

Made by Enzo maioli per <http://www.wintricks.it/>

I N D I C E G E N E R A L E

Come nasce internet	pag.4
Commutazione di pacchetto	pag.5
Protocollo di rete e comunicazione	pag.5
Requests for comment (RFC)	pag.5
Esempi di RFC che non specifica protocolli	pag.6
Tipi di rete	pag.6
Lan	pag.6
Man	pag.6
Wan	pag.7
Reti paritetiche	pag.7
Reti basate su server	pag.7
Tipologia di rete	pag.8
Reti a bus	pag.8
Reti a stella	pag.9
Reti ad anello	pag.10
Reti a doppio anello	pag.11
Reti a commutazione di pacchetto	pag.11
I sistemi aperti	pag.12
Il modello a strati	pag.12
Il modello ISO/OSI	pag.12
Strato fisico	pag.13
Strato del collegamento dati	pag.13
Strato della rete	pag.13
Strato del trasporto	pag.13
Strato della sessione	pag.14
Strato della presentazione	pag.14

Strato dell'applicazione	pag.14
Modello a strati TCP/IP	pag.15
Internet	pag.16
Trasporto	pag.16
Confronto TCP/IP modello OSI	pag.17
Protocollo nello strato internet	pag.17
ARP	pag.18
Esempi su ICMP e ping – TTL	pag.19
L'indirizzo IP	pag.19
Rappresentazione binaria e dot notation	pag.20
Classi di indirizzi IP	pag.21
Maschere di sottorete	pag.22
Indirizzi di rete locale	pag.22
Telnet	pag.22
FTP	pag.23
FTP e FTP anonimo	pag.23
http	pag.24
Posta Elettronica E-Mail	pag.24
SMTP	pag.25
POP3	pag.25
Uuencode	pag.25
World Wide Web (WWW)	pag.26
Browser	pag.26
Link	pag.27
Formati dei file su web	pag.27
Pagina web	pag.27
Gli indirizzi internet (URL)	pag.28
URL con numeri di porta	pag.28
HTML	pag.29

Un po' di storia dell'HTML	pag.29
SGML il padre dell'html	pag.29
Evoluzione dell'HTML	pag.29
Html funzioni principali	pag.30
I TAG	pag.30

Come nasce Internet

1957	Nasce ARPA (Advanced Research Projects Agency)
1962	Su incarico della USAF (United States Air Force) viene progettata una rete a commutazione di pacchetto
1968	BBN (Bolt, Beranek e Newman) riceve l'incarico da ARPA di costruire una rete a commutazione di pacchetto. Nasce ARPAnet. Inizialmente sono collegati 4 nodi.
1972	ARPA net si allarga a 32 nodi
1973	Inizia lo sviluppo di un insieme di protocolli di rete che diventeranno nel tempo TCP/IP
1976	La Xerox sviluppa Ethernet che permise lo sviluppo di reti locali (LAN, Local Area Network)
1982	Nasce il termine Internet . Fino ad allora si chiamavano internet un gruppo di reti interconnesse con protocollo TCP/IP. Con Internet si definisce l'insieme delle internet TCP/IP connesse tra di loro
1983	ARPAnet viene divisa in ARPAnet e MILNET. Viene introdotto un metodo più efficiente per la risoluzione dei nomi e degli indirizzi di rete: il DNS (Domain Name Space)
1985	La National Science Foundation inizia a stendere le nuove linee a 1,544 Mbps. La nuova generazione di reti viene denominata NSFNET.
1996	Il traffico in questi anni è continuato a crescere in modo esponenziale, si passa ad una nuova dorsale a 145 Mbps: ATM (Asynchronous Transmission Mode)
1996	Il 30 aprile questa dorsale diventa privata

Commutazione di pacchetto

Con *commutazione di pacchetto* si intende la suddivisione dei dati da spedire in cosiddetti *datagrams* (o pacchetti).

Ogni pacchetto è etichettato in modo da contenere gli indirizzi di origine e di destinazione del pacchetto.

In caso di indisponibilità di una connessione iniziale i pacchetti devono essere in grado di trovare una strada alternativa per giungere a destinazione.

Il ricevente nel caso di mancata ricezione di un pacchetto può richiedere al mittente la rispedizione del pacchetto andato perduto.

