computer, di un modem e degli opportuni software oppure una rete locale da interconnettere attraverso un router, deve instaurare una connessione con l'Internet Service Provider (ISP), a seguito della stipulazione di un contratto di servizio, che gli fornisce un accesso attraverso una linea di telecomunicazione dedicata cablata o wireless (ADSL, HDSL, VDSL, GPRS, HSDPA, Wi-Fi, ecc.) o una linea telefonica della Rete Telefonica Generale (POTS, ISDN, GSM, UMTS ecc.) ed iniziare così la rispettiva sessione di navigazione attraverso l'utilizzo di un web browser.

Gli ISP sono connessi a loro volta a ISP di livello superiore che utilizzano *router* ad alta velocità e *link* fisici in fibra ottica nella rete di trasporto.

In molti Stati la possibilità di accesso ad Internet da parte dell'utente viene vista sotto l'obbligo di servizio universale al pari della connessione alla rete telefonica.

Funzionamento logico

Generalmente definita «*la Rete delle reti*» o semplicemente «*Rete*», Internet è infatti costituita da tutta una serie di reti, private, pubbliche, aziendali, universitarie, commerciali, interconnesse tra di loro. In effetti, già prima della sua nascita, esistevano reti locali, principalmente nei centri di ricerca internazionali e nei dipartimenti universitari, che operavano ciascuna secondo modalità o protocolli propri di comunicazione.



Il grande risultato della nascita e dell'affermazione di Internet è stata quindi la creazione di uno standard *de facto* tra i protocolli di comunicazione che, senza eliminare ovvero in aggiunta ai protocolli di rete locale, interoperasse e gestisse in maniera affidabile ad un livello logico superiore tutte le varie reti interagenti, consentendo ai più diversi enti e agenti (diversi governi, diverse società nazionali o sovranazionali, tra i vari dipartimenti universitari) di scambiarsi dati grazie a un protocollo comune, il TCP/IP, relativamente indipendente da specifiche hardware proprietarie, da sistemi operativi e dai formati dei linguaggi di comunicazione degli apparati di rete (*modem, repeater, hub, multiplexer, bridge, switch, router, gateway*).

Dal punto di vista trasmissivo e informativo ciò che viaggia in Internet sono i pacchetti dati, che costituiscono l'unità minima di informazione in questo vasto sistema di comunicazione. Tali pacchetti viaggiano nei link e nodi interni di rete usando una tecnica di commutazione nota come commutazione di pacchetto (*packet switching*) che consente di condividere più di un possibile cammino piuttosto che fare uso di un percorso unico dedicato come accade invece nella classica commutazione di circuito della rete telefonica. In pratica i pacchetti dati di una comunicazione che partono da un host e giungono ad un altro host non seguono percorsi di instradamento predefiniti, ma quelli più congeniali nel preciso momento di attraversamento in base alla disponibilità fisica di collegamento dei link della rete e/o alle condizioni di congestione della rete stessa. Di conseguenza i pacchetti di una stessa comunicazione possono seguire percorsi diversi verso lo stesso destinatario.

Per potersi collegare a Internet ed usufruire dei relativi servizi, il solo requisito logico-funzionale richiesto a un qualsiasi agente o dispositivo elettronico (tipicamente detto client) è quello di poter "dialogare" con il destinatario e i nodi interni di rete per mezzo di opportuni protocolli di rete che, nel caso in questione, fanno parte della cosiddetta suite di protocolli Internet, regolando opportunamente l'invio e la ricezione dei pacchetti informativi e implementando a livello software tutte le funzionalità richieste in una tipica architettura di rete a strati o livelli (*layer*).

I protocolli più importanti di tale suite, cioè quelli che garantiscono l'interoperabilità e il buon funzionamento tra le diverse sottoreti, sono il Transmission Control Protocol ("Protocollo di Controllo di trasmissione dati", TCP), l>User Datagram Protocol (UDP) e l'Internet Protocol ("Protocollo Internet", IP): il primo ha funzionalità di controllo di trasmissione, il secondo di inoltra semplice, il terzo ha funzionalità di indirizzamento/instradamento nei nodi interni di commutazione.

Come detto, la struttura della comunicazione è a strati (simile al modello di architettura ISO-OSI) in una pila o stack protocollare di 5 livelli dal livello fisico al livello applicativo: secondo tale struttura, il protocollo TCP o UDP occupa il livello superiore (livello di trasporto) rispetto a IP (livello di rete). Al di sopra di questi ci sono i protocolli di tipo applicativo connessi al particolare servizio richiesto da espletare, al di sotto ci sono i protocolli di trasporto tipici delle reti locali, metropolitane e geografiche da interconnettere, dei collegamenti di raccordo (*backhauling*) e di dorsale (*backbone*), altri ancora sono collocati al loro stesso livello.

In sostanza un pacchetto dati iniziale (payload) che parte a livello applicativo da un host attraversa verticalmente dall'alto al basso tutti i vari strati protocollari che aggiungono al pacchetto stesso, in una procedura detta di imbustamento, via via informazioni aggiuntive (*header*) in una struttura di informazioni di servizio (overhead); quando il pacchetto totale così ottenuto, una volta trasmesso a livello fisico sul mezzo trasmissivo, raggiunge la destinazione, ovvero viene ricevuto, avviene uno spaccettamento inverso dal basso verso l'alto e ogni livello legge e elabora le informazioni del rispettivo header.

Tale struttura logica di servizio si basa sugli *Internet Standard* sviluppati dall'Internet Engineering Task Force (IETF) con documenti rigorosamente approvati noti come Request for Comments ("Richiesta di commenti", RFC) e, a livello applicativo, dai protocolli del World Wide Web Consortium (W3C).

In particolare, dal punto di vista della fruizione di servizi di livello applicativo, l'architettura logica della rete Internet può essere di tipo client-server oppure peer-to-peer. Ciascun terminale o host di rete, per la sua raggiungibilità, è inoltre identificato da un indirizzo IP statico o dinamico (cioè assegnato manualmente dall'amministratore di rete o dal DHCP), mentre le risorse da fruire sono tipicamente presenti sui server, raggiungibili dal client sotto l'indicazione mnemonica fornita dal cosiddetto URL, grazie all'utilizzo di un web browser e dei suddetti protocolli di rete forniti/implementati nel sistema operativo della macchina terminale in uso.

La conversione da indirizzo mnemonico URL a indirizzo IP, necessaria per l'instradamento effettivo in rete nei nodi di commutazione, è fornita dal cosiddetto DNS, mentre la comunicazione tra client e server si instaura in seguito alla definizione dei cosiddetti socket in cui oltre agli indirizzi IP di client e server vengono specificate anche le cosiddette porte coinvolte nel servizio di comunicazione da espletare.

Le modalità di utilizzo di Internet differiscono a seconda del tipo di servizio che si richiede e al tipo di server a cui ci si collega; per citarne solo alcune:

- Posta elettronica (e-mail)

consente di inviare e ricevere (a/da utenti) messaggi contenenti testo e altri formati (es.: immagini, video, audio). La modalità di funzionamento dei *server* di posta elettronica e di molti programmi *client* viene detta *store-and-forward*;

- File Transfer Protocol ("Protocollo di trasferimento dati", FTP)

consente di inviare e ricevere (a/da sistemi) file, cioè insiemi di informazioni codificate in maniera binaria (es.: testi, immagini, filmati, programmi, ecc.);

- Hyper Text Transfer Protocol ("Protocollo di trasferimento ipertesti", HTTP)

consente di organizzare le informazioni e le risorse presenti in rete in maniera non-sequenziale (Collegamento ipertestuale), come meglio descritto in seguito.

Nata come rete di comunicazione dati tra utenti è diventata in seguito anche una rete per trasferimento di dati vocali grazie allo sviluppo della tecnologia VOIP.

Storia di Internet

L'origine di Internet risale agli anni sessanta, su iniziativa degli Stati Uniti, che misero a punto durante la guerra fredda un nuovo sistema di difesa e di controspionaggio.

La prima pubblicazione scientifica in cui si teorizza una rete di computer mondiale ad accesso pubblico è *On-line man computer communication* dell'agosto 1962, pubblicazione scientifica degli statunitensi Joseph C.R. Licklider e Welden E. Clark. Nella pubblicazione Licklider e Clark, ricercatori del Massachusetts Institute of Technology, danno anche un nome alla rete da loro teorizzata: "Intergalactic Computer Network".

Prima che tutto ciò cominci a diventare una realtà pubblica occorrerà attendere il 1991 quando il governo degli Stati Uniti d'America emana la *High performance computing act*, la legge con cui per la prima volta viene prevista la possibilità di ampliare, per opera dell'iniziativa privata e con finalità di sfruttamento commerciale, una rete Internet fino a quel momento rete di computer mondiale di proprietà statale e destinata al mondo scientifico. Questo sfruttamento commerciale viene subito messo in atto anche dagli altri Paesi.

ARPANET (1969)

Il progenitore e precursore della rete Internet è considerato il progetto ARPANET, finanziato dalla *Defence Advanced Research Projects Agency* (inglese: DARPA, Agenzia per i Progetti di ricerca avanzata per la Difesa), una agenzia dipendente dal Ministero della Difesa statunitense (*Department of Defense* o DoD degli Stati Uniti d'America). In una nota del 25 aprile 1963, Joseph C.R. Licklider aveva espresso l'intenzione di collegare tutti i computer e i sistemi di *time-sharing* in una rete continentale. Avendo lasciato l'ARPA per un posto all'IBM l'anno seguente, furono i suoi successori che si dedicarono al progetto ARPANET.

Il contratto fu assegnato all'azienda da cui proveniva Licklider, la Bolt, Beranek and Newman (BBN) che utilizzò i minicomputer di Honeywell come supporto. La rete venne fisicamente costruita nel 1969 collegando quattro nodi: l'Università della California di Los Angeles, l'SRI di Stanford, l'Università della California di Santa Barbara, e l'Università dello Utah. L'ampiezza di banda era di 50 Kbps. Negli incontri per definire le caratteristiche della rete, vennero introdotti i fondamentali *Request for Comments*, tuttora i documenti fondamentali per tutto ciò che riguarda i protocolli informatici della rete e i loro sviluppi. La super-rete dei giorni nostri è risultata dall'estensione di questa prima rete, creata sotto il nome di ARPANET.

I primi nodi si basavano su un'architettura *client/server*, e non supportavano quindi connessioni dirette (*host-to-host*). Le applicazioni eseguite erano fondamentalmente Telnet e i programmi di File Transfer Protocol (FTP). Il servizio di posta elettronica fu inventata da Ray Tomlinson della BBN nel 1971, derivando il programma da altri due: il SENDMSG per messaggi interni e CPYNET, un programma per il trasferimento dei file. L'anno seguente Arpanet venne presentata al pubblico, e Tomlinson adattò il suo programma per funzionarvi: divenne subito popolare, grazie anche al contributo di Larry Roberts che aveva sviluppato il primo programma per la gestione della posta elettronica, RD.

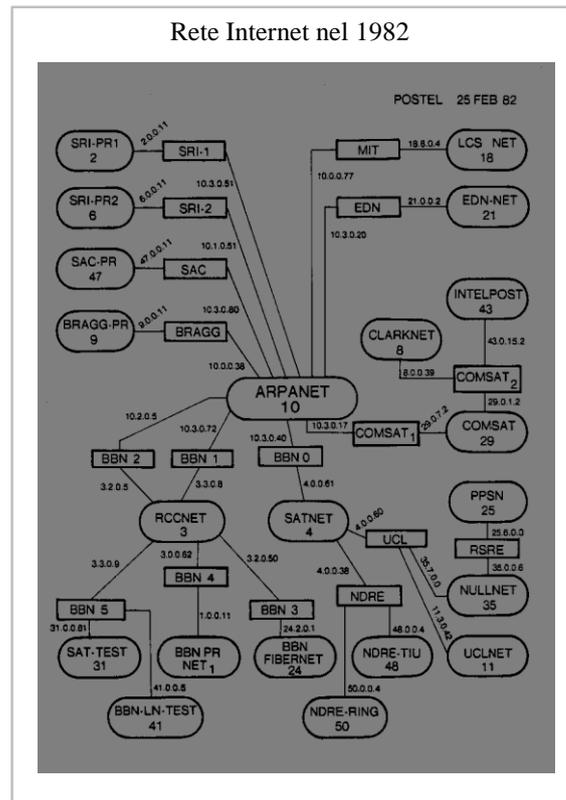
Da Arpanet a Internet (anni settanta e ottanta)

In pochi anni, ARPANET allargò i suoi nodi oltreoceano, contemporaneamente all'avvento del primo servizio di invio pacchetti a pagamento: Telenet della BBN. In Francia inizia la costruzione della rete CYCLADES sotto la direzione di Louis Pouzin, mentre la rete norvegese NORSAR permette il collegamento di Arpanet con lo *University College* di Londra. L'espansione proseguì sempre più rapidamente, tanto che il 26 marzo del 1976 la regina Elisabetta II d'Inghilterra spedì un'email alla sede del *Royal Signals and Radar Establishment*.

Gli *Emoticon* vennero istituiti il 12 aprile 1979, quando Kevin MacKenzie suggerì di inserire un simbolo nelle mail per indicare gli stati d'animo.

Tutto era pronto per il cruciale passaggio a Internet, compreso il primo virus telematico: sperimentando sulla velocità di propagazione delle e-mail, a causa di un errore negli header del messaggio, Arpanet venne totalmente bloccata: era il 27 ottobre

1980. Wikipedia: Cita le fonti Definendo il *Transmission Control Protocol (TCP)* e l'*Internet Protocol (IP)*, DCA e ARPA diedero il via ufficialmente a Internet come l'insieme di reti interconnesse tramite questi protocolli.



Internet in Italia

L'Italia fu il terzo Paese in Europa a connettersi in rete, dopo Norvegia e Inghilterra, grazie ai finanziamenti del Dipartimento della Difesa degli Stati Uniti. La connessione avvenne dall'Università di Pisa, dove era presente un gruppo di ricerca fra i più avanzati in Europa. Alcuni dei componenti del gruppo avevano lavorato a contatto con quelli che poi sarebbero stati considerati i padri di Internet, Robert Kahn e Vinton Cerf. Fu proprio Kahn a convincere i suoi superiori a finanziare l'acquisto delle tecnologie necessarie (Butterfly Gateway) per il gruppo di Pisa. Il collegamento avvenne il 30 aprile 1986.³⁰

Nascita del World Wide Web (1991)

Nel 1991 presso il CERN di Ginevra il ricercatore Tim Berners-Lee definì il protocollo HTTP (*HyperText Transfer Protocol*), un sistema che permette una lettura ipertestuale, *non-sequenziale* dei documenti, saltando da un punto all'altro mediante l'utilizzo di rimandi (link o, più propriamente, hyperlink). Il primo browser con caratteristiche simili a quelle attuali, il Mosaic, venne realizzato nel 1993. Esso rivoluzionò profondamente il modo di effettuare le ricerche e di comunicare in rete. Nacque così il World Wide Web.

Nel World Wide Web (WWW), le risorse disponibili sono organizzate secondo un sistema di librerie, o pagine, a cui si può accedere utilizzando appositi programmi detti *web browser* con cui è possibile *navigare* visualizzando file, testi, ipertesti, suoni, immagini, animazioni, filmati.

³⁰Intervista su Repubblica a Luciano Lenzini, docente dell'Infn di Pisa

Nel 1993 il CERN, l'istituzione europea dove nasce il World Wide Web, decide di rendere pubblica la tecnologia alla base del World Wide Web in modo che sia liberamente implementabile da chiunque. A questa decisione fa seguito un immediato e ampio successo del World Wide Web in ragione delle funzionalità offerte, della sua efficienza e, non ultima, della sua facilità di utilizzo. Da tale successo ha inizio la crescita esponenziale di Internet che in pochissimi anni porterà la rete delle reti a cambiare la società umana rivoluzionando il modo di relazionarsi delle persone come quello di lavorare, tanto che nel 1998 si arriverà a parlare di "nuova economia".

La facilità d'utilizzo connessa con l'HTTP e i browser, in coincidenza con una vasta diffusione di computer per uso anche personale (vedi Storia del Personal Computer), hanno aperto l'uso di Internet a una massa di milioni di persone, anche al di fuori dell'ambito strettamente informatico, con una crescita in progressione esponenziale.

Diffusione

Se prima del 1995 Internet era dunque relegata ad essere una rete dedicata alle comunicazioni all'interno della comunità scientifica e tra le associazioni governative e amministrative, dopo tale anno si assiste alla diffusione costante di accessi alla rete da parte di computer di utenti privati fino al boom degli anni 2000 con centinaia di milioni di computer connessi in rete in parallelo alla diffusione sempre più spinta di PC al mondo, all'aumento dei contenuti e servizi offerti dal Web e a modalità di navigazione sempre più usabili, accessibili e user-friendly nonché a velocità di trasferimento dati a più alta velocità di trasmissione passando dalle connessioni ISDN e V.90 alle attuali e ben note connessioni a banda larga tramite sistemi DSL.

Principali servizi disponibili su Internet

- World Wide Web
- Motori di ricerca
- File Transfer Protocol (FTP)
- Telnet
- Hosting
- Gopher
- E-Mail
- Mailing list
- Newsgroup
- Podcast
- Chat
- Forum
- Blog
- Social network
- File sharing
- Video sharing
- Webcast
- Streaming o Web Tv
- IPTV
- Web Radio

- VoIP
- Bitcoin
- E-Commerce
- E-Learning
- E-Government
- E-health
- Home banking
- Multiplayer

Protocollo di rete

Nelle telecomunicazioni, per **protocollo di rete** (o anche *protocollo di comunicazione*) si intende la definizione formale a priori delle modalità di interazione che, due o più apparecchiature elettroniche collegate tra loro, devono rispettare per operare particolari funzionalità di elaborazione necessarie all'espletamento di un certo servizio di rete. In termini equivalenti si tratta dunque della descrizione a livello logico del processo di comunicazione (meccanismi, regole o schema di comunicazione) tra terminali e apparati preposti al funzionamento efficace della comunicazione in rete. Queste apparecchiature possono essere host, computer clienti, telefoni intelligenti, Personal Digital Assistant (PDA), monitor, stampanti, sensori ecc.

In particolare un protocollo prevede la definizione dei linguaggi costituiti dai messaggi scambiati, messaggi che devono potersi interpretare correttamente. L'aderenza ai protocolli garantisce che due software in esecuzione su diverse macchine possano comunicare efficacemente, anche se sono stati realizzati indipendentemente cioè interoperabilità. È evidente l'importanza della standardizzazione dei protocolli di rete. L'implementazione informatica dei protocolli di rete definisce, all'interno dell'architettura di rete, il cosiddetto software di rete, presente usualmente all'interno del sistema operativo ed elaborato dalla scheda di rete.

In senso più lato, un protocollo di comunicazione si può definire come un insieme di regole che vengono stabilite per instaurare una comunicazione corretta: ad esempio un italiano e un cinese per comunicare potrebbero mettersi d'accordo nell'utilizzo della lingua inglese.

Indirizzo IP

Un **indirizzo IP** (dall'inglese **Internet Protocol address**) è un'etichetta numerica che identifica univocamente un dispositivo (host) collegato a una rete informatica che utilizza l'Internet Protocol come protocollo di comunicazione. Un indirizzo IP assolve essenzialmente a due funzioni principali: identificare un dispositivo sulla rete e di conseguenza fornirne il percorso per la sua raggiungibilità da un altro terminale o dispositivo di rete in una comunicazione dati a pacchetto.

Più esattamente l'indirizzo IP viene assegnato a una *interfaccia* (ad esempio una scheda di rete) che identifica l'host di rete, che può essere un personal computer, un palmare, un router, elettrodomestici in generale, ecc. Va considerato, infatti, che un host può contenere più di una interfaccia: ad esempio, un router ha diverse interfacce (minimo due) per ognuna delle quali occorre un indirizzo IP.

L'indirizzo IP è un indirizzo di livello 3 del modello ISO-OSI che si rende necessario per l'*instradamento indiretto* *Wikipedia:Cita le fonti* tramite l'omonimo protocollo di rete IP ovvero Aiuto:Chiarezza per interconnettere (*internetworking*) più sottoreti diverse all'interno del paradigma TCP-IP della rete Internet. Si contrappone all'indirizzo fisico di livello 2 o indirizzo MAC utilizzato invece per l'*instradamento diretto* all'interno di una sottorete locale.

Gli indirizzi IP

La distribuzione mondiale degli indirizzi IP

Gli indirizzi IP pubblici e i range di indirizzi sono rilasciati e regolamentati dall'ICANN tramite una serie di organizzazioni delegate³¹. A causa della saturazione dello spazio di IPv4 l'ICANN per ovviare a tale disagio ha proceduto alla definizione della nuova versione del protocollo IP: IPv6³². Tale versione è basata su indirizzi a 128 bit anziché a 32 e ciò permetterà l'assegnazione di un numero decisamente maggiore di indirizzi.

Ma la difficile implementazione a livello globale dell'IPv6 ha portato all'introduzione di nuovi concetti che hanno rivoluzionato la teoria delle reti. Vanno citati l'abbandono del concetto di classi di indirizzi IP e il conseguente utilizzo sempre maggiore di indirizzi classless (privi del concetto di classe)³³, la subnet mask, la riorganizzazione gerarchica degli indirizzi mediante utilizzo massivo di Network address translation (NAT).

Composizione e formati

Ogni indirizzo IP è suddiviso in due parti:

- La prima parte che identifica la rete, chiamata *network* o *routing prefix* (Net_ID) ed è utilizzato per l'instradamento a livello di sottoreti.
- La seconda parte che identifica l'host (o l'interfaccia in IPv6) e le eventuali sottoreti (Host_ID) ed è utilizzato per l'instradamento a livello locale dell'host una volta raggiunta la sottorete locale di destinazione, cui segue la traduzione o risoluzione in indirizzo MAC per l'effettiva consegna del pacchetto dati al destinatario con i protocolli della rete locale.

Grazie all'introduzione del CIDR è possibile creare delle sottoreti, a differenza del sistema a classi, che prevedeva il rigoroso rispetto delle reti imposte con conseguente spreco di indirizzi IP³⁴.

IPv4

L'indirizzo IPv4 è costituito da 32 bit (4 byte) suddiviso in 4 gruppi da 8 bit (1 byte), separati ciascuno da un punto (notazione *dotted*) (es. 11001001.00100100.10101111.00001111). Ciascuno di questi 4 byte è poi convertito in formato decimale di più facile identificazione (quindi ogni numero varia tra 0 e 255 essendo $2^8=256$). Un esempio di indirizzo IPv4 è 195.24.65.215.

IPv6

L'indirizzo IPv6

³¹La struttura dell'ICANN

³²Annuncio passaggio IPV6

³³Il CIDR

³⁴Il CIDR

L'indirizzo IPv6 è costituito da 128 bit (16 byte), viene descritto da 8 gruppi di 4 numeri esadecimali che rappresentano 2 byte ciascuno (quindi ogni numero varia tra 0 e 65535) separati dal simbolo "due punti". Un esempio di indirizzo IPv6 è 2001:0DB8:0000:0000:0000:0000:0000:0001, che può essere abbreviato in 2001:DB8::1 (i due punti doppi rappresentano la parte dell'indirizzo che è composta di soli zeri consecutivi. Si può usare una sola volta, per cui se un indirizzo ha due parti composte di zeri la più breve andrà scritta per esteso).

I dispositivi connessi ad una rete IPv6 ottengono un indirizzo di tipo unicast globale, vale a dire che i primi 48 bit del suo indirizzo sono assegnati alla rete a cui esso si connette, mentre i successivi 16 bit identificano le varie sottoreti a cui l'host è connesso. Gli ultimi 64 bit sono ottenuti dall'indirizzo MAC dell'interfaccia fisica.

Assegnazione degli indirizzi IP

La parte Net_ID degli indirizzi è assegnata dall'ICANN mentre l'assegnazione della parte Host_ID è delegata al richiedente che eventualmente può suddividerla ulteriormente per la creazione di altre sottoreti logiche (subnetting) evitando duplicazioni e sprechi di indirizzi.

Gli indirizzi IP possono essere assegnati in maniera permanente (per esempio un server che si trova sempre allo stesso indirizzo) oppure in maniera temporanea, da un intervallo di indirizzi disponibili.

In particolare l'assegnazione dell'Host_Id può essere di due tipi: dinamica oppure statica.

Indirizzi dinamici

Gli indirizzi dinamici vengono utilizzati per identificare dispositivi non permanenti in una LAN. Un server presente nella LAN assegna dinamicamente e automaticamente l'indirizzo scegliendolo casualmente da un range preimpostato. Si può scegliere l'intervallo di indirizzi a seconda del numero delle utenze della rete impostando la netmask, ossia dicendo al server DHCP quanti bit dell'indirizzo sono assegnabili dinamicamente a ogni singolo client che fa accesso. Per esempio, se la netmask ha valore 255.255.255.0 (dove ogni blocco separato da puntini denota un gruppo di 8 bit) solo gli ultimi 8 bit sono assegnabili agli host.

Gli Internet Service Provider (ISP), per esempio, utilizzano un numero di indirizzi assegnabili ristretto per una vasta clientela facendo leva sul concetto che non tutti i client saranno connessi nello stesso momento. Wikipedia: Cita le fonti Questo sistema viene utilizzato soprattutto per gli accessi dial-up, Wi-fi o in qualsiasi altro accesso temporaneo permettendo, per esempio a un portatile, di connettersi a un'ampia varietà di servizi senza la necessità di dover conoscere i dettagli di indirizzamento di ogni rete.

Indirizzi statici

Gli indirizzi statici vengono utilizzati per identificare dispositivi semi-permanenti con indirizzo IP permanente. Server, stampanti di rete ecc... utilizzano tipicamente questo metodo di indirizzamento³⁵. Tipicamente si può ricorrere anche ad un'assegnazione statica anziché dinamica per dispositivi di rete non permanenti se il numero di host della sottorete è contenuto e/o per motivi di sicurezza potendo tenere sotto controllo le azioni di ciascun host e del relativo utente. Sul fronte della sicurezza informatica l'assegnazione di un IP statico rende però il dispositivo più soggetto ad attacchi informatici.

L'indirizzo statico può essere configurato direttamente sul dispositivo, oppure come parte di una configurazione DHCP che associa all'Indirizzo MAC il corrispondente indirizzo IP statico.

³⁵L'utilizzo degli indirizzi statici

I servizi correlati

Sistemi di risoluzione dei nomi

Per rendere maggiormente user-friendly la tecnologia IP sono stati implementati alcuni servizi che associano un nome leggibile, e più semplice da ricordare, a un indirizzo ip.

DNS (Domain Name System)

Il DNS è un servizio di directory utilizzato per la risoluzione dei nomi dei server da indirizzi logici e testuali (URL) in indirizzi IP. Questa funzione è essenziale per l'usabilità di Internet, visto che gli esseri umani hanno più facilità a ricordare nomi testuali, mentre i dispositivi di instradamento (interfacce di rete e router di livello 2 e superiore) lavorano su indirizzi binari. Permette inoltre ad una qualsiasi entità di cambiare o riassegnare il proprio indirizzo IP, senza dover notificare tale cambiamento a nessuno, tranne che al proprio server DNS di riferimento.

Un'altra delle peculiarità del DNS è quella di consentire, ad esempio ad un sito web, di essere ospitato su più server (ognuno con il proprio indirizzo IP), con una conseguente divisione del carico di lavoro.

Server

In informatica il termine *server* (dall'inglese, letteralmente *servitore*), indica genericamente un componente o sottosistema informatico che fornisce, a livello logico e a livello fisico, un qualunque tipo di servizio ad altre componenti (tipicamente chiamate *client*, cioè "cliente") attraverso una rete di computer, all'interno di un sistema informatico o direttamente in locale su un computer. Rappresenta dunque un nodo terminale della rete opposto all'host client. In parole semplici si tratta di un Computer o di un programma che fornisce i dati richiesti da altri elaboratori, facendo quindi da Host per la trasmissione delle informazioni virtuali.

Al termine *server*, così come per il termine *client*, possono dunque riferirsi sia la componente hardware che la componente software che forniscono le funzionalità o servizi di cui sopra. La macchina hardware server si dice anche capace di offrire risorse in termini di servizio al client che ne fa domanda/richiesta. Esso fa parte dunque dell'architettura logica di rete a livello applicativo detta client-server.



Descrizione

In base al contesto, il termine *server* può indicare quindi:

1. un computer utilizzato per fornire servizi ad altri computer, a prescindere dalle sue caratteristiche hardware.
2. un computer specifico appartenente alla fascia di mercato dedicata all'uso come server, caratterizzato da alta affidabilità e maggiori prestazioni.

3. un processo (ovvero un programma in esecuzione) che fornisca servizi ad altri processi (es. Server Web).

In generale quindi si tende ad indicare come *server* il componente hardware che attraverso un software specifico eroga un servizio e *client* il componente hardware con relativo software che ne fa uso su richiesta dell'utente.

Caratteristiche

Tipicamente una macchina hardware server può ospitare uno o più servizi ovvero implementare logicamente uno o più server. Ciascun server può essere di tipo:

- *iterativo*, cioè accogliere e soddisfare sequenzialmente una sola richiesta di servizio alla volta da parte dei client con una tipica procedura a coda di attesa dei processi da gestire (caso di pochi client: situazione ai primordi della rete Internet);
- *concorrente*, ovvero soddisfare più richieste di servizio da parte di più client attraverso procedure tipiche del multithreading (situazione diventata ora comune nella rete Internet).

Funzionamento

I server offrono servizi al client tipicamente in modalità *richiesta/risposta* lavorando a livello applicativo del modello ISO-OSI e del modello TCP/IP (come anche i client) secondo i dettami del particolare protocollo di rete che regola il suo funzionamento per l'espletamento del servizio al client. Si appoggiano dunque agli strati protocollari inferiori dello stack di rete, in particolare sui protocolli di livello di trasporto quali TCP e UDP e sul livello di rete (IP) per l'instradamento.

Tipicamente l'espletamento del servizio per il client è preceduta da una fase di definizione di un socket (coppia indirizzo IP/porta) e successiva instaurazione della connessione con il server tramite TCP o UDP, con il server che possiede un indirizzo IP statico a causa dei legami intrinseci con il nome di dominio (che è inevitabilmente fisso) attraverso il DNS. Le porte dedicate ai server sono tipicamente porte note e statiche (*well known ports*), dipendenti dal particolare servizio espletato.

La creazione di applicazioni di rete, per quanto riguarda la connessione, ricade all'interno degli ambiti della cosiddetta programmazione socket. Tipicamente tale ambito di programmazione fa uso di opportune chiamate di sistema o API Socket al sistema operativo del server e del client per realizzare la connessione affidabile sia in modalità iterativa sia concorrente gestendo anche tutti i possibili errori o eccezioni.

Architettura client-server o N-tier

Il complesso logico formato da uno o più server e, solitamente, svariati client è denominato *Modello client-server* tipico delle Rete Internet e dei sistemi informatici aziendali. Negli ultimi anni è stato esteso alla interazione tra più macchine o servizi eterogenei, e in tal caso si parla di *Architettura 3-tier* (o più in generale ad *N-tier*).

Il server riceve dai client delle richieste di servizio, e restituisce loro la risposta con i risultati dopo averne effettuato l'elaborazione relativa. Questo ha l'indubbio vantaggio di concentrare su una sola macchina centrale tutte le principali risorse di elaborazione, il software, la manutenzione, le informazioni critiche o sensibili, gli accorgimenti atti a garantire affidabilità (come i backup); I client, al contrario, possono in generale essere anche macchine con risorse e affidabilità inferiori, che hanno l'unico compito di interagire con l'utente, e che non contengono informazioni critiche o sensibili.

Si noti che dal punto di vista software un computer non è mai puramente un server, in quanto oltre ad erogare servizi ad altri si appoggia solitamente a servizi forniti da terzi, e quindi è contemporaneamente server e client. Ad esempio, un server di posta elettronica si appoggia ad altri server (es. *Server DNS*) per ottenere informazioni sulla rete, e quindi è simultaneamente sia un server di e-mail sia un client per quanto riguarda altri servizi relativi alla gestione della rete

Quando invece non si è in questa situazione ovvero ciascun terminale di rete è in una gerarchia paritaria con gli altri si parla di modello o architettura di rete peer-to-peer.

Tipicamente più server diversi possono risiedere anche sulla stessa macchina hardware con risparmio sui costi totali, previo opportuno dimensionamento per soddisfare le maggiori richieste dei client, a prezzo di una maggiore vulnerabilità dovuta alla centralizzazione delle risorse. Dal punto di vista hardware i server possono risiedere su un normale PC oppure su macchine hardware dedicate con prestazioni, affidabilità e costi superiori.