Protocollo di rete e di comunicazione

Il *protocollo di rete* è un linguaggio utilizzato per le comunicazione tra due host remoti.

Il protocollo definisce il modo in cui devono essere impacchettati i dati per la trasmissione in rete, in modo tale da poter essere spaccchettati nella maniera opportuna dal ricevente.

Analogamente un *protocollo di comunicazione* stabilisce il linguaggio di comunicazione tra due applicativi posti eventualmente su computer remoti, al fine di poter dialogare proficuamente.

Esempi di protocolli di comunicazione attualmente largamente diffusi e noti sono:

FTP	File Transfer Protocol
HTTP	Hyper Text Transmission Protocol
POP	Post Office Protocol (attualmente largamente usato nella versione 3 POP3)
SMTP	Simple Mail Transfer Protocol

Tutti questi protocolli di comunicazione si basano sul protocollo di rete TCP/IP.

Requests for Comment

Le Requests for Comment (RFC) sono letteralmente delle richieste di commento ad una proposta.

L'Internet Architecture Board pubblica tutti gli standard TCP/IP sotto forma di RFC.

Sebbene tutti gli standard siano pubblicati come RFC, non tutte le RFC sono specifiche di standard.

Alcune delle specifiche dei protocolli maggiormente usati sono:

<i>RFC</i>	<i>PROTOCOLLO</i>
RFC959	FTP
RFC854, RFC855	TELNET
RFC821	SMTP
RFC1939	POP3

Esempio di RFC che non specifica protocolli

L'esempio seguente riporta una RFC (la 527) che non solo non specifica alcun protocollo ma che addirittura è la presa in giro del linguaggio gergale usato dagli amministratori di rete.

Nel leggere questa RFC si tenga conto che è stata pubblicata nel maggio 1973, per cui termini come per esempio *bit* e *byte* che allora sembravano bizzarrie di tecnici affezionati al loro gergo, oggi sono di uso talmente comune che sono comprensibili anche ai non addetti ai lavori.

Tipi di Rete

Le reti vengono divise a seconda della loro estensione geografica in:

LAN: Local Area Network
MAN: Metropolitan Area Network
WAN: Wide Area Network

LAN

Le LAN (Local Area Network) sono il tipo di rete più ampiamente diffuso negli uffici. Esse si estendono su un piano di un edificio, o su intero edificio. Una LAN si può anche arrivare ad estendersi su più edifici vicini.

- Tutti i siti sono vicini tra di loro
- Ampia velocità di trasmissione
- Bassa frequenza di errori
- Tutti i dati fanno parte della rete locale

MAN

Le MAN (Metropolitan Area Network) sono reti che collegano aree metropolitane quali

- Pubbliche amministrazioni,
- Università,
- Reti civiche,
- Agenzie di servizi.

Sono caratterizzate da:

- Alte velocità di trasmissione
- Costi elevati

In effetti una rete MAN consente agli utenti dislocati in punti geografici diversi di utilizzare le risorse condividendole come se facessero tutti parte della stessa rete locale. Ovvio comunque che le reti MAN hanno una complessità maggiore rispetto alle LAN; una conseguenza dell'impiego di linee telefoniche ad alta velocità o di hardware specializzato è il maggior costo della MAN rispetto a una LAN.

WAN

Le WAN (Wide Area Network) sono nella maggior parte la combinazione di una serie di reti su area locale (LAN) opportunamente connesse tra di loro mediante collegamenti aggiuntivi per permettere la comunicazione tra di loro.

Sono nate per collegare tra di loro siti di ricerca distanti tra di loro.

Sono caratterizzate da:

- Costi bassi
- Velocità basse
- Utilizzano linee telefoniche standard come mezzo di comunicazione principale

Reti paritetiche

Le reti paritetiche operano senza server dedicati sulla rete, di conseguenza ciascuna macchina funge contemporaneamente sia da client che da server.

Questo tipo di rete si adatta bene a piccole organizzazioni in cui non vi sono particolari problemi di sicurezza, ed in cui generalmente non è possibile accedere dall'esterno.

È il caso tipico di uffici in cui ciascun computer è autosufficiente, quindi dotato di stampante e di qualunque altra cosa possa servire per poter operare.

In questi casi spesso ciascun impiegato oltre che essere un utente è anche amministratore del proprio computer.