Gestione

Tipicamente i server in ambienti distribuiti come sistemi informatici e reti informatiche sono gestiti da amministratori di sistema da remoto (*telegestione*) attraverso connessioni remote realizzate attraverso varie possibili applicazioni dedicate (Virtual Network Computing).

Servizi e tipologie di server

Tra i servizi che vengono tipicamente erogati da un server o più server e che ne identificano a sua volta le varie tipologie, si possono citare:

- File server, permette agli utenti di accedere ai file situati sul server come se fossero sul proprio calcolatore, agevolando la condivisione di informazioni;
- Database server, permette di gestire intere banche dati;
- FTP server, fornisce alla rete accesso a cartelle pubbliche o con autenticazione;
- Web server, usato per ospitare un sito web (es. server HTTP);
- Application server, usato per far funzionare un programma applicativo sul Web (applicazione Web) e condividerne le funzionalità tra gli utenti;
- Mail server, usato per la gestione della posta elettronica;
- Print server, permette di mettere in comune una o più stampanti tra gli utenti di una rete con la eventuale gestione dei diritti di accesso;
- Gestione di una rete informatica, locale o geografica;
 - DHCP server, per l'assegnazione automatica di indirizzi IP ai computer host;
 - DNS server, fornisce la risoluzione dei nomi di dominio dei siti (per esempio it.wikipedia.org) nei loro indirizzi IP;
- Server grafico o *display server* – vedere X Window System;
- Server di autenticazione, permette di autenticare l'accesso ad altre risorse;
- VNC server, fornisce, attraverso un client VNC un desktop remoto;
- Proxy server, fornisce una cache di accesso al Web e la possibilità di controlli di autenticazione (ACL) e di filtro;
- Game server, ospita risorse per rendere possibili i giochi multiutente in linea;
- VPN server, fornisce attraverso un client VPN, l'accesso alla LAN da remoto.

Client

In informatica, con **client** (in italiano detto anche *cliente*) si indica una componente che accede ai servizi o alle risorse di un'altra componente detta server. In questo contesto si può quindi parlare di client riferendosi all'hardware oppure al software. Esso fa parte dunque dell'architettura logica di rete detta client-server.

Descrizione

Un computer collegato ad un server tramite una rete informatica (locale o geografica) ed al quale richiede uno o più servizi, utilizzando uno o più protocolli di rete è un esempio di client hardware.

Il termine client indica anche il software usato sul computer client per accedere alle funzionalità offerte dal server.

Sono sempre di più i software, come il web, l'e-mail, i database, che sono divisi in una parte client (residente ed in esecuzione sul pc client) ed una parte server (residente ed in esecuzione sul server). Un programma di posta elettronica è un esempio di client software.

Ad esempio, nel web il software client è il browser che comunica con il server web attraverso il protocollo HTTP; per l'e-mail il client è detto in gergo *mail user agent* o MUA (ad esempio, Outlook, Mozilla Thunderbird, Eudora, ...) che comunica con il server attraverso il protocollo SMTP e POP o IMAP; il client per la consultazione o la modifica del database (spesso costituito da librerie software utilizzate da una applicazione) comunica con il DBMS, che gestisce il database e risponde alle interrogazioni del client.

Transmission Control Protocol

In telecomunicazioni e informatica il **Transmission Control Protocol (TCP)**, anche chiamato **Transfer Control Protocol**, è un protocollo di rete a pacchetto di livello di trasporto, appartenente alla suite di protocolli Internet, che si occupa di *controllo di trasmissione*. È definito nella RFC 793 e su di esso si appoggiano gran parte delle applicazioni della rete Internet. È presente solo sui terminali di rete (host) e non sui nodi interni di commutazione della rete di trasporto, implementato all'interno del rispettivo sistema operativo e vi si accede automaticamente dal browser attraverso l'uso di opportune chiamate di sistema definite nelle API di sistema.

Descrizione

Il TCP può essere classificato al livello trasporto (OSI level 4) del modello di riferimento OSI, e di solito è usato in combinazione con il protocollo di livello rete (OSI level 3) IP. La corrispondenza con il modello OSI non è perfetta, in quanto il TCP e l'IP nascono prima del suddetto modello. La loro combinazione è indicata come TCP/IP e, alle volte, è erroneamente considerata un unico protocollo. Da qui, la difficoltà di una classificazione univoca per un protocollo che comprende, a pieno titolo, due livelli dello stack OSI (o *pila ISO/OSI* in italiano).

In linea con i dettami del livello di trasporto stabiliti dal modello ISO-OSI e con l'intento di superare il problema della mancanza di affidabilità e controllo della comunicazione sorto con l'interconnessione su vasta scala di reti locali in un'unica grande rete geografica, TCP è stato progettato e realizzato per utilizzare i servizi offerti dai protocolli di rete di livello inferiore (IP e protocolli di livello fisico e livello datalink) che definiscono efficacemente il modo di trasferimento sul canale di comunicazione, ma che non offrono alcuna garanzia di affidabilità sulla consegna in termini di ritardo, perdita ed errore dei pacchetti informativi trasmessi, sul controllo di flusso tra terminali e sul controllo della congestione di rete, supplendo quindi ai problemi di cui sopra e costruendo così un **canale di comunicazione affidabile tra due processi applicativi di rete**. Il canale di comunicazione così costruito è costituito da un flusso bidirezionale di byte a seguito dell'instaurazione di una connessione agli estremi tra i due terminali in comunicazione. Inoltre alcune funzionalità di TCP sono vitali per il buon funzionamento complessivo di una rete IP. Sotto questo punto di vista TCP può essere considerato come un elemento di rete che si occupa di garantire una qualità di servizio minima su una rete IP sotto che è di tipo *best-effort*.

Il TCP nacque nel 1970 come frutto del lavoro di un gruppo di ricerca del dipartimento di difesa statunitense. I suoi punti di forza sono l'alta affidabilità e robustezza. La sua popolarità si deve anche grazie ad una sua implementazione diffusa dalla Università di Berkeley, rilasciata in California sotto forma di sorgenti (*TCP Berkeley*). Molte tuttavia sono le implementazioni e sviluppi che si sono succedute nel tempo come evoluzioni e miglioramenti (es. *TCP Tahoe*, *TCP Reno*, *TCP New Reno*).

Caratteristiche principali

- TCP è un protocollo **orientato alla connessione**, ovvero prima di poter trasmettere dati deve stabilire la comunicazione, negoziando una connessione tra mittente e destinatario, che rimane attiva anche in assenza di scambio di dati e viene esplicitamente chiusa quando non più necessaria. Esso quindi possiede le funzionalità per creare, mantenere e chiudere/abbattere una connessione.
- Il servizio offerto da TCP è il trasporto di un **flusso di byte** bidirezionale tra due applicazioni in esecuzione su host differenti. Il protocollo permette alle due applicazioni di trasmettere contemporaneamente nelle due direzioni, quindi il servizio può essere considerato "**Full-Duplex**" anche se non tutti i protocolli applicativi basati su TCP utilizzano questa possibilità.
- Il *flusso di byte* viene **frazionato** in blocchi per la trasmissione dall'applicazione a TCP (che normalmente è implementato all'interno del sistema operativo), per la trasmissione all'interno di *segmenti* TCP, per la consegna all'applicazione che lo riceve, ma questa divisione in blocchi non è necessariamente la stessa nei diversi passaggi.
- TCP garantisce che i dati trasmessi, se giungono a destinazione, lo facciano in ordine e una volta sola ("**at most once**"). Più formalmente, il protocollo fornisce ai livelli superiori un servizio equivalente ad una connessione fisica diretta che trasporta un flusso di byte. Questo è realizzato attraverso vari meccanismi di acknowledgment e di ritrasmissione su timeout.
- TCP offre funzionalità di controllo di errore sui pacchetti pervenuti grazie al campo checksum contenuto nella sua PDU.
- TCP possiede funzionalità di controllo di flusso tra terminali in comunicazione e controllo della congestione sulla connessione, attraverso il meccanismo della finestra scorrevole. Questo permette di ottimizzare l'utilizzo della rete anche in caso di congestione, e di condividere equamente la capacità disponibile tra diverse sessioni TCP attive su un collegamento.
- TCP fornisce un servizio di moltiplicazione delle connessioni su un host, attraverso il meccanismo delle porte.

World Wide Web

« Il World Wide Web ha le potenzialità per svilupparsi in un'enciclopedia universale che copra tutti i campi della conoscenza e in una biblioteca completa di corsi per la formazione.³⁶ »

(Richard Stallman)

Il **World Wide Web**³⁷ (ellissi diffusa nel linguaggio comune: "Web";³⁸ abbreviazione diffusa nel linguaggio comune: "WWW";³⁹ appellativo diffuso nel linguaggio comune: "Grande Ragnatela Mondiale")⁴⁰ è un servizio di Internet⁴¹ che permette di navigare ed usufruire di un insieme vastissimo di contenuti (multimediali e non) collegati tra loro attraverso legami (link), e di ulteriori servizi accessibili a tutti o ad una parte selezionata degli utenti di Internet.



Immagine dell'home page del primo sito web della storia

Caratteristiche generali

Caratteristica principale della rete *Web* è che i nodi che la compongono sono tra loro collegati tramite i cosiddetti *link* (collegamenti), formando un enorme ipertesto. E i suoi servizi possono essere resi disponibili dagli stessi utenti di Internet.⁴² Per quanto riguarda i contenuti, quindi, il *Web* possiede la straordinaria peculiarità di offrire a chiunque la possibilità di diventare editore e, con una spesa estremamente esigua, di raggiungere un pubblico potenzialmente vastissimo distribuito in tutto il mondo.⁴³

³⁶L'enciclopedia universale libera e le risorse per l'apprendimento, Richard Stallman.

³⁷Nel linguaggio comune "Web", per la sua maggiore brevità, è quasi sempre preferito a "World Wide Web".

³⁸"Web", essendo un nome proprio, andrebbe scritto con l'iniziale maiuscola ma, come tutti i nomi propri non di persona molto diffusi nel linguaggio comune, viene anche scritto con l'iniziale minuscola. In particolare, quando è utilizzato in funzione aggettivale ("pagina web", "browser web", "server web", ecc.), viene prevalentemente scritto con l'iniziale minuscola.

³⁹"WWW" è l'acronimo di "World Wide Web". Per la scarsa compatibilità delle possibili pronunce dell'acronimo "WWW" con la fonologia della lingua italiana, "WWW" è uno dei pochi acronimi che normalmente si utilizza solo nel linguaggio scritto (in caso di lettura di un testo scritto dove è presente l'acronimo "WWW", solitamente l'acronimo viene letto "World Wide Web" o "tripla W"). Anche nel linguaggio scritto comunque, nella maggior parte dei casi, è preferita l'ellissi "Web" all'acronimo "WWW".

⁴⁰"Grande Ragnatela Mondiale" è un appellativo che nasce come traduzione letterale di "World Wide Web" (anche se la traduzione letterale più accurata sarebbe "ragnatela grande quanto il mondo"). L'appellativo "Grande Ragnatela Mondiale" si è diffuso nei primi anni di affermazione del World Wide Web e oggi è tendenzialmente in disuso.

⁴¹In particolare il *World Wide Web* è uno dei principali servizi di Internet: assieme alla posta elettronica certamente il servizio di Internet più utilizzato e conosciuto.

⁴²Ovviamente sono necessarie opportune risorse (un computer, opportuni *software*, un accesso a Internet, ecc.) e conoscenze. Se i contenuti da pubblicare sul *Web* non sono troppo complessi le risorse e conoscenze necessarie sono comunque di esigua entità. Ad esempio, già dotandosi di un comune *personal computer* attualmente disponibile in commercio e sottoscrivendo un servizio di accesso a Internet con uno dei più popolari *Internet service provider*, si ha a disposizione tutto il necessario per poter pubblicare sul *Web* contenuti non troppo complessi. I più popolari *Internet service provider* offrono infatti la possibilità di pubblicare sul *Web* contenuti non troppo complessi attraverso il *Web* stesso senza la necessità di conoscere il linguaggio di *markup* del *Web*. Lo stesso vale anche per i *blog*, molti sono infatti i siti *web* che offrono gratuitamente la possibilità di creare un *blog* in modo semplice e immediato attraverso il *Web* stesso. In caso invece di contenuti più complessi è necessario dotarsi anche di un *editor web* *WYSIWYG* se si vuole evitare, o almeno semplificare, l'apprendimento del linguaggio di *markup* del *Web*.

⁴³Gli utenti di Internet nel 2008 hanno superato il miliardo e cinquecento milioni, quasi un quarto della popolazione

Il *Web* è stato inizialmente implementato da Tim Berners-Lee mentre era ricercatore al CERN, sulla base di idee dello stesso Berners-Lee e di un suo collega, Robert Cailliau, e oggi gli standard su cui è basato, in continua evoluzione, sono mantenuti dal *World Wide Web Consortium* (W3C).

La nascita del *Web* risale al 6 agosto 1991, giorno in cui Berners-Lee mise *on-line* su Internet il primo sito *Web*. Inizialmente utilizzato solo dalla comunità scientifica, il 30 aprile 1993 il CERN decide di rendere pubblica la tecnologia alla base del *Web*. A tale decisione fa seguito un immediato e ampio successo del *Web* in virtù della possibilità offerta a chiunque di diventare editore, della sua efficienza e, non ultima, della sua semplicità. Con il successo del *Web* ha inizio la crescita esponenziale e inarrestabile di Internet ancora oggi in atto, nonché la cosiddetta "era del *Web*".

Descrizione

Il *Web* è uno spazio elettronico e digitale di Internet destinato alla pubblicazione di contenuti multimediali (testi, immagini, audio, video, ipertesti, ipermedia, ecc.) nonché uno strumento per implementare particolari servizi come ad esempio il download di software (programmi, dati, applicazioni, videogiochi, ecc.). Tale spazio elettronico e tali servizi sono resi disponibili attraverso particolari computer di Internet chiamati server web.

Chiunque disponga di un computer, di un accesso ad Internet, degli opportuni programmi e del cosiddetto spazio web, porzione di memoria di un server web destinata alla memorizzazione di contenuti web e all'implementazione di servizi web, può, nel rispetto delle leggi vigenti nel Paese in cui risiede il server web, pubblicare contenuti multimediali sul *Web* e fornire particolari servizi attraverso il *Web*. I contenuti del *Web* sono infatti costantemente *on-line* quindi costantemente fruibili da chiunque disponga di un computer, di un accesso a Internet, e degli opportuni programmi (in particolare del cosiddetto browser web, il programma che permette, come si dice in gergo, di "navigare" nel *Web*, cioè di fruire dei contenuti e dei servizi del *Web*.)

Non tutti i contenuti e i servizi del *Web* sono però disponibili a chiunque in quanto il proprietario dello spazio web, o chi ne ha delega di utilizzo, può renderli disponibili solo a determinati utenti, gratuitamente o a pagamento, utilizzando il sistema degli account.

I contenuti

I contenuti principali del *Web* sono costituiti da testo e grafica rappresentati in un insieme ristretto di standard definito dal W3C. Tali contenuti sono quelli che tutti i browser web devono essere in grado di fruire autonomamente, cioè senza software aggiuntivo.

I contenuti pubblicati sul *Web* possono essere però di qualunque tipo e in qualunque standard. Alcuni di questi contenuti sono pubblicati per essere fruiti attraverso il browser web e, non essendo in uno degli standard appartenenti all'insieme definito dal W3C, per poterli fruire attraverso il browser web questo deve essere integrato con i cosiddetti *plug-in*, software che integrano le funzionalità di un programma i quali, per quanto riguarda il browser web, normalmente sono scaricabili dal *Web*. Il resto dei contenuti del *Web* è utilizzabile con programmi autonomi. Ad esempio si può trattare di un file eseguibile sul sistema operativo che si sta utilizzando o di un documento di testo in formato Microsoft Word.

mondiale (fonte: *Internet: 2,2 miliardi di utenti nel 2013. Asia e Cina guideranno la crescita mentre l'Italia risulta ancora tra i Paesi poco connessi*).