Reti basate su server

Nelle *reti basate su server*, almeno una macchina è dedicata alla funzione di server.

Tra le varie funzioni che un *server* può svolgere, si possono citare:

- *Server di file e di stampa*, dedicati a fornire un'area sicura per il deposito dei propri dati; forniscono inoltre l'accesso ad una stampante gestendone la coda di stampa
- *Server di applicazioni*, dedicati a fornire la parte server di applicazioni client-server; esempi di tali server sono i server HTTP ed i server di database
- *Server di posta*
- *Server di sicurezza*, dedicato a gestire la sicurezza sulla rete locale controllandone opportunamente gli accessi; esempi di tali server sono i *firewall* ed i *server proxy*

Tipologia di rete

Nell'implementazione di una rete locale, bisogna prendere in considerazione diversi aspetti della rete, tra cui la collocazione del computer, l'ubicazione dei cavi, l'hardware richiesto per la connessione.

Al momento attuale si utilizzano comunemente quattro topologie di rete:

- reti a bus
- reti a stella
- reti ad anello
- reti a doppio anello

Reti a bus

La rete a bus è il metodo più semplice per permettere la connessione tra più computer.

Consiste di un singolo cavo che connette tutti i computer. Quando un host deve comunicare con un altro host immette sulla rete i propri dati.

Questi dati arrivano a tutti computer sulla rete e ciascuno di essi quindi, li esamina per individuare se sono diretti a lui. Se non lo sono le scarta, e così via fino a quando non raggiungono l'host destinatario.

Se più host iniziano l'invio contemporaneamente avviene una *collisione* e gli host lo possono scoprire.

Per il collegamento di host su una rete a bus si utilizzano i connettori BNC.

Uno degli svantaggi di questo tipo di rete, è che la disconnessione di un computer dalla rete può portare al blocco dell'intera rete.

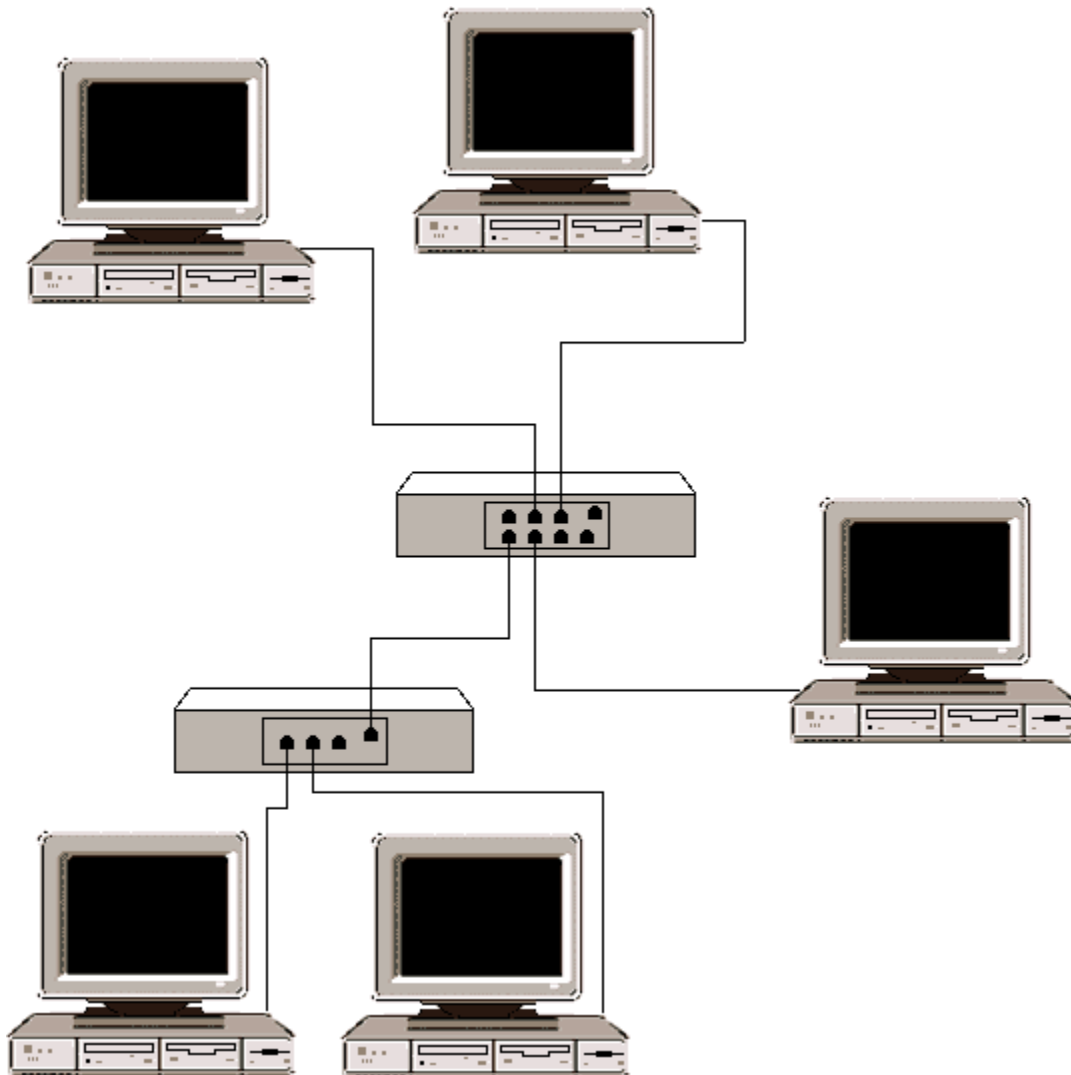
La rete inoltre deve essere sempre terminata da un apposito *terminatore* (volgarmente *tappo*) per impedire che la rete lasciata aperta si blocchi.

È probabilmente il tipo di rete più economico, ma supporta un numero limitato di host.

Reti a stella

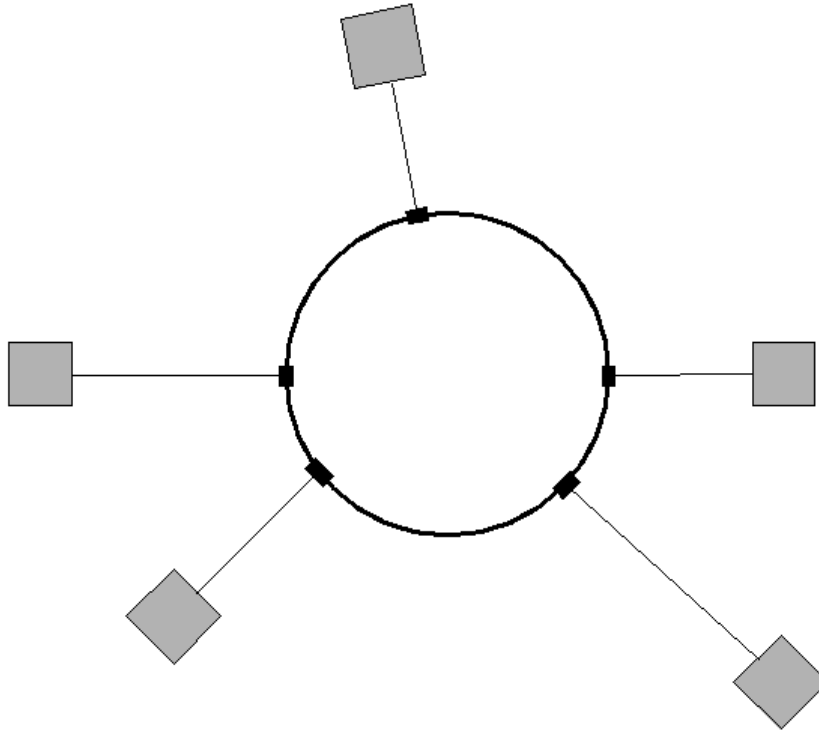
Tipo LAN centralizzata in cui i nodi, costituiti da stazioni (workstation) di lavoro, sono collegati a un computer centrale o hub (vedi figura), dispositivo che consente alla rete di avere un unico punto di collegamento per tutti gli altri dispositivi. I costi di collegamento sono più elevati rispetto ad altre topologie di rete; inoltre, dato che tutti i messaggi passano attraverso l'hub, in caso di un suo malfunzionamento l'intera rete va fuori uso.

Rispetto alla rete a bus ha però il vantaggio che la disconnessione di un singolo computer non comporta alcun impatto sul buon funzionamento della rete.