L'organizzazione dei contenuti

I contenuti del Web sono organizzati nei cosiddetti *siti web* a loro volta strutturati nelle cosiddette *pagine web* le quali si presentano come composizioni di testo e/o grafica visualizzate sullo schermo del computer dal browser web. Le pagine web, anche appartenenti a siti diversi, sono collegate fra loro in modo non sequenziale attraverso i cosiddetti *link* (anche chiamati *collegamenti*), parti di testo e/o grafica di una pagina web che permettono di accedere ad un'altra pagina web, di scaricare particolari contenuti, o di accedere a particolari funzionalità, cliccandoci sopra con il mouse, creando così un ipertesto.

Tutti i siti web, sono identificati dal cosiddetto *indirizzo web*, una sequenza di caratteri univoca chiamata in termini tecnici URL che ne permette la rintracciabilità nel Web.

Non è previsto un indice aggiornato in tempo reale dei contenuti del Web, quindi nel corso degli anni sono nati ed hanno riscosso notevole successo i cosiddetti *motori di ricerca*, siti web da cui è possibile ricercare contenuti nel Web in modo automatico sulla base di parole chiave inserite dall'utente, e i cosiddetti *portali web*, siti web da cui è possibile accedere ad ampie quantità di contenuti del Web selezionati dai redattori del portale web attraverso l'utilizzo di motori di ricerca o su segnalazione dei redattori dei siti web.

I servizi

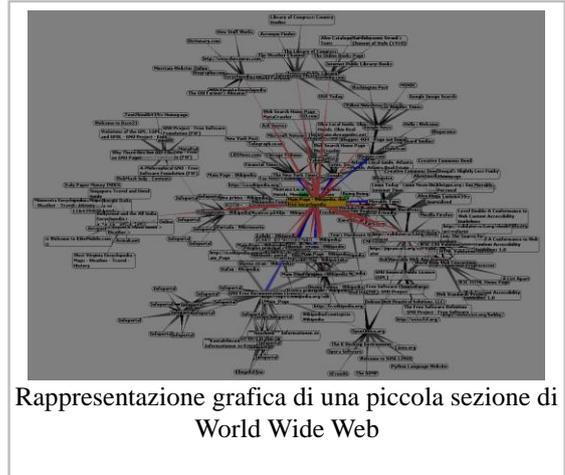
Oltre alla pubblicazione di contenuti multimediali il Web permette di offrire servizi particolari implementabili dagli stessi utenti del Web. I servizi implementabili sono innumerevoli, in pratica limitati solo dalla velocità della linea di telecomunicazioni con cui l'utente e chi fornisce il servizio sono collegati e dalla potenza di calcolo dei loro computer. Di seguito quindi sono elencati solo quelli contraddistinti da una denominazione generica:

- download: la distribuzione di software;
- web mail: la gestione della casella di posta elettronica attraverso il Web;
- streaming: la distribuzione di audio/video in tempo reale;
- web TV: la televisione fruita attraverso il Web;
- web radio: la radio fruita attraverso il Web;
- web chat: la comunicazione testuale in tempo reale tra più utenti di Internet, tramite pagine web;

Implementazione

Il Web è implementato attraverso un insieme di standard, i principali dei quali sono i seguenti:

- HTML (e suoi derivati): il linguaggio di markup con cui sono scritte e descritte le pagine web;
- HTTP: il protocollo di rete appartenente al livello di applicazione del modello ISO/OSI su cui è basato il Web;



- URL: lo schema di identificazione, e quindi di rintracciabilità, dei contenuti e dei servizi del Web.

La peculiarità dei contenuti del Web è quella di non essere memorizzati su un unico computer ma di essere distribuiti su più computer, caratteristica da cui discende efficienza in quanto non vincolati ad una particolare localizzazione fisica. Tale peculiarità è realizzata dal protocollo di rete HTTP il quale permette di vedere i contenuti del Web come un unico insieme di contenuti anche se fisicamente risiedono su una moltitudine di computer di Internet sparsi per il pianeta.

Funzionamento

La visione di una pagina web inizia digitandone l'URL nell'apposito campo del browser web oppure cliccando su un collegamento ipertestuale presente in una pagina web precedentemente visualizzata o in altra risorsa come ad esempio un'e-mail. Il browser web a quel punto dietro le quinte inizia una serie di messaggi di comunicazione con il web server che ospita quella pagina con lo scopo di visualizzarla sul terminale utente.

Per prima cosa la porzione di server-name dell'URL è risolta in un indirizzo IP usando il database globale e distribuito conosciuto come Domain Name System (in sigla DNS). Questo indirizzo IP è necessario per inviare e ricevere pacchetti dal server web.

A questo punto il browser richiede le informazioni inviando una richiesta a quell'indirizzo. In caso di una tipica pagina web, il testo HTML di una pagina è richiesto per primo ed immediatamente interpretato dal browser web che, successivamente, richiede eventuali immagini o file che serviranno per formare la pagina definitiva.

Una volta ricevuti i file richiesti dal web server, il browser formatta la pagina sullo schermo seguendo le specifiche HTML, CSS, o di altri linguaggi web. Ogni immagine e le altre risorse sono incorporate per produrre la pagina web che l'utente vedrà.

Accessibilità e usabilità

Cruciali nell'evoluzione del Web sono diventati i concetti, di accessibilità e usabilità a favore di ogni tipologia di utente, relativi alla progettazione, organizzazione e implementazione dei contenuti secondo specifici requisiti, in comune in generale con le linee evolutive di tutti i prodotti hardware e software in ambito ICT.

Storia

La nascita

La prima proposta di un sistema ipertestuale si può far risalire agli studi di Vannevar Bush, poi pubblicati nell'articolo *As We May Think* (in italiano "Come potremmo pensare") del 1945.

La data di nascita del World Wide Web viene comunemente indicata nel 6 agosto 1991, giorno in cui l'informatico inglese Tim Berners-Lee pubblicò il primo sito web dando così vita al fenomeno "WWW" (detto anche "della tripla W").

L'idea del World Wide Web era nata due anni prima, nel 1989, presso il CERN (Conseil Européen pour la Recherche Nucléaire) di Ginevra, il più importante laboratorio di fisica europeo. Il ricercatore inglese fu colpito da come alcuni colleghi italiani usavano trasmettere informazioni tramite linea telefonica da un piano all'altro dell'istituto visualizzando informazioni tramite video. Il 13 marzo 1989 Tim Berners-Lee presentò infatti al proprio supervisore il documento *Information Management: a Proposal* che fu valutato «vago ma interessante». Alla sua base vi era il progetto dello stesso Berners-Lee e di un suo collega, Robert Cailliau, di elaborare un software per la condivisione di documentazione scientifica in formato elettronico indipendentemente dalla piattaforma informatica utilizzata, con il fine di migliorare la comunicazione, e quindi la cooperazione, tra i ricercatori dell'istituto. A lato della creazione del software, iniziò anche la definizione di standard e protocolli per scambiare documenti su reti di calcolatori: il linguaggio HTML e il protocollo di rete HTTP. Questi standard e protocolli supportavano inizialmente la sola gestione di pagine HTML *statiche*, vale a dire file ipertestuali -preparati precedentemente- visualizzabili e, soprattutto, navigabili utilizzando opportune applicazioni (browser web).

Dopo i primi anni in cui era stato usato solo dalla comunità scientifica, il 30 aprile 1993 il CERN decise di mettere il WWW a disposizione del pubblico rinunciando ad ogni diritto d'autore. La semplicità della tecnologia decretò un immediato successo: in pochi anni il WWW divenne la modalità più diffusa al mondo per inviare e ricevere dati su Internet, facendo nascere quella che oggi è nota come "era del web".

Dal web statico ai web service

Per superare le limitazioni del progetto iniziale, furono subito definiti strumenti capaci di generare pagine HTML dinamiche (ad es. utilizzando dati estratti da un database). La prima soluzione di questo tipo furono le CGI (*Common Gateway Interface*). Attraverso una CGI è possibile richiedere ad un Web server di invocare un'applicazione esterna e presentare il risultato come una qualsiasi pagina HTML. Questa soluzione, sebbene molto semplice da realizzare, presenta numerose limitazioni di progetto (l'applicativo esterno viene eseguito ad ogni richiesta utente e non è prevista alcuna ottimizzazione, non vi è alcuna gestione dello stato della sessione, etc.).



computer utilizzato da Tim Berners-Lee per realizzare il primo server web

Per dare al web una maggiore interattività e dinamicità sono state perseguite due strade. Da un lato sono state aumentate le funzionalità dei browser attraverso un'evoluzione del linguaggio HTML e la possibilità d'interpretazione di linguaggi di scripting (come il JavaScript). Dall'altro, si è migliorata la qualità di elaborazione dei server attraverso una nuova generazione di linguaggi integrati con il Web Server (come JSP, PHP, ASP, etc.), trasformando i Web Server in quelli che sono oggi più propriamente noti come Application Server.

La diffusione di queste soluzioni ha consentito di avviare l'utilizzo del web come piattaforma applicativa che oggi trova la sua massima espressione nei Web Service, alla cui realizzazione e diffusione sta lavorando l'intera industria mondiale del software per la gestione d'azienda, dai grandi nomi commerciali (come SAP e Oracle) fino alle comunità Open Source. L'utilizzo dei web-service all'interno dell'architettura di integrazione SOA permetterà anche alle piccole imprese di gestire senza grandi sforzi i propri processi aziendali.

Scopo dei Web Service è di limitare il più possibile le attività di implementazione, consentendo di accedere a servizi software resi disponibili in rete, assemblarli secondo le proprie necessità e pagarli soltanto per il loro utilizzo effettivo, metodologia individuata nella terminologia anglosassone come *pay per use, on demand software, just in time software, on tap software, etc.*

È chiaro, quindi, che i web-service e il loro successo hanno ed avranno un legame strutturale ed intrinseco con i processi aziendali che dovranno supportare nell'ambito di una nuova organizzazione basata sui processi.

Dal web statico al web semantico

Nonostante tutte queste evoluzioni, il web rimane, ancora e soprattutto, una gigantesca biblioteca di pagine HTML statiche. Wikipedia: Cita le fonti on-line. Però, se da un lato lo standard HTML con la sua semplicità ha contribuito all'affermazione del web, dall'altro ha la grossa limitazione di occuparsi solo ed esclusivamente della formattazione dei documenti, tralasciando del tutto la struttura e il significato del contenuto.

Questo pone notevoli difficoltà nel reperimento e riutilizzo delle informazioni. Per rendersi conto di questo è sufficiente eseguire una ricerca utilizzando uno dei molti motori disponibili in rete e ci si accorgerà che, delle migliaia di documenti risultanti dalla query, spesso solo una piccola percentuale è d'interesse per la ricerca che s'intendeva fare. Ad esempio, per un qualsiasi motore di ricerca, non esiste alcuna differenza fra il termine *Rossi* nel contesto *Il Sig. Rossi* ed il termine *rossi* nel contesto *capelli rossi*, rendendo la ricerca molto difficile.

La risposta a questo problema è venuta, ancora una volta, dal fisico inglese Tim Berners-Lee, che, abbandonato il CERN, ha fondato il consorzio W3C che ha assunto il ruolo di governo nello sviluppo di standard e protocolli legati al web. Egli nel 1998 ha definito lo standard XML (eXtensible Markup Language), un metalinguaggio derivante dall'SGML, che consente la creazione di nuovi linguaggi di marcatura (ad es. lo stesso HTML è stato ridefinito in XML come XHTML). Sua caratteristica innovativa è la possibilità di aggiungere informazioni semantiche sui contenuti attraverso la definizione di opportuni tag.

I principali obiettivi di XML, dichiarati nella prima specifica ufficiale (ottobre 1998), sono pochi ed espliciti: utilizzo del linguaggio su Internet, facilità di creazione dei documenti, supporto di più applicazioni, chiarezza e comprensibilità. Con queste semplici caratteristiche l'XML fornisce un modo comune di rappresentare i dati, cosicché i programmi software siano in grado di eseguire meglio ricerche, visualizzare e manipolare informazioni nascoste nell'oscurità contestuale.

È per questo che, nonostante la sua giovane età, l'XML è alla base di tutte le nuove specifiche tecnologiche rilasciate dal W3C ed è stato adottato come standard di rappresentazione dati da tutta l'industria informatica. Aiuto: Chiarezza (dai *file di configurazione* delle *applicazioni* alla definizione di *formati di interscambio dei dati*). Wikipedia: Cita le fonti

Le specifiche XML hanno però una lacuna molto importante: non definiscono alcun meccanismo univoco e condiviso per specificare relazioni tra informazioni espresse sul web per una loro elaborazione automatica (ad es. più documenti che parlano dello stesso argomento, persona, organizzazione, oggetto), rendendo molto difficile la condivisione delle informazioni.

Anche in questo caso la soluzione al problema è venuta dal W3C di Berners-Lee, attraverso la formalizzazione del web semantico. Il W3C considera l'ideale evoluzione del web dal *machine-representable* al *machine-understandable*. L'idea è di generare documenti che possano non solo essere letti e apprezzati da esseri umani, ma anche accessibili e interpretabili da *agenti* automatici per la ricerca di contenuti.

A tale scopo sono stati definiti alcuni linguaggi, quali Resource Description Framework (RDF) e Web Ontology Language (OWL), entrambi basati su XML, che consentono di esprimere le relazioni tra le informazioni rifacendosi alla logica dei predicati mutuata dall'intelligenza artificiale. Questi standard sono già disponibili, ma continuano ad essere ulteriormente sviluppati insieme a formalismi e strumenti per dotare il web di capacità di inferenza.

Quello appena esposto è un processo solo apparentemente tecnico, ma ben visibile nella sua portata, che ha come obiettivo l'approdo all'intelligenza condivisa del web che promette, a breve, l'uso più efficiente dei siti internet e, a più lungo termine, una autentica trasformazione nella natura del software e dei servizi.

Tanto interesse per queste tecnologie è da ravvisare nel fatto che tutti (utenti, produttori di software e di servizi piccoli e grandi) hanno da avvantaggiarsi dalla diffusione piena di questi standard. La formazione nel corpo del web di una vasta rete "semantica" è, infatti, la condizione chiave per il decollo di un nuovo modo di intendere ed usare il web.

Tim Berners-Lee

Sir Timothy John Berners-Lee (Londra, 8 giugno 1955) è un informatico britannico, co-inventore insieme a Robert Cailliau del World Wide Web.

Biografia

Nel 1976 si laureò in fisica al The Queen's College dell'Università di Oxford, dove, in seguito, costruì il suo primo computer. Successivamente trascorse due anni alla Plessey Telecommunications Ltd, per passare nel 1978 alla D.G Nash Ltd.

Nel 1980 trascorse sei mesi, da giugno a dicembre, al CERN come consulente nel campo dell'ingegneria del software. Lì realizzò, per uso interno nella diffusione di informazioni fra i diversi centri del CERN, il primo software per immagazzinare informazioni usando associazioni casuali: Enquire. Tale prototipo, mai pubblicato, formerà la base concettuale per il futuro sviluppo del World Wide Web.

Dal 1981 al 1984 lavorò al John Poole's Image Computer Systems Ltd. Nel 1984 ritornò al CERN con una borsa di studio per lavorare sui sistemi distribuiti real-time per l'acquisizione di dati scientifici e sistemi di controllo.

Nel 1989 propose⁴⁴ un progetto globale sull'ipertesto, poi noto come World Wide Web.

Il 27 luglio 2012 partecipa alla cerimonia d'apertura della XXX Olimpiade moderna a Londra.

⁴⁴cern.info.ch - Tim Berners-Lee's proposal

World Wide Web Consortium

World Wide Web Consortium



World Wide Web Consortium

Tipo	ONG
Fondazione	ottobre 1994
Altre sedi	380 organizzazioni membri
Area di azione	Mondo 
Presidente	Tim Berners-Lee
(EN)	(IT)
« Leading the Web to Its Full Potential... »	« Guidare il World Wide Web fino al massimo del suo potenziale... »

(Motto del W3C)

Il **World Wide Web Consortium**, anche conosciuto come **W3C**, è un'organizzazione non governativa internazionale che ha come scopo quello di sviluppare tutte le potenzialità del World Wide Web.

Al fine di riuscire nel proprio intento, la principale attività svolta dal W3C consiste nello stabilire standard tecnici per il World Wide Web inerenti sia i linguaggi di markup che i protocolli di comunicazione.

Storia

Il World Wide Web Consortium è stato fondato nell'ottobre del 1994 al MIT (Massachusetts Institute of Technology) dal "padre" del Web, Tim Berners-Lee, in collaborazione con il CERN (il laboratorio dal quale Berners-Lee proveniva).

Nell'aprile del 1995 l'INRIA (Istituto Nazionale di Ricerca Informatica ed Automatica) divenne il primo membro europeo del W3C.