Reti ad anello

Le reti ad anello sono un tipo di rete locale decentralizzata i cui nodi, costituiti da stazioni di lavoro periferiche condivise e server di file, sono collegati da un cavo chiuso ad anello.



Ciò permette una maggiore velocità di trasferimento dati (data transfer rate) rispetto a una rete a bus, che raggiunge 16 MBPS (contro i 10 dello standard Ethernet). Come in una rete a bus, ogni stazione invia i suoi messaggi a tutte le altre; ogni nodo ha un indirizzo unico, e la sua circuiteria di ricezione monitorizza costantemente il bus in attesa di messaggi, ignorando quelli inviati agli altri.

I dati viaggiano sulla rete con un metodo chiamato a passaggio di testimone o di gettone (token-ring).

Il gettone per trasmettere i dati è unico, ma questo non rallenta la velocità di trasmissione. Si consideri che su una rete di circa 400 m il gettone può fare il giro dell'intero anello circa 5000 volte in un secondo.

Reti a doppio anello

Le reti a doppio anello sono simili a quelle ad anello, avendo la sostanziale differenza di utilizzare due anelli anziché uno:

- un anello primario e
- un anello secondario

Una ulteriore differenza è l'utilizzo di fibre ottiche (FDDI – Fiber Distributed Data Interface).

Nelle normali condizioni i dati fluiscono solo sull'anello primario, utilizzando il secondario solo in caso di guasto del primario.

Ogni computer deve ovviamente essere connesso ad entrambi gli anelli, per poter commutare dal primario al secondario in caso di guasto.

Generalmete in questo tipo di rete non tutti i computer sono collegati ad entrambi gli anelli.

Reti a commutazione di pacchetto

Le reti a *commutazione di pacchetto* consentono di trasmettere dati su una qualsiasi area geografica con una connessione *chiunque a chiunque*.

In una rete a commutazione di pacchetto l'informazione da trasmettere è suddivisa in pacchetti di dimensione abbastanza piccola; ad ognuno di essi viene aggiunta un'intestazione che contiene tutta l'informazione necessaria affinché il pacchetto sia inoltrato alla sua destinazione finale. I pacchetti sono inviati individualmente attraverso la rete e vengono poi riassemblati nella loro forma originale quando arrivano sul computer destinazione.

Poiché ogni pacchetto porta con sé la sua identificazione, una data rete può trasportare nello stesso tempo pacchetti provenienti da computers differenti. La commutazione di pacchetto permette quindi a più utenti di inviare informazioni attraverso la rete in modo efficiente e simultaneo, risparmiando tempo e costi sulle linee telefoniche, sulle connessioni radio e via satellite.

E poiché i pacchetti possono prendere strade alternative sulla rete, la trasmissione dei dati è facilmente mantenuta anche se parti della rete sono danneggiate o non funzionano efficacemente.

I sistemi aperti

Il concetto di sistema aperto deriva dalla necessità di standardizzazione.

Nella progettazione di una rete ci si può trovare di fronte ad una serie di prodotti offerti da fornitori diversi, che potrebbero vincolare definitivamente al fornitore scelto qualsiasi scelta progettuale futura.

Un sistema aperto permette di poter cambiare fornitore senza dover necessariamente riprogettare l'intera rete.

Il TCP/IP è un esempio di sistema aperto per un insieme di protocolli.

Grazie alle RFC tutti gli standard del TCP/IP sono completamente definiti.

Il modello a strati

Nella tecnologia di rete si utilizzano spesso i modelli a strati, per poter rappresentare le varie funzioni di messa in rete che bisogna svolgere.

I motivi principali per l'utilizzo di un tale modello sono:

- Suddividere tutte le funzioni di una operazione di rete in elementi meno complessi
- Consentire ai fornitori di concentrare la propria attenzione su aree specifiche
- Far sì che gli strati non siano influenzati dai cambiamenti che avvengono negli altri strati, essendo ben definita l'interfaccia tra di loro
- Fornire una piattaforma per lo sviluppo delle reti

Progettare in un modello stratificato significa preoccuparsi solo dei due strati immediatamente adiacenti, e tralasciare completamente gli altri.

Il modello ISO/OSI

Agli inizi degli anni 80 la International Standard Organization (ISO) vide la necessità di sviluppare un modello di rete che aiutasse i fornitori a creare soluzioni di messa in rete che potessero operare insieme.