Nel 1996 il W3C approdò in Asia, grazie alla Keio University di Tokio.

Nel 2003 l'ERCIM (Consorzio Europeo di Ricerca in Informatica e Matematica) prese il ruolo di host europeo del W3C dall'INRIA (che è l'istituzione francese dell'ERCIM).

Membri

Il W3C comprende circa 380 membri (al 2012), tra i quali:

- aziende informatiche di primaria importanza, come Adobe, Apple, Cisco Systems, Google, IBM, Intel, Microsoft, Oracle, Siemens, Sony e Sun Microsystems;
- compagnie telefoniche come Ericsson, Nokia, NTT DoCoMo;
- società di grandi dimensioni appartenenti ai più svariati settori, ma strategicamente interessate alla crescita del Web: American Express, Agfa-Gevaert N. V., Boeing, Chevron-Texaco;
- organizzazioni non-profit come la Mozilla Foundation e The Open Group;
- università e istituzioni per la ricerca: il CSAIL del MIT, Inria e altri membri dell'ERCIM e Keio University; altre istituzioni ospitano gli uffici nazionali del Consorzio: per l'Italia l'ISTI di Pisa del CNR; sono numerose le università e gli istituti di ricerca tra i più prestigiosi: Accademia Sinica, la Library of Congress, il Los Alamos National Laboratory, il National Institute of Standards and Technology.

L'importanza dei suoi membri fa del W3C un organismo di grande autorevolezza e molti sono portati a chiamarlo "il Consorzio", per antonomasia.

Obiettivi del W3C

1. Il web ormai non è più uno strumento per «appassionati», ma è diventato parte integrante della vita comune delle persone. Attualmente esistono diversi tipi di apparecchi (come cellulari e PDA) che accedono ad internet. Ciò è possibile solo grazie ad un «comune linguaggio di comunicazione» (nello specifico, il protocollo di comunicazione e linguaggio in cui è scritto il file) tra server, PC ed altri dispositivi. Il W3C si occupa di aggiornare e creare queste specifiche.
2. Il Web ha un potenziale praticamente illimitato, ed apre nuove strade ai portatori di handicap, anche gravi. Il W3C cerca di studiare i modi per rendere quanto più agevole l'accesso al web da parte di questa categoria di persone.
3. Il web è unico perché è libero. Chiunque può creare un documento html e metterlo liberamente online. Il W3C cerca di evitare che interessi di qualsiasi genere possano porre un freno a questa assoluta libertà ed è anche un sito per aiutare a sviluppare le potenzialità comunicative degli esseri umani.

Le raccomandazioni

La definizione di uno standard nell'ambito del W3C passa attraverso gli stadi di:

- *Working Draft* (bozza)
- *Last Call* (ultimo appello)
- *Proposed Recommendation* (proposta di raccomandazione)
- *Candidate Recommendation* (raccomandazione candidata).

Il processo si conclude con la pubblicazione di una Raccomandazione. Una Raccomandazione può venire migliorata da documenti di *Errata* pubblicati separatamente, fino a che il complesso delle modifiche non richiede di produrre una nuova versione della Raccomandazione. Ad es. l'XML è ora alla sua terza versione. Talvolta una Raccomandazione viene ritirata per essere riorganizzata attraverso un nuovo procedimento: ciò è accaduto alla Raccomandazione per l'RDF. Il W3C inoltre pubblica Note informative che non pretendono di essere trattate come standard.

Il Consorzio non è un organismo di standardizzazione, come invece lo sono l'ISO, ITU-T o la Commissione Elettrotecnica Internazionale (IEC): il suo compito è invitare i produttori a seguire le proprie raccomandazioni. A differenza degli organismi di standardizzazione, il W3C non ha un programma di certificazione e molti dei suoi standard di fatto non definiscono formalmente i livelli di conformità. Di conseguenza molte Raccomandazioni sono implementate solo parzialmente.

Browser

In informatica un *web browser* (/ˈbraʊzə(r)/) o **navigatore** è un programma che consente di usufruire dei servizi di connettività in Rete e di navigare sul World Wide Web, appoggiandosi sui protocolli di rete forniti dal sistema operativo (a partire da quelli di livello applicativo come HTTP, FTP ecc.) attraverso opportune API, permettendo di visualizzare i contenuti delle pagine dei siti web, specificandone l'URL, e di interagire con essi. Quest'ultima funzionalità è supportata dalla capacità del browser di interpretare l'HTML — il codice con il quale sono scritte la maggior parte delle pagine web — e di visualizzarlo in forma di ipertesto grazie al motore di rendering.

Descrizione

Esso rappresenta dunque il sistema software di interfacciamento dell'utente con la rete che rende la navigazione dell'utente tipicamente user-friendly, sebbene ai primordi della rete siano esistiti anche browser testuali da riga di comando su shell.

I browser vengono principalmente utilizzati su personal computer, ma anche su altri dispositivi che consentono la navigazione in Internet, come i palmari e gli smartphone. Quelli più noti e diffusi sono Internet Explorer, Mozilla Firefox, Google Chrome e Opera.

Storia

Il primo browser fu sviluppato da Tim Berners-Lee (tra i primi precursori del concetto di WWW e fondatore del W3C), e venne chiamato WorldWideWeb. Serviva a scopi dimostrativi, era disponibile solo per sistema operativo NeXT e perciò in seguito fu chiamato Nexus.

Il primo browser a raggiungere un'apprezzabile popolarità internazionale fu Mosaic, sviluppato da NCSA, seguito poi da Netscape Navigator, che crebbe in fretta e fu la piattaforma su cui vennero messe a punto alcune innovazioni diffuse a larga scala col passare del tempo (come ad esempio JavaScript). La netta prevalenza di Netscape presso l'utenza, a metà degli anni novanta, fu incrinata dalla cosiddetta *guerra dei browser*, una competizione inizialmente commerciale e poi di immagine cominciata da Microsoft quando le innovazioni introdotte da Netscape divennero così sofisticate da costituire una minaccia potenziale per i propri interessi. Per vincere la concorrenza, Microsoft incluse Internet Explorer nel proprio sistema operativo Windows (il più diffuso al mondo), stroncando sul nascere i possibili concorrenti. Questa mossa fu motivo di numerose cause legali per la difesa della libera concorrenza e contro la nascita di monopoli informatici.

Netscape reagì rilasciando nel 1998 il proprio codice con una licenza open source. Il progetto Mozilla che ne derivò fornisce il codice che è alla base di diversi browser, fra i quali Netscape, Mozilla Suite, Galeon, Firefox e uno dei programmi di accesso della AOL.

La presenza di browser diversi con funzioni differenti, ha portato alcuni webmaster a realizzare siti web destinati a essere visitati con un browser preferenziale, talvolta impedendo l'accesso a utenti che utilizzassero un browser differente da quello scelto. Questa consuetudine, oltre a essere contro la filosofia portante del World Wide Web, che vede l'accessibilità come uno dei pilastri portanti e aver quindi suscitato forti critiche e campagne di protesta e sensibilizzazione (come Campaign for a Non-Browser Specific WWW), ha costretto browser molto comuni (tra i quali Opera) a "*fingersi*" un altro browser (modificando il proprio user agent), per poter aver accesso a un maggior numero di documenti.

Il browser più diffuso a livello mondiale era Internet Explorer, prodotto da Microsoft, che viene fornito gratuitamente e "di serie" con ogni sistema operativo Windows (dati al 2006). Il secondo browser in ordine di utilizzo era Mozilla nelle sue varianti (e in particolare Firefox), prodotto e distribuito gratuitamente dalla Mozilla Foundation; il terzo era Safari di Apple⁴⁵, distribuito assieme al sistema operativo Mac OS X (di cui più recentemente è stata messa a disposizione una versione gratuita anche per Windows). Altri browser molto diffusi sono Opera (terzo in Italia nel 2006) e il già citato Netscape Navigator, il cui sviluppo è ormai sospeso.

Fra le principali funzionalità dei browser disponibili, si possono ricordare: navigazione a schede (*Tabbed browsing*), supporto alla navigazione *off-line* tramite la memoria cache e *plugin* dedicati per mantenere i link tra le pagine salvate, funzione di *download manager* con arresto/ripresa sempre tramite la memoria cache, anteprima delle pagine da scaricare, sintesi vocale, integrazione dei *feed RSS* e di *client* di posta elettronica.

La diversa diffusione dei vari browser, con tecnologie compatibili tra loro, ma con alcune peculiarità uniche, ha portato in passato con Internet Explorer 6 e attualmente con i sistemi basati sul WebKit a progettare le pagine web tenendo conto di un solo browser causando in alcuni casi il funzionamento solo su specifici browser, il che può portare a creare uno standard ancora prima che venga approvato dagli organi competenti⁴⁶.

HTML

In informatica l'**HyperText Markup Language (HTML)** (traduzione letterale: linguaggio a marcatori per ipertesti) è il linguaggio di markup solitamente usato per la formattazione di documenti ipertestuali disponibili nel World Wide Web sotto forma di pagine web.

In generale una pagina web, per essere visibile e intelligibile sul Web, deve essere opportunamente formattata. Il linguaggio di formattazione è l'HTML. La formattazione consiste nell'inserimento nel testo di marcatori o etichette, detti *tag*, che descrivono caratteristiche come la funzione, il colore, le dimensioni, la posizione relativa all'interno della pagina.

Il contenuto delle pagine web solitamente consiste dunque di un documento HTML e dei file ad esso correlati che un web browser scarica da uno o più web server per elaborarli, interpretando il codice sorgente, al fine di generare la visualizzazione, sullo schermo del computer-client, della pagina desiderata, grazie al motore di rendering del browser stesso.

⁴⁵Safari terzo nella classifica dei browser | Macitynet

⁴⁶Daniel Glazman del W3C: Google e Apple peggio di Internet Explorer 6 coi loro browser mobile

L'HTML non è un linguaggio di programmazione (in quanto non prevede alcuna definizione di variabili, strutture dati, funzioni, strutture di controllo), ma solamente un *linguaggio di formattazione* che descrive cioè le modalità di impaginazione o visualizzazione grafica (layout) del contenuto, testuale e non, di una pagina web attraverso *tag di formattazione*. Tuttavia, l'HTML supporta l'inserimento di script e oggetti esterni quali immagini o filmati. *Punto* HTML (.html) o *punto* HTM (.htm) è anche l'estensione comune per riconoscere i documenti in questo formato.

Descrizione

L'HTML è un linguaggio di pubblico dominio la cui sintassi è stabilita dal World Wide Web Consortium (W3C), e che è basato su un altro linguaggio avente scopi più generici, l'SGML.

È stato sviluppato verso la fine degli anni ottanta da Tim Berners-Lee al CERN di Ginevra assieme al noto protocollo HTTP che supporta invece il trasferimento di documenti in tale formato. Verso il 1994 ha avuto una forte diffusione in seguito ai primi utilizzi commerciali del web.

Nel corso degli anni, seguendo lo sviluppo di Internet, l'HTML ha subito molte revisioni, ampliamenti e miglioramenti, che sono stati indicati secondo la classica numerazione usata per descrivere le versioni dei software. Attualmente l'ultima versione disponibile è la versione 4.01, resa pubblica il 24 dicembre 1999.

Dopo un periodo di sospensione, in cui il W3C si è focalizzato soprattutto sulle definizioni di XHTML (applicazione a HTML di regole e sintassi in stile XML) e dei fogli di stile (CSS), nel 2007 è ricominciata l'attività di specifica con la definizione, ancora in corso, di HTML5, attualmente allo stato di release candidate all'ultima votazione.

Un'ulteriore ed importante caratteristica di HTML è che esso è stato concepito per definire il contenuto logico e non l'aspetto finale del documento. I dispositivi che possono accedere ad un documento HTML sono molteplici e non sempre dotati di potenti capacità grafiche. Proprio per questo gli sviluppatori di HTML hanno optato per un linguaggio che descrivesse dal punto di vista logico, piuttosto che grafico, il contenuto dei documenti. Questo significa che non esiste alcuna garanzia che uno stesso documento venga visualizzato in egual modo su due dispositivi. Se da una parte questo ha imposto in passato dei forti limiti agli sviluppatori di pagine Web, ha dall'altro garantito la massima diffusione di Internet ed evitato che essa diventasse un medium di élite.

Attualmente i documenti HTML sono in grado di incorporare molte tecnologie, che offrono la possibilità di aggiungere al documento ipertestuale controlli più sofisticati sulla resa grafica, interazioni dinamiche con l'utente, animazioni interattive e contenuti multimediali. Si tratta di linguaggi come CSS, JavaScript e jQuery, XML, JSON, o di altre applicazioni multimediali di animazione vettoriale o di streaming audio o video.

Al giorno d'oggi molti web designer delegano la scrittura del codice HTML ad applicazioni specifiche, come per esempio i cosiddetti *editor WYSIWYG* che permettono al designer di occuparsi dell'aspetto grafico finale della pagina mentre il codice vero e proprio viene generato automaticamente. Gli sviluppatori puri preferiscono invece utilizzare direttamente il codice HTML, in modo da avere un maggior controllo sul risultato finale e sulla pulizia del codice scritto, cosa che gli editor WYSIWYG odierni nonostante siano sempre più avanzati non possono sempre garantire, vuoi anche per la diversa resa dei browser sulle parti di codice HTML.

```
1 <!DOCTYPE html PUBLIC "-//W3C/DTD HTML
2 <html>
3   <head>
4     <title>Example</title>
5     <link href="screen.css" rel="sty
6   </head>
7   <body>
8     <h1>
9       <a href="/">Header</a>
10    </h1>
11    <ul id="nav">
12      <li>
13        <a href="one/">One</a>
14      </li>
15      <li>
16        <a href="two/">Two</a>
17      </li>
```

Un esempio di codice HTML con sintassi evidenziata. Questa funzione viene solitamente eseguita da un editor di pagine web o da alcuni browser, come ad esempio Mozilla Firefox.

Accesso via Internet

I documenti HTML vengono immagazzinati sui dischi rigidi di macchine elaboratrici (computer-server) costantemente collegate e connesse alla rete Internet. Su queste macchine è installato un software specifico (web server) che si occupa di produrre e inviare i documenti ai browser degli utenti che ne fanno richiesta usando il protocollo HTTP per il trasferimento dati.

Spesso il documento HTML viene generato del tutto o parzialmente tramite un codice eseguibile residente sul server Internet (elaborazione lato server) in grado di interagire con altre applicazioni presenti sul server stesso, come per esempio un database, e inviare poi al browser il risultato finale, realizzando le cosiddette *pagine web dinamiche* con cui un utente può compiere operazioni interattive avanzate (ad es. filtrare gli articoli all'interno di un catalogo on-line, inviare e registrare dati, fare login ecc.). È il caso di documenti scritti in linguaggi come ASP, PHP, Perl o Java. In altri casi invece l'elaborazione è operata lato client con linguaggi come Javascript.

Elementi sintattici

Ogni documento ipertestuale scritto in HTML deve essere contenuto in un file, la cui estensione è tipicamente .htm o .html (l'estensione htm dipende dal fatto che all'epoca il sistema operativo MS-DOS non permetteva di utilizzare estensioni superiori a 3 caratteri).

Il componente principale della sintassi di questo linguaggio è l'*elemento*, inteso come struttura di base a cui è delegata la funzione di formattare i dati o indicare al browser delle informazioni.

Ogni elemento è racchiuso all'interno di marcature dette *tag*, costituite da una sequenza di caratteri racchiusa tra due parentesi angolari o uncinata (<>), cioè i segni minore e maggiore (Es.:
; il tag di questo esempio serve per indicare un ritorno a capo).

Quando il tag deve essere applicato a una sezione di testo o di codice, l'ambito di applicazione deve essere delimitato fra un tag di apertura ed uno di chiusura (chiusura esplicita), che coincide col tag di apertura preceduto da una barra (/) dopo la parentesi angolare aperta (Es.: testo testo, in questo caso, il testo compreso tra questi due tag verrà visualizzato in grassetto dal browser).

Alcuni tag presentano un'applicazione puntuale, come per esempio il tag che serve per inserire un'immagine in un determinato punto della pagina, e in quanto tali non richiedono il tag di chiusura; in questo caso si parla di *tag a chiusura implicita*. In XHTML, invece, la chiusura implicita è proibita e tutti i tag devono essere sempre chiusi esplicitamente tramite un tag di chiusura o, nel caso dei tag puntuali, usando il carattere '/' alla fine del tag stesso (per esempio
). Per questi tag, i browser sono in grado di accettare entrambe le modalità, per motivi di compatibilità.