Sviluppò quindi un modello oggi universalmente conosciuto come OSI (Open System Interconnection).

Il modello OSI è costituito da sette strati distinti:

- strato dell'applicazione
- strato della presentazione
- strato della sessione
- strato del trasporto
- strato della rete
- strato del collegamento dati
- strato fisico

Strato fisico

Lo strato fisico definisce le correnti elettriche, gli impulsi fisici o gli impulsi ottici che sono coinvolti nel trasporto dei dati dalla NIC (Network Interface Card) di un host al sistema di comunicazione.

I requisiti e le caratteristiche necessarie per la trasmissione sono documentate in genere in standard tipo V.35 o RS-232.

In prati lo strato fisico è il responsabile dei bit da un computer all'altro.

Strato del collegamento dati

Questo strato si occupa dell'invio dei *frame* dei dati dallo strato della rete a quello fisico.

Quando riceve i bit dallo strato fisico, li traduce in *frame* di dati.

Un frame comprende in genere le seguenti componenti:

- ID del destinatario. Questo ID è in genere l'indirizzo MAC* dell'host di destinazione o del gateway predefinito.
- ID del mittente. In genere è il MAC del mittente.
- Informazioni di controllo. Includono informazioni quali l'effettivo tipo di frame e notizie riguardo l'instradamento e la segmentazione.
- CRC (Cyclic Redundancy Check). Effettua la correzione degli errori e verifica che il frame di dati sia arrivato intatto all'host destinatario.

Strato della rete

Lo strato della rete determina il modo migliore per spostare i dati da un host all'altro.

Gestisce l'indirizzamento dei messaggi e la traduzione degli indirizzi logici (es. gli indirizzi IP) in indirizzi fisici (gli indirizzi MAC).

Strato del trasporto

Lo strato del trasporto segmenta e riassembla i dati in un flusso di dati.

Provvede ad una connessione tra l'host mittente e l'host destinatario.

I dati da trasferire sono spezzettati in segmenti e spediti al destinatario numerandoli sequenzialmente.

Il destinatario, alla ricezione dei segmenti, invia un segnale di avvenuta trasmissione.

Nel caso di avvenuta ricezione di un segmento, il destinatario può richiederne la ritrasmissione.

In questo modo sussiste il controllo degli errori nel trasporto dei dati.

* Il MAC (Media Access Control) è un numero esadecimale univoco di 12 cifre assegnato dal costruttore dell'hardware. È a carico del costruttore assicurare l'univocità del MAC.

Strato della sessione

Lo strato della sessione consente a due applicazioni su host *eventualmente* separati di stabilire una connessione per la comunicazione, chiamata sessione.

La sessione assicura che i messaggi inviati dall'uno all'altro siano ricevuti con un alto grado di attendibilità.

La sessione svolge funzioni di sicurezza, per assicurare che i due host siano autorizzati a comunicare attraverso la rete.

Alcuni esempi di di protocolli ed interfacce che operano a livello di sessione sono:

- **Winsock.** È una interfaccia di programmazione che stabilisce le porte i protocolli e gli indirizzi di due host che si apprestano a comunicare su rete.
- **Remote Procedure Call (RPC).** È un meccanismo che consente ad un host client di costruire una richiesta che verrà poi eseguita su un host server al livello di sicurezza del client.
- **X-Window.** Permette a terminali intelligenti di comunicare con computer UNIX come se fossero direttamente attaccati.

Strato della presentazione

Lo strato di presentazione determina il modo in cui i dati sono formattati nello scambio tra due computer in rete.

I dati ricevuti dallo strato dell'applicazione sono tradotti in un formato intermedio comunemente riconosciuto.

Lo strato di presentazione è responsabile anche di per le traduzioni e le codifiche dei dati e conversioni dei set di caratteri e dei protocolli.

Alcuni formati di presentazione gestiti dallo strato di presentazione sono:

ASCII. L'*American Standard Code for Information Interchange* è un set di caratteri codificati ad 8 bit, usato per definire tutti i caratteri alfanumerici.

EBCDIC. L'*Extended Binary Coded Decimal Interchange Code* è il metodo di rappresentazione dei testi usato abitualmente sui MainFrame e sui Mini dall'IBM.