Struttura di un documento HTML

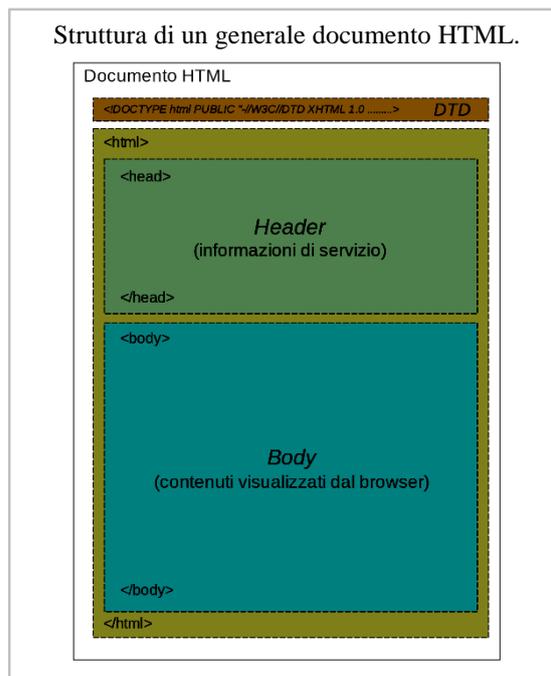
Un documento HTML comincia con l'indicazione della definizione del tipo di documento (*Document Type Definition* o DTD), la quale segnala al browser l'URL delle specifiche HTML utilizzate per il documento, indicando quindi, implicitamente, quali elementi, attributi ed entità si possono utilizzare e a quale versione di HTML si fa riferimento. Di fatto, questa informazione serve al browser per identificare le regole di interpretazione e visualizzazione appropriate per lo specifico documento. Questa definizione deve pertanto precedere tutti i tag relativi al documento stesso.

Dopo il DTD, il documento HTML presenta una struttura ad albero annidato, composta da sezioni delimitate da tag opportuni che al loro interno contengono a loro volta sottosezioni più piccole, sempre delimitate da tag.

La struttura più esterna è quella che delimita l'intero documento, eccetto la DTD, ed è compresa tra i tag `<html>` e `</html>`.

All'interno dei tag `<html>` lo standard prevede sempre la definizione di due sezioni ben distinte e disposte in sequenza ordinata:

- la sezione di intestazione o *header*, delimitata tra i tag `<head>` e `</head>`, che contiene informazioni di controllo normalmente non visualizzate dal browser, con l'eccezione di alcuni elementi
- la sezione del corpo o *body*, delimitata tra i tag `<body>` e `</body>`, che contiene la parte informativa vera e propria, ossia il testo, le immagini e i collegamenti che costituiscono la parte visualizzata dal browser.



Al di sotto di questa suddivisione generale, lo standard non prevede particolari obblighi per quanto riguarda l'ordine e il posizionamento delle ulteriori sottosezioni all'interno dell'*header* o del *body*, a parte l'indicazione del rispetto dei corretti annidamenti (le sottosezioni non si devono sovrapporre, ossia ogni sottosezione deve essere chiusa prima di iniziare la sottosezione successiva), lasciando così completa libertà allo sviluppatore o al progettista per quanto riguarda la strutturazione e l'organizzazione successive.

Pagina web

Una **pagina web** è il modo in cui vengono rese disponibili all'utente finale le informazioni del World Wide Web della rete Internet tramite un web browser. Un insieme di pagine web, tra loro relazionate secondo una gerarchica e una struttura ipertestuale e riferibili, di norma, ad un unico web server, costituiscono un sito web.

Descrizione

Informalmente e logicamente una pagina web si può suddividere in una parte relativa ai contenuti ed una parte di formattazione (layout), presentazione grafica o organizzazione dei contenuti stessi all'utente (testo e/o immagini digitali). La pagina web principale e iniziale di un sito web è la cosiddetta home page, da cui dipartono tramite collegamenti o link interni (*ancore*), spesso riuniti in menù e widget, verso pagine secondarie del sito.

Tipologie

Esistono due tipologie di pagine web che identificano rispettivamente due diversi paradigmi di programmazione Web:

- pagina web statica afferente al paradigma del cosiddetto Web statico;
- pagina web dinamica afferente al paradigma del cosiddetto Web dinamico;

Nel primo caso le informazioni, con la loro descrizione, sono di solito contenute in ipertesti in formato HTML o XHTML che consentono di accedere ad altre pagine web o altri contenuti informativi tramite collegamenti ipertestuali (*link* o rimandi). Nel secondo caso si fa invece uso di linguaggi di scripting.

A seconda della tecnologia impiegata, le pagine web possono contenere elementi grafici o multimediali e elementi interattivi e dinamici.

URL

Di solito una pagina web ha uno o più URL (*Universal Resource Locator*), un link permanente che ne permette il deep linking; qualche volta ha solo un URL temporaneo che si riferisce ad un'area cache. Questo è ad esempio il caso di quando una pagina è il risultato di uno zoom o di uno spostamento di una mappa oppure una pagina può essere il risultato dovuto a qualche azione. Una pagina di questo tipo, come una pagina con un form riempito, non potrebbe essere precisamente chiamata una pagina web, in quanto non è nel web ma creata da qualcosa che è nel web.

L'URL digitato nel web browser dell'utente è convertito dal DNS in indirizzo IP necessario per l'instradamento della comunicazione tra web server e client nei nodi interni di commutazione (router).

Grafica

Il formato file della grafica in una pagina web è di solito JPEG per fotografie e GIF o PNG per altre immagini (come diagrammi, disegni, grafici, etc.). Gli ultimi due formati possono essere utilizzati per foto, ma non sono convenienti per lo scopo come JPEG, che è un formato a perdita d'informazione, ma a maggior livello di compressione, mentre GIF e PNG non lo sono, cioè meno performanti a livello di compressione. GIF è utilizzata per animazioni, GIF e PNG per immagini con pixel trasparenti, PNG per immagini con pixel parzialmente trasparenti (anche se Internet Explorer 6 non supporta le trasparenze del formato PNG).

Tutti questi formati appartengono alla grafica *raster*. I modi più comuni di impiegare grafica vettoriale nelle pagine web sono tre:

- utilizzando il moderno formato SVG (non ancora supportato da molti browser);
- inserendo dei file PDF, visualizzabili con un plug-in del browser o con un programma separato;
- inserendo dei file di Flash, che permette l'animazione degli elementi e l'utilizzo di musica ed effetti sonori.

Ad esempio, per creare una mappa, può essere utile una combinazione di livelli a grafica vettoriale e testo, possibilmente con un livello di grafica raster. Questo fornisce migliori risultati nello zoom rispetto ad un'immagine GIF o PNG (JPEG dovrebbe dar luogo a risultati peggiori dovuti agli artefatti di compressione).

Creare una pagina web

Per creare una pagina web statica, c'è bisogno di un generico web editor, testuale direttamente in HTML oppure visuale (es. i cosiddetti WYSIWYG) (o anche un semplice convertitore in formato HTML come consente di fare, ad esempio, Microsoft Word in fase di salvataggio di un pagina o documento), e un programma FTP per effettuare l'upload della pagina verso il server web che ospiterà (hosting) la pagina web. Si può utilizzare anche il browser web per effettuare l'invio delle pagine web al server, ma non è consigliabile.

Si può utilizzare anche un modello web già pronto. I modelli web sono pagine html già fatte pronte all'uso. I creatori di pagine web devono solo inserire il titolo ed i contenuti. Inoltre questi modelli sono riusabili.

Wiki è invece un modo particolare di creare o modificare e inviare una pagina web senza utilizzare un programma FTP ed effettuare l'upload del file, riempiendo solamente un modulo di testo in una pagina web. Wikipedia è un esempio di applicazione della tecnologia wiki.

Per creare una pagina web dinamica occorre invece conoscere almeno un linguaggio di scripting per programmazione Web ed utilizzare i relativi editor.

Standard per pagine web

Dal punto di vista tecnologico i vari aspetti caratteristici di una pagina vengono realizzati utilizzando i seguenti standard:

- l'HTML, che è il linguaggio di descrizione dei documenti Web;
- l'XHTML, che è l'evoluzione dell'HTML, e si poggia sulla tecnologia XML;
- i fogli di stile (CSS), che sono utilizzati per controllare la presentazione delle pagine su diverse tipologie di dispositivi (testuali, grafici, di stampa, ecc.)
- linguaggi di scripting come JavaScript, PHP, ASP, .NET, Python ecc... per pagine web dinamiche.

La "normalizzazione" dell'output da parte del Browser web può ovviamente variare in funzione della correttezza semantica del linguaggio di programmazione usato. Al fine di uniformare la rasterizzazione della pagina web esiste un "organo competente" che si pone l'obiettivo di determinare quando il codice sia compilato in maniera da rendersi identico su tutti i vari browser, stiamo parlando del W3C. Tramite un semplice processo di validazione del codice tramite invio dell'url che si intende controllare siamo così in grado di capire se l'output html, xhtml, css etc.. sia "valido" ovvero uguale su tutti i browser.

Sito web

Un **sito web** o **sito Internet** (anche abbreviato in **sito** se chiaro il contesto informatico) è un insieme di pagine web correlate, ovvero una struttura ipertestuale di documenti che risiede, tramite hosting, su un web server e accessibile all'utente client che ne fa richiesta tramite un web browser sul World Wide Web della rete Internet, digitando in esso il rispettivo URL o direttamente l'indirizzo IP.

Descrizione

Dal punto di vista dell'organizzazione o struttura dei contenuti tipicamente un sito web è strutturato in una home page, che rappresenta la pagina principale di presentazione del sito e dei suoi contenuti e nelle pagine web secondarie dei contenuti, raggiungibili dai menù e widget presenti nell'home page stessa tramite link interni (*ancore*) al sito stesso. Quando il sito web costituisce un punto di partenza, una porta di ingresso, ad un gruppo consistente di risorse di Internet o di una Intranet si parla più propriamente di portale web. Spesso inoltre un sito web si accoppia/accompagna ad un database sullo stesso web server per la memorizzazione strutturata di dati e informazioni.

Accesso alle pagine web

Le pagine di un sito web sono accessibili tramite una radice comune (detta "nome di dominio", per esempio "*www.kernel.org*"), seguita da una serie opzionale di "sotto cartelle" e dal nome della pagina. Il nome completo di ogni pagina è detto "indirizzo web" o, più tecnicamente, URI (o URL). L'home page di un sito è la prima pagina che si ottiene digitando il solo nome di dominio.

Per esempio, nell'indirizzo **www.w3c.org/Consortium/Offices/role.html**:

- **www.w3c.org** è la radice, o nome di dominio
- **/Consortium/Offices/** sono le sottocartelle, separate dal simbolo "/"
- **role.html** è il nome della pagina

Solitamente le pagine di un sito risiedono tutte sullo stesso Web server, ossia lo spazio fisico in cui risiedono i files che compongono il sito, e la ramificazione in sottocartelle dell'indirizzo corrisponde ad una uguale ramificazione nell'hard disk dello stesso server.

Ovviamente l'accesso ad un sito o alle pagine web per eventuali modifiche su contenuti e/o forma da parte degli sviluppatori è del tutto riservato all'amministratore del sito tramite codici o chiavi di accesso.

Per l'interpretazione del codice che compone una pagina web vengono comunemente utilizzati i browser, ovvero quei software che hanno il compito di interpretare il codice e restituirne un output grafico. Tra i browser più comunemente utilizzati possiamo trovare (a titolo indicativo e non esaustivo) Mozilla Firefox, Google Chrome, Opera, Microsoft Internet Explorer e Safari. I file interpretati da questi software all'accesso della pagina web sono solitamente i fogli di stile (CSS) e i vari script che vengono richiamati all'accesso della pagina. Questi determinano dunque l'aspetto grafico finale restituito all'utente all'accesso della pagina.

Tipologie di siti web

Dal punto di vista della struttura e della tecnologia utilizzata, i siti web si possono distinguere sommariamente in due tipologie principali:

- *siti statici*
- *siti dinamici*

I siti web *statici* formati da pagine statiche presentano contenuti di sola ed esclusiva lettura. Solitamente vengono aggiornati con una bassa frequenza e sono mantenuti da una o più persone che agiscono direttamente sul codice HTML della pagina (tramite appositi editor web). Si tratta storicamente della prima generazione di siti web.

I siti web *dinamici* formati da pagine web dinamiche presentano invece contenuti redatti *dinamicamente* (in genere grazie al collegamento con un database) e forniscono contenuti che possono variare in base a più fattori. I siti web dinamici sono caratterizzati da un'alta interazione fra sito e utente; alcuni elementi che caratterizzano la dinamicità di un sito possono essere: l'interazione con uno o più database, la visualizzazione dell'ora server, operazioni varie sul file system (tipo creazione dinamica di documenti, ecc.), il cloaking basandosi su vari parametri (tra i quali lo user agent, o il browser, o la risoluzione dell'utente), la visualizzazione o stampa o catalogazione degli indirizzi IP degli utenti, e molto altro. Essi possono essere scritti tramite linguaggi di scripting come ASP, PHP e molti altri linguaggi web. Particolarità dei siti dinamici è che le "pagine" che costituiscono il sito vengono generate all'atto della chiamata da parte dell'utente (fatta eccezione per i sistemi che adottano una cache) in quanto i contenuti non sono memorizzati direttamente nel codice sorgente della pagina, bensì in sistemi di memorizzazione esterni allo stesso, come le già citate basi di dati.

La modifica dei contenuti, che spesso possono essere aggiornati grazie ad alcuni editor grafici (detti *WYSIWYG*) anche senza agire direttamente sul codice, è generalmente frequente. Un chiaro esempio di sito dinamico è www.wikipedia.org.

Categorie

I siti web vengono poi comunemente divisi in categorie per inquadrarne il settore di operatività o i servizi offerti:

- *sito personale* - condotto da individui o piccoli gruppi di persone (come una famiglia), che contiene informazioni prevalentemente autobiografiche o focalizzate sui propri interessi personali (ad es. un blog)
- *sito aziendale* - funzionale alla promozione di un'azienda o di un servizio
- *sito di commercio elettronico (o "e-commerce")* - specializzato nella vendita di beni e/o servizi via internet
- *sito comunità* - un sito in cui utenti possono comunicare fra loro, per esempio tramite chat o forum, formando una comunità virtuale
- *sito di download* - ampie raccolte di link per scaricare software (dimostrativi di videogiochi, immagini, ecc.)
- *forum*: luogo in cui discutere tramite la pubblicazione e la lettura di messaggi, organizzati per discussioni (*thread*) e messaggi (*post*)
- *sito informativo* - con contenuti tesi a informare l'utente, ma non strettamente per scopi commerciali
- *motore di ricerca* - registra i contenuti degli altri siti e li rende disponibili per la ricerca
- *database* - un sito il cui utilizzo principale è di ricercare e mostrare il contenuto di uno specifico database (ad esempio l'*Internet Movie Database* per i film)
- *sito ludico* - un sito che è in sé un gioco oppure serve da arena per permettere a più persone di giocare
- *aggregatore di notizie* - presenta contenuti prelevandoli automaticamente dalla rete da molte fonti contemporaneamente, ed è simile in questo ai motori di ricerca che sfruttano i risultati delle ricerche degli utenti aggregandoli e immagazzinandoli in un database con criteri quali la tipologia dei media (foto, video, testo etc.) e il contenuto testuale. Il risultato è un sito che cresce esponenzialmente e si arricchisce all'infinito di pagine con contenuto inerente all'argomento di base, definito in fase di installazione ma affinato sulle preferenze degli utenti, il tutto in modo automatico.

- *link farm* - siti creati unicamente per proporre collegamenti verso altri siti a scopo pubblicitario (spesso in cambio di canoni in denaro)

Molti siti sono un incrocio tra due o più delle precedenti categorie. Ad esempio, un sito aziendale può allo stesso tempo pubblicizzare i propri prodotti e pubblicare materiale informativo.

Prende il nome di *portale* un sito web che ha assunto dimensioni tali da costituire un punto di partenza, una porta di ingresso ad un gruppo consistente di risorse per gli utenti del web; l'argomento del portale può essere più o meno generalista oppure specializzato (*portale verticale*).

Taluni siti web presentano *aree riservate*, ovvero sezioni accessibili solamente previo pagamento di un canone, registrazione, o assegnazione di una password.