XDR. L'*eXtended Data Representation* viene usata da applicazioni come **NFS** e **NIS** per fornire un formato universale per la trasmissione di testi tra due computer che si servono di due rappresentazioni diverse (per es. ASCII ed EBCDIC).

Strato dell'applicazione

Lo strato dell'applicazione consente ai programmi di accedere ai servizi di rete.

Per usare lo strato dell'applicazione, un programma deve avere una componente della comunicazione che richieda risorse di rete.

- **Posta Elettronica.** Lo strato dell'applicazione consente ai programmi di accedere ai servizi di comunicazione di rete. Esempi di prodotti di questo tipo sono Lotus Notes e MS Exchange Server.
- **Applicazioni per teleconferenze.** Lo strato dell'applicazione consente agli utenti di utilizzare applicazioni per riunirsi, come per esempio video, dati vocali, etc. Un programma di questo tipo è MS Net Meeting.

- *World Wide Web*. Attraverso i browser gli utenti possono accedere ad informazioni provenienti da località remote in varietà di formati diversi (testo, immagini, video, suoni). Applicazioni di questo genere sono Apache Web Server, e tutti gli altri tradizionali Web server.

Modello a strati TCP/IP

Il modello a strati TCP/IP si basa su una rete a quattro strati:

- Applicazione
- Trasporto
- Internet
- Rete

Lo *strato di rete* immette sulla rete i frame in partenza e raccoglie quelli in arrivo. Prima di immettere sulla rete i frame, aggiunge ad essi una testata ed un controllo ciclico di ridondanza (CRC) per assicurare che i dati non siano corrotti durante il trasferimento.

Lo *strato Internet* svolge tre funzioni principali:

- l'indirizzamento
- la suddivisione in pacchetti
- l'instradamento

In questo strato risiede l'IP che offre la consegna di informazioni senza connessione e non garantita.

Lo *strato del trasporto* fornisce una comunicazione tra host utilizzando le cosiddette *porte*.

In questo strato si trovano i due protocolli

- TCP (Transmission Control Protocol) e
- UDP (User Datagram Protocol)

Il cui compito è proprio quello di trasportare dati.

Nello *strato delle applicazioni* si trovano infine le applicazioni che si basano sulla rete. Applicazioni di questo tipo sono le applicazioni Winsock dell'ambiente Windows quali *FTP* e *Telnet*.

Internet

Come detto precedentemente il protocollo IP non svolge alcun tipo di controllo per assicurarsi del buon esito del trasferimento dei dati. Di conseguenza i pacchetti possono andare perduti o non arrivare in sequenza.

Il ruolo svolto da IP è quello di aggiungere a ciascun pacchetto una intestazione contenente una serie di informazioni per poter effettuare il corretto instradamento dei dati.

L'intestazione contiene:

- l'indirizzo IP dell'origine: è l'indirizzo IP assegnato al mittente
- l'indirizzo IP del destinatario: è l'indirizzo IP assegnato al destinatario
- il protocollo di trasporto (TCP o UDP): serve ad indicare all'host destinatario il tipo di trasporto e di conseguenza il modo in cui manipolare i dati ricevuti
- checksum: è il CRC calcolato sui dati trasferiti/da trasferire che permette di verificare l'integrità dei dati
- TTL (Time-To-Live): il tempo di durata in vita di un datagram; alla partenza viene assegnato un valore predefinito che diminuisce ad ogni attraversamento di un router; quando il TTL raggiunge il valore zero viene il datagram viene tolto dalla rete;

Trasporto

Nello strato del trasporto si trovano i due protocolli TCP e UDP che hanno il compito di trasportare i dati da un host all'altro.

Il TCP si occupa delle comunicazioni orientate alla connessione. Questo significa che quando due host comunicano utilizzando come trasporto il TCP, è *necessario* che tra di essi si instauri una *sessione*.

Nel corso di trasmissioni di questo tipo ci si serve di numeri sequenziali e di conferme per assicurare il trasferimento con successo dei dati.

In pratica ciascun frammento di dati viene numerato e spedito al destinatario.

Il destinatario una volta ricevuti i frammenti provvede ad inviare al mittente un messaggio di avvenuta ricezione dei frammenti.

A questo punto il mittente provvede ad inviare i frammenti successivi.