Progettazione e realizzazione

La progettazione di un sito web è analoga a larghe linee al modello di sviluppo del software tipica dell'ingegneria del software con la definizione/pianificazione iniziale di ciò che deve essere realizzato a livello logico (analisi/definizione delle specifiche, strutturazione dei contenuti e grafica/formattazione), seguita dalla fase implementativa e dalla fase di testing e successiva eventuale correzione delle incongruenze con le specifiche o l'usabilità del sito stesso.

La realizzazione/gestione di un sito web fa ampio uso di tecniche e conoscenze proprie della programmazione Web o di format predefiniti direttamente a disposizione dell'utente amministratore forniti da strumenti appositi o reperibili in rete e dall'utilizzo tipicamente user friendly.

I linguaggi

Il linguaggio più diffuso con cui i siti web sono realizzati è l'HTML (*Hyper Text Markup Language*), i suoi derivati e i CSS. In particolare l'HTML viene interpretato da particolari software chiamati web browser. Alcuni plugin per i browser, come Flash, Shockwave o applet Java, permettono la visualizzazione di contenuti animati,

Per i siti web dinamici dove alcuni contenuti possono essere generati dinamicamente sul browser dell'utente, in aggiunta ai precedenti linguaggi si utilizzano linguaggi di scripting come JavaScript o Dynamic HTML, tecnologie supportate per impostazione predefinita da tutti i browser recenti (programmazione web client-side).

Per la costruzione/realizzazione di siti web dinamici in grado invece di estrapolare dati da database, inviare email, gestire informazioni, ecc., i linguaggi di scripting più diffusi sono PHP, Ruby, Python e ASP in alternativa al Java, solitamente utilizzato per applicazioni enterprise (programmazione web server-side).

Per la creazione di siti internet molti webmaster utilizzano comunemente strumenti automatizzati, chiamati webeditor (es. WYSIWYG), grazie ai quali possono redigere il codice anche senza avere un'approfondita conoscenza dei linguaggi.

Sempre più la realizzazione e la gestione di siti web complessi e ricchi di contenuti passa attraverso l'uso di Content Management System CMS, delle vere piattaforme web (sorta di evoluzione dei web editor) per le quali non è richiesta alcuna conoscenza di tecniche di programmazione web ed utili anche per forum, blog e portali di e-commerce. Queste piattaforme spesso si appoggiano su piattaforme XAMPP.

Accessibilità dei siti web

I siti web devono soddisfare dei requisiti di accessibilità che rendano i suoi contenuti fruibili a chiunque, compresi i soggetti affetti da disabilità psichica e fisica, definiti in regole precise denominate Web Content Accessibility Guidelines (WCAG) definite dal World Wide Web Consortium (W3C).

Usabilità del web

L'**usabilità del web** o *web usability* è l'applicazione dei principi dell'usabilità in quegli ambiti in cui la navigazione del web può essere considerata come un paradigma generale (o "metafora") per costruire una interfaccia grafica (GUI). L'usabilità del web è una proprietà esterna al prodotto-sito poiché riguarda l'interazione tra un utente, le sue caratteristiche e quel sito. La normativa ISO/IEC 2001a definisce l'usabilità come "*la capacità di un sistema di essere compreso, appreso, utilizzato e attraente per gli utenti quando usato sotto condizioni specificate*".

- La *Comprensibilità* di un sito riguarda lo sforzo richiesto per capire il sistema
- L'*apprendibilità* riguarda lo sforzo necessario all'utente per imparare ad usare il sistema
- L'*utilizzabilità* si riferisce allo sforzo richiesto all'utente per utilizzare il sistema agendo sui suoi controlli.

Un'altra definizione di usabilità tratta dalle indicazioni ISO 9241-11:1998 la indica come:

"Il grado in cui un prodotto può essere usato da particolari utenti per raggiungere certi obiettivi con efficacia, efficienza, soddisfazione in uno specifico contesto d'uso" intendendo

- *Efficacia* come precisione e completezza con cui gli utenti raggiungono specifici obiettivi
- *Efficienza* come risorse impiegate in relazione alla precisione e completezza cui gli utenti raggiungono specifici obiettivi
- *Soddisfazione* come libertà dal disagio e attitudine positiva con cui gli utenti raggiungono specifici obiettivi attraverso l'uso del prodotto.

In questo contesto si parla di usabilità in termini di interazione efficace e soddisfacente per l'utente sia in termini di efficienza sia in termini di benessere. L'usabilità ha quindi come obiettivi quello di economizzare lo sforzo cognitivo dell'utente, proponendo prodotti che siano facili da comprendere, da imparare, da usare, da ricordare, che evitino o rendano recuperabili gli errori e che quindi gratifichino l'utilizzatore.

Evoluzione del concetto di usabilità

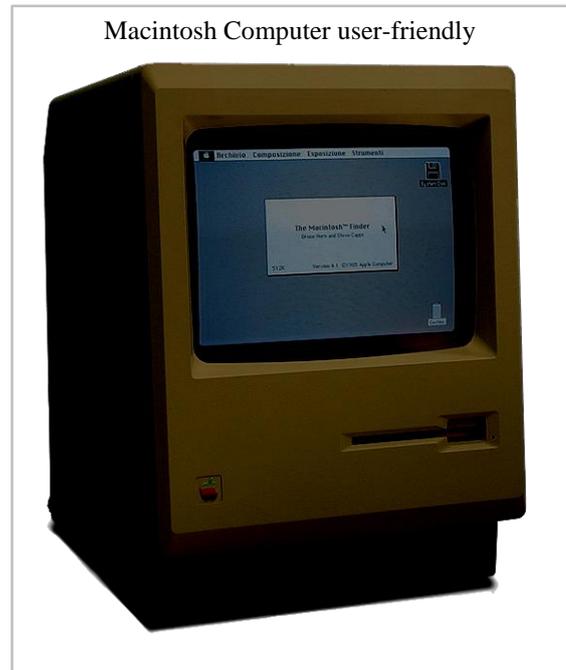
Il concetto di usabilità del web è strettamente connesso a quello di usabilità generale che ha cominciato a diffondersi negli anni 60 quando si inizia a parlare di ergonomia cognitiva, la scienza che studia le interazioni dell'uomo con l'ambiente in relazione ai suoi vincoli fisici e cognitivi. Per tutti gli anni 70, poiché il computer non costituiva un prodotto di massa, l'usabilità non costituiva un problema: chi progettava un sistema sapeva anche usarlo. Il concetto si viene a sviluppare a partire dalla seconda metà degli anni '80, con l'espansione delle tecnologie informatiche a livello degli ambienti lavorativi e delle famiglie e si configura con maggior precisione negli anni '90 con la diffusione del personal computer. In questo nuovo contesto il concetto di usabilità viene ad inserirsi nell'ambiente Information Communication Technology (ICT), come disciplina che coniuga psicologia, intelligenza artificiale, informatica e design. In questo periodo nasce anche la Human Computer Interaction (HCI), la scienza che studia l'interazione tra uomo e computer, in cui l'usabilità riguarda la componente psicologica del software. In questi anni Apple Macintosh realizza il primo computer completamente usabile (user-friendly) da masse di inesperti, con un sistema operativo visuale basato sulla metafora della scrivania, seguito poco dopo da Windows che attuerà una politica di vendita più aggressiva, raggiungendo un numero molto alto di utenti.

L'avvento di Internet, a partire dalla metà degli anni '90 ha ulteriormente ampliato il concetto di usabilità che si sposta dalle caratteristiche del software a quelle dell'interazione. L'usabilità dei siti web viene a configurarsi come elemento determinante perché è condizione preliminare al loro effettivo utilizzo da parte dell'utente e, in termini di produttività, al realizzarsi stesso del guadagno.

Web Usability e progettazione

La web usability è un approccio della progettazione per rendere i siti web facili da usare per l'utente finale, senza richiedere all'utente di sottostare a una formazione specifica. È un processo che cerca di rendere l'ambiente web attraente e amichevole per l'utente che deve navigare con fluidità recuperando facilmente i contenuti. L'utente deve essere in grado di mettere in relazione in modo intuitivo le azioni che ha bisogno di compiere nella pagina web con altre interazioni che svolge fisicamente nella vita quotidiana, come ad esempio premere un pulsante per ottenere un'azione. La web usability si pone come obiettivi:

1. Presentare l'informazione all'utente in modo chiaro e conciso, evitando termini tecnici o specialistici
2. Semplificare la struttura del compito
3. Offrire all'utente le scelte corrette, in una maniera che risulti ovvia
4. Organizzare ogni pagina in modo che l'utente riconosca la posizione e le azioni da compiere
5. Eliminare ogni ambiguità relativa alle conseguenze di un'azione (es. fare clic su cancella/rimuovi/compra)



6. Mettere la cosa più importante nella posizione giusta della pagina web o dell'applicazione web.
7. Fare in modo che l'utente abbia un rapido feedback (informazione di ritorno) ad ogni azione compiuta
8. Rendere la grafica accattivante ed interessante dal punto di vista visivo attraverso l'uso di diagrammi, tabelle, sezioni informative
9. Ridurre gli sforzi cognitivi dell'utente

L'obiettivo finale della web usability è quello di rendere invisibile la tecnologia sottostante, *trasparente* per l'utente che deve concentrare la propria attenzione e lo sforzo cognitivo principalmente sul compito, senza disperderli sull'utilizzo del mezzo. Per ottenere siti web usabili è necessario conoscere e coinvolgere gli utenti durante la fase della progettazione. La web usability va perciò considerata e valutata in ogni momento della progettazione: nelle fasi iniziali, in itinere e prima della pubblicazione; in ogni fase con strumenti diversi (riferimenti ai principi, interviste, confronti, simulazione d'uso, questionari e test su campione di utenti target) e con il coinvolgimento continuo dell'utente.

Indicazioni per l'usabilità

La web usability deve tener conto:

- del compito che l'utente deve svolgere (la navigazione delle pagine di un sito)
- dell'utente che lo svolge (target cui si rivolge il sito)
- all'ambiente d'uso (interfaccia, hardware, browser, velocità di connessione)

A questo proposito, la progettazione di siti web tiene conto degli apporti, in termini di esperienza, del lavoro di Jakob Nielsen, Bruce Tognazzini e Steve Krug che, con i loro lavori, hanno dettato euristiche e principi per la web usability. Dall'analisi della letteratura relativa all'argomento sono stati individuati sei requisiti che, nella loro interazione, permettono di realizzare un sito usabile:

- navigabilità
- tempi di attesa
- completezza dei contenuti
- comprensibilità delle informazioni
- efficacia comunicativa
- attrattiva grafica

Motore di ricerca

Un **motore di ricerca** (in inglese *search engine*) è un sistema automatico che analizza un insieme di dati spesso da esso stesso raccolti e restituisce un indice dei contenuti disponibili classificandoli in base a formule statistico-matematiche che ne indichino il grado di rilevanza data una determinata chiave di ricerca.

Uno dei campi in cui i motori di ricerca trovano maggiore utilizzo è quello dell'information retrieval e nel web.

Motori di ricerca per il web

Esistono numerosi motori di ricerca attivi sul web. Il più utilizzato, su scala mondiale (con un indice che supera gli 8 miliardi di pagine), è Google; molto usati anche Live e Bing (motori di ricerca di Microsoft), Yahoo!, Ask. Da segnalare il tentativo di creare il primo motore di ricerca europeo, Quaero concorrente di Google con una iniziativa franco-germanica. Il progetto, stimato attorno ai 400 milioni di dollari, è stato abbandonato dopo pochi mesi per la rinuncia da parte della compagnia tedesca.

Fra i motori di ricerca nati in Italia quelli maggiormente utilizzati nel nostro paese sono Arianna, attivo nel portale Libero e Virgilio. Tuttavia non sono veri motori di ricerca giacché si limitano a riutilizzare Google. Arianna e altri ne evidenziano chiaramente il logo, mentre Virgilio ne usa i risultati senza evidenziarne la fonte, limitandosi solo ad aggiungere alcuni propri risultati sponsorizzati.

La maggior parte dei motori di ricerca che opera sul web è gestito da compagnie private che utilizzano algoritmi proprietari e database tenuti segreti. Esistono comunque diversi tentativi di dar vita a motori di ricerca fondati sul software libero, alcuni esempi sono:

- Lucene
- Nutch
- Wikia Search
- YaCy
- OpenIndexer

Le fasi

Il lavoro dei motori di ricerca si divide principalmente in tre fasi:

- **analisi** del campo d'azione (tramite l'uso di crawler appositi);
- **catalogazione** del materiale ottenuto;
- **risposta** alle richieste dell'utente;

Catalogazione

Dopo l'analisi delle pagine, a seconda di criteri che variano da motore a motore, alcune di esse vengono inserite nel database e nell'indice del motore di ricerca.

La parte testuale archiviata durante la fase di analisi verrà in seguito analizzata per fornire le risposte alle ricerche degli utenti. Molti motori di ricerca sul web rendono anche disponibile una copia dei dati testuali di ogni pagina archiviata per quando la risorsa originale sia irraggiungibile: questa funzione è detta copia cache.

Risposta

Rispondere alle richieste degli utenti implica la necessità di elencare i siti in ordine di rilevanza rispetto alla richiesta ricevuta.

Per stabilire la rilevanza di un sito vengono cercati nel database quei documenti che contengono la parola chiave inserita dall'utente, dopodiché ogni motore di ricerca sfrutta propri algoritmi per classificare le pagine, controllando, per esempio, quante volte le parole chiave vengono ripetute, quanti link riceve quel documento, in quali punti della pagina sono poste le parole chiave, quanti siti del database contengono link verso quella pagina, o quante volte un utente ha visitato quel sito dopo una ricerca.

Risultati sponsorizzati

I motori di ricerca forniscono anche risultati sponsorizzati, ovvero mostrano in maggiore evidenza nelle SERP (Search Engine Result Pages, Pagine dei risultati dei motori di ricerca) siti web di aziende che pagano per risultare tra i primi risultati quando si cercano termini (detti keyword o parole chiave) che sono in relazione all'ambito di competenza dell'azienda stessa. I risultati sponsorizzati dei motori possono apparire anche sui siti che partecipano al loro programma di affiliazione. In particolar modo, Google permette di far apparire nelle proprie SERP (chiaramente distinti dai risultati "naturali") risultati a pagamento comprati con il programma AdWords. In aggiunta a questo offre anche un servizio di sponsorizzazione che si rivolge a tutti i siti che hanno determinati requisiti, chiamato AdSense. Google AdSense (spesso abbreviato con Google AS) usa le capacità del motore di ricerca di interpretare il tema della pagina in cui è posizionato l'apposito codice per fornire annunci a tema. Yahoo! Search ha annunciato l'arrivo di un programma analogo chiamato Panama.

Raffinazione della ricerca

La possibilità di raffinazione della ricerca varia da motore a motore, ma la maggior parte permette di utilizzare operatori booleani: ad esempio è possibile cercare "*Ganimede AND satellite NOT coppiere*" per cercare informazioni su Ganimede inteso come pianeta e non come figura mitologica.

Su Google e sui motori più moderni è possibile raffinare la ricerca a seconda della lingua del documento, delle parole o frasi presenti o assenti, del formato dei file (Microsoft Word, PDF, PostScript, ecc.), a seconda della data di ultimo aggiornamento, e altro ancora. È anche possibile cercare contenuti presenti in un determinato sito, ad esempio "*Ganimede site:nasa.gov*" cercherà le informazioni su Ganimede presenti sul sito della NASA.

Su Exalead si trova una parte speciale per raffinare la ricerca più intuitivamente.

Motori di ricerca più utilizzati

- Google, oltre ad essere di gran lunga il più usato direttamente, è anche quello con il maggior numero di siti che ne utilizzano il database, anche a causa del programma di sponsorizzazione "AdSense per la ricerca" che appartiene al più vasto programma AdSense.
- Yahoo!
- Windows Live Search (*Live Search* è il motore creato dalla Microsoft ed usato da Microsoft Network, meglio noto come MSN).
- Ixquick (*Ixquick* è uno dei pochi motori di ricerca al mondo a garantire la privacy, poiché non conserva né registra gli ip di chi effettua ricerche).
- Ask.com, usato dalla versione italiana di Excite per un rapporto di partnership.
- FileByType è una raccolta di moduli di ricerca basati sulle categorie.