In caso di mancata conferma dell'avvenuta ricezione da parte del destinatario il mittente provvede a spedire di nuovo i frammenti.

Questo meccanismo ovviamente appesantisce la trasmissione, ma assicura un elevato grado di affidabilità.

Utilizzando invece il protocollo UDP non vi è certezza dell'avvenuta ricezione da parte del destinatario dei dati spediti, ma in compenso la trasmissione risulta più semplice e veloce.

Quest'ultimo protocollo viene utilizzato nei casi in cui la velocità di trasmissione sia più importante della sicurezza della trasmissione (per es. in applicazioni *real-time*)

Confronto TCP/IP – modello OSI

Riportando in tabella i modelli a strati TCP/IP ed OSI si possono fare una serie di confronti.

Modello OSI	Modello TCP/IP
Applicazione	Applicazione
Presentazione	
Sessione	
Trasporto	Trasporto
Rete	Internet
Collegamento dati	Rete
Fisico	

1. nel modello TCP/IP gli strati *fisico* e *collegamento dati* sono fusi nello strato della *rete*. In pratica non viene fatta alcuna distinzione tra la scheda di rete ed i loro driver, e questo consente di implementare il TCP/IP in qualsiasi topologia di rete.
2. Lo strato *Internet* del TCP/IP, in cui viene implementato il protocollo IP, corrisponde allo strato di *rete* del modello OSI: entrambi si occupano dell'instradamento dei dati.
3. lo strato di *trasporto* nei due modelli è analogo, ed entrambi permettono che tra i due host si stabilisca una sessione.
4. lo strato dell'*applicazione* TCP/IP, infine è il merge degli strati di *applicazione*, *presentazione* e *sessione* del modello OSI

Protocolli nello strato Internet

Lo strato Internet comprende il protocollo IP, ma quest'ultimo non è l'unico protocollo implementato in questo strato.

Altri protocolli situati nello strato Internet sono:

- ARP (Address Resolution Protocol)
- ICMP (Internet Control Message Protocol)
- IGMP (Internet Group Management Protocol)

ARP

Il protocollo ARP si occupa di mettere in relazione gli indirizzi IP con gli indirizzi MAC. Alloquando viene richiesta la spedizione di dati ad una macchina con un dato indirizzo IP, ARP immette una richiesta sulla rete mediante la quale chiede all'host con quell'indirizzo di rispondergli notificandogli il proprio indirizzo MAC.

Ricevuta la risposta la corrispondenza IP -> MAC viene mantenuta in una cache al fine di evitare di ripetere l'operazione nel caso venisse richiesta una spedizione allo stesso indirizzo.

La cache ha ovviamente una dimensione predefinita. Nel caso in cui la cache fosse piena, l'aggiunta di una nuova voce porta all'eliminazione della voce più vecchia presente.

Con il comando *arp* è possibile visualizzare il contenuto della cache.

Esempi su ICMP e ping – TTL

ICMP fornisce un meccanismo di monitoraggio sugli errori e messaggi di controllo per l'insieme dei protocolli TCP/IP.

Il protocollo ICMP può svolgere le seguenti funzioni:

- fornire messaggi di echo e di risposta per verificare l'attendibilità di una connessione tra due host
- reindirizzare il traffico per fornire un instadamento più efficiente nel caso di intasamento di un router
- emettere un messaggio di time-out quando il TTL di un datagram viene superato
- fornire un messaggio di inibizione dell'origine per dire ad un host di rallentare le proprie comunicazioni per non intasare un router

Uno dei programmi più comunemente usati che fanno uso di questo protocollo è il ping (Packet INternet Groper).

L'indirizzo IP

L'indirizzo IP identifica univocamente un host su una Inter-Net TCP/IP.

Con il termine host si intende genericamente un computer, un terminale, un router, un hub.

Gli indirizzi IP su Internet sono assegnati da un comitato il cui compito è proprio quello di gestire tali indirizzi.

In genere però non ci si rivolge a tale ente (IANA – Internet Assigned Number Authority) ma al proprio provider (ISP – Internet Service Provider) che ha precedentemente provveduto a richiedere un insieme di indirizzi.

Di conseguenza l'assegnazione di un indirizzo IP non è un processo arbitrario, ma deve essere autorizzato dallo IANA (direttamente o indirettamente).

Questo