Le prime quattro tecnologie proprietarie sono utilizzate da una quantità sterminata di provider e di metamotori di ricerca (il cui nome deriva proprio dal pescare i propri risultati da più motori, come il metamotore incorporato nel portale Excite). A più riprese Microsoft ha provato a comprare Yahoo!, cosa che avrebbe portato ad un ulteriore accorpamento e riduzione delle tecnologie proprietarie in campo, a causa della fusione di Live (il motore della Microsoft) e Yahoo! Search. L'ultimo rifiuto di Yahoo è riportato dal Wall Street Journal in data 6/5/2007 (a fronte di un'offerta di circa 50 miliardi di dollari da parte della società fondata da Bill Gates).

L'unico motore con una tecnologia proprietaria in qualche modo affiancabile come utenza ai quattro big è il cinese Baidu www.baidu.com (la cui inferiorità tecnologica è palese Wikipedia:Cita le fonti, ma che attinge ad un bacino di utenza tanto vasto quanto in crescita, sebbene appartenente ad una sola nazione).

Parte delle descrizioni dei siti presentate nei risultati di ricerca dai principali motori sono importate da DMOZ - ODP. ODP, acronimo di Open Directory Project, non è un motore ma una open directory (basata su listing e recensioni fatte da esseri umani, anche se esistono alcuni meccanismi automatici per eliminare i siti estinti). È stata creata da Netscape, a sua volta comprata da AOL nel 1998 per quasi 25 miliardi di dollari, ed appartiene tuttora ad AOL (che è la divisione internet di Time Warner, divisione nella quale Google ha una modesta partecipazione azionaria).

In Italia

- Google è utilizzato anche da Libero, Yootube-find, Arianna e Tiscali (per un rapporto di partnership). In Italia il divario fra l'utilizzo di Google e degli altri è talmente vasto che, pur essendo i dati ovviamente variabili, la percentuale delle ricerche effettuate su Google è pari ad un multiplo della somma di tutti gli altri motori di ricerca messi assieme.
- Yahoo! è utilizzato anche da AltaVista, che venne acquistata indirettamente da Yahoo! quando quest'ultima acquisì Overture, nonché da Kataweb per un rapporto di partnership.
- Virgilio.it è legato a Google da un rapporto di partnership da diversi anni e ne importa (cosa immediatamente verificabile da chiunque) i risultati di ricerca con minime modifiche come aggiungere propri risultati sponsorizzati ed accorpare quelli che fanno riferimento allo stesso dominio.
- istella è un motore di ricerca sviluppato da Tiscali.

Iper testo

Un **ipertesto** è un insieme di documenti messi in relazione tra loro per mezzo di parole chiave. Può essere visto come una rete; i documenti ne costituiscono i nodi. La caratteristica principale di un ipertesto è che la lettura può svolgersi in maniera non lineare: qualsiasi documento della rete può essere "il successivo", in base alla scelta del lettore di quale parola chiave usare come collegamento. È possibile, infatti, leggere all'interno di un ipertesto tutti i documenti collegati dalla medesima parola chiave. La scelta di una parola chiave diversa porta all'apertura di un documento diverso: all'interno dell'ipertesto sono possibili praticamente infiniti percorsi di lettura.

L'ipertesto informatico è la versione di ipertesto più usata e più diffusa oggi. Il computer ha automatizzato il passaggio da un documento all'altro. I documenti sono leggibili a video grazie a un'interfaccia elettronica, le parole chiave in esso contenute appaiono marcate (sottolineate oppure evidenziate, ecc) in maniera da renderle riconoscibili.



Prototipo di una console per ipertesto del 1969

L'ipertesto consiste in un collegamento alla parola chiave (opportunamente evidenziata allo scopo), che talvolta appare nello schermo anche sotto forma di icona o immagine. Selezionando o posizionandosi su tale parola o oggetto e facendo clic con il mouse oppure dando l'invio (per navigazione basata sulla tastiera) si ottiene come conseguenza l'apertura di un altro documento, che si può trovare sullo stesso server o altrove. Quindi le parole chiave funzionano come collegamenti ipertestuali (*hyperlink* in inglese), che consentono all'utente di navigare verso informazioni aggiuntive.

Dopo la nascita del World Wide Web (1993) l'ipertesto ha avuto un notevolissimo sviluppo. Tutto il web, infatti, è stato concepito dal suo inventore, l'inglese Tim Berners-Lee, come un ipertesto globale in cui tutti i siti mondiali possono essere consultati da tutti. La pagina web è il singolo documento e la "navigazione" è il passaggio da un sito all'altro tramite i "link" (ma anche da una pagina all'altra dello stesso sito o in un'altra parte della stessa pagina). L'interfaccia per visualizzare i siti web (e le pagine ipertestuali contenute) è il browser.

Collegamenti ipertestuali

Il World wide web utilizza il linguaggio HTML (*HyperText Markup Language*) per definire all'interno del testo istruzioni codificate per il suo funzionamento. Se una pagina del web contiene informazioni su una parola (o termine di qualsiasi natura) utilizzata su un nuovo documento HTML è possibile creare un legame, chiamato *collegamento ipertestuale*, tra questo nuovo documento e la pagina già esistente. Per esempio, nel caso di un indice, allo stesso modo è possibile indirizzare parole di un documento a parti del medesimo.

Se si seleziona con il mouse la parola sottolineata dal tag di rimando o collegamento si visualizza il testo che contiene la definizione o l'informazione aggiuntiva.

Storia

Agostino Ramelli, ingegnere svizzero-italiano nato nel 1531, ideò la "ruota dei libri", un leggio multiplo rotante, ideato per consentire l'agevole lettura contemporanea di più testi e che si può considerare una prima forma di ipertesto.

Nel 1929 il regista sovietico Sergej Michajlovič Ejzenštejn nella prefazione di un suo libro si rammaricò di non potere raccogliere in un volume i suoi saggi in modo che fossero "percepiti tutti insieme simultaneamente" per "comparare ciascun saggio direttamente con gli altri, di passare dall'uno all'altro avanti e indietro". Per Ejzenštejn l'unico tipo di libro in grado di soddisfare queste condizioni dovrebbe avere la forma di una sfera, "dove i settori coesistono simultaneamente in forma di sfera, e dove, per quanto lontani possano essere, è sempre possibile un passaggio diretto dall'uno all'altro attraverso il centro della sfera".⁴⁷

Il concetto di ipertesto è stato rivalutato dall'informatica, a cui si è interessata fin dalle sue origini. Nel 1945 l'ingegnere americano Vannevar Bush scrisse un articolo intitolato *Come potremmo pensare* nel quale descriveva un sistema di informazione interconnesso chiamato Memex. Molti esperti non considerano tuttavia il Memex un vero e proprio sistema ipertestuale. Nonostante ciò, il Memex è considerato la base degli ipertesti perché il saggio di Vannevar Bush influenzò direttamente Ted Nelson e Douglas Engelbart, universalmente riconosciuti come gli inventori dell'ipertesto.

⁴⁷In una nota contenuta in: Sergej Michajlovič Ejzenštejn, *Il montaggio*, Venezia, Marsilio Editori, 1986. ISBN 88-317-4932-3

Secondo Ted Nelson, che coniò il termine *hypertext* nel 1965, la definizione riveste un significato più ampio, coinvolgendo qualsiasi sistema di scrittura non lineare che utilizza l'informatica.

Nel 1980 il programmatore Bill Atkinson realizzò per la Apple *HyperCard*, un'applicazione software che gestiva in maniera semplice grandi quantità di informazioni sotto forma di testo o di immagini, dotato di un avanzato linguaggio di programmazione ipertestuale, *HyperTalk*. Malgrado fosse disponibile solo per la piattaforma MacOS, *HyperCard* divenne uno dei più diffusi sistemi di realizzazione di ipertesti prima dell'avvento del World Wide Web,

Nel 1989 Tim Berners-Lee, ricercatore inglese del CERN, inventò il World Wide Web con l'intento di dare una risposta alla necessità espressa dalla comunità scientifica di un sistema di condivisione delle informazioni tra diverse università e istituti di tutto il mondo. All'inizio del 1993 il National Center for Supercomputing Applications (NCSA) all'Università dell'Illinois rese pubblica la prima versione del loro browser *Mosaic*. Mosaic girava in ambiente X Window, popolare nella comunità scientifica, e offriva un'interfaccia di facile utilizzo. Il traffico web esplose, passando da soli 500 web server noti nel 1993 a oltre 10mila nel 1994 dopo la pubblicazione della versione che girava sia in ambiente Windows che MacOS.

Database

In informatica, il termine *database*, **banca dati** o **base di dati**, indica un archivio dati, o un insieme di archivi, in cui le informazioni in esso contenute sono strutturate e collegate tra loro secondo un particolare *modello logico* (relazionale, gerarchico, reticolare o a oggetti) e in modo tale da consentire la gestione/organizzazione efficiente dei dati stessi grazie a particolari applicazioni software dedicate (DBMS), basate su un'architettura di tipo client-server, e ai cosiddetti query language per l'interfacciamento con le richieste dell'utente (query di ricerca o interrogazione, inserimento, cancellazione ed aggiornamento).

Il termine *database* può dunque indicare contemporaneamente:

- l'archivio a livello fisico (hardware) cioè il sistema con i supporti di memorizzazione che contengono i dati stessi e il processore per l'elaborazione di questi (database server);
- l'archivio a livello logico cioè i dati strutturati e la parte software cioè il *database management system* (DBMS) ovvero quella vasta categoria di applicazioni che consentono la creazione, manipolazione (gestione) ed interrogazione efficiente dei dati.

Informalmente e impropriamente, la parola "database" viene spesso usata quasi unicamente per indicare il *database management system* (DBMS) riferendosi dunque alla sola parte software.

Nei database più moderni ovvero quelli basati sul modello relazionale i dati vengono suddivisi per argomenti (in apposite tabelle) e poi tali argomenti vengono suddivisi per categorie (campi) con tutte le possibili operazioni di cui sopra. Tale suddivisione e funzionalità rende i database notevolmente più efficienti rispetto ad un archivio di dati creato ad esempio tramite file system di un sistema operativo su un computer almeno per la gestione di dati complessi.

La diffusione dei database, e dei relativi sistemi DBMS di gestione, nei sistemi informativi moderni è enorme e capillare essendo un componente base di un sistema informativo attraverso il rispettivo sistema informatico: si pensi a tutte le attività commerciali di gestione di magazzino, gestione clienti, a sistemi di immagazzinamento di dati personali o pubblici nella pubblica amministrazione e nelle imprese private (es. banche e aziende), contabilità ecc. La gestione e lo sviluppo dei database attraverso DBMS è diventata nel tempo una branca a tutti gli effetti dell'informatica moderna.

Gestione delle informazioni

La base di dati deve contenere anche le informazioni sulle loro rappresentazioni e sulle relazioni che li legano. Spesso, ma non necessariamente, una base dati contiene le seguenti informazioni:

- Strutture dati che velocizzano le operazioni frequenti, di solito a spese di operazioni meno frequenti.
- Collegamenti con dati esterni, cioè riferimenti a file locali o remoti non facenti parte del database.
- Informazioni di sicurezza, che autorizzano solo alcuni profili utente ad eseguire alcune operazioni su alcuni tipi di dati.
- Programmi che vengono eseguiti, automaticamente o su richiesta di utenti autorizzati, per eseguire elaborazioni sui dati. Un tipico automatismo consiste nell'eseguire un programma ogni volta che viene modificato un dato di un certo tipo.

In un sistema informatico, una base di dati può essere usata direttamente dai programmi applicativi, interfacciandosi con il supporto di memorizzazione (insomma, agendo direttamente sui file). Questa strategia era quella adottata universalmente fino agli anni sessanta, ed è tuttora impiegata quando i dati hanno una struttura molto semplice, o quando sono elaborati da un solo programma applicativo.

Dalla fine degli anni sessanta, tuttavia, per gestire basi di dati complesse condivise da più applicazioni si sono utilizzati appositi sistemi software, detti **sistemi per la gestione di basi di dati** (in inglese "Database Management System" o "DBMS"). Uno dei vantaggi di questi sistemi è la possibilità di non agire direttamente sui dati, ma di vederne una rappresentazione concettuale.

La ricerca nel campo delle basi di dati studia le seguenti problematiche:

- Progettazione di basi di dati.
- Progettazione e implementazione di DBMS.
- Interpretazione (o analisi) di dati contenuti in database.

Le basi di dati spesso fanno uso di tecnologie derivate da altre branche dell'informatica. È usuale utilizzare tecniche derivate dall'intelligenza artificiale, come ad esempio il data mining, per cercare di estrarre relazioni o più in generale informazioni presenti nelle banche dati ma non immediatamente visibili.

Struttura o modello logico

Le basi di dati possono avere varie strutture, di solito, in ordine cronologico:

1. *gerarchica* (rappresentabile tramite un albero – anni sessanta),
2. *reticolare* (rappresentabile tramite un grafo – anni sessanta),
3. *relazionale* (attualmente il più diffuso, rappresentabile mediante tabelle e relazioni tra esse – anni settanta),
4. *ad oggetti* (estensione alle basi di dati del paradigma "Object Oriented", tipico della programmazione a oggetti – anni ottanta),
5. *semantica* (rappresentabile con un grafo relazionale – inizio anni duemila).

Il formato XML, oltre che per scambi di dati su web, si sta diffondendo per la definizione di vere e proprie basi di dati. XML ha una struttura gerarchica, pare quindi un "ritorno alle origini" dei modelli di dati.

Un requisito importante di una buona base dati consiste nel non duplicare inutilmente le informazioni in essa contenute: questo è reso possibile dai gestori di database relazionali (teorizzati da Edgar F. Codd), che consentono di salvare i dati in tabelle che possono essere collegate.

La funzionalità di un database dipende in modo essenziale dalla sua progettazione: la corretta individuazione degli scopi del database e quindi delle tabelle, da definire attraverso i loro campi e le relazioni che le legano, permette poi una estrazione dei dati più veloce e, in generale, una gestione più efficiente.

Linguaggi per basi di dati

Nell'ambito dell'utilizzo/amministrazione del database, una qualunque operazione sulla base di dati da parte dell'utente/amministratore su DBMS è ottenibile tramite un opportuno linguaggio attraverso un DBMS manager ad interfaccia grafica o ad interfaccia a riga di comando. In generale è possibile distinguere più linguaggi per basi di dati (ciascuno con lessico e sintassi che definiscono tutte le istruzioni possibili) a seconda del loro utilizzo o finalità a partire dalla creazione/progettazione, gestione, ristrutturazione, sicurezza, fino all'interrogazione del database stesso:

- Data Definition Language (DDL) - consente di definire la struttura o organizzazione logica della base di dati (schema di database) e le autorizzazioni per l'accesso.
- Device Media Control Language (DMCL) - permette di controllare i supporti (memoria di massa) dove vengono memorizzati i dati.
- Data Manipulation Language (DML) - permette di inserire, modificare, cancellare i dati (transazione).
- Data Control Language (DCL) - permette di gestire gli utenti e i permessi.
- Query language (QL) - permette di interrogare il database, cioè di leggere i dati.

La sintassi di questi linguaggi varia a seconda del particolare DBMS, mantenendo tra loro un'uniformità concettuale.

Inoltre è possibile suddividere i linguaggi come:

- Linguaggi testuali interattivi, come l'SQL, di cui sono stati pubblicati diversi standard, che rappresenta attualmente il linguaggio più utilizzato.
- Linguaggi testuali interattivi immersi in linguaggi di programmazione comuni, quali C, Basic ecc.
- Linguaggi testuali interattivi immersi in linguaggi di programmazione proprietari.
- Linguaggi grafici e *user-friendly*, come QBE (Query By Example), che possono essere utilizzati anche dai meno esperti.

Server di database

Un Server di database è la parte del DBMS (e, per estensione, il server su cui il programma opera) che si occupa di fornire i servizi di utilizzo del database ad altri programmi e ad altri computer secondo la modalità client/server. Il server memorizza i dati, riceve le richieste dei client ed elabora le risposte appropriate.

Tra i più diffusi DBMS open source troviamo:

- MySQL
- Firebird SQL
- PostgreSQL

I più diffusi sistemi commerciali sono:

- Oracle
- IBM DB2
- Microsoft SQL Server
- Sybase

I database server sono complessi sistemi software concepiti in un linguaggio binario, oltre che per memorizzare i dati, anche per fornire un accesso rapido ed efficace a una pluralità di utenti contemporaneamente e garantire protezione sia dai guasti che dagli accessi indebiti (sicurezza o protezione del database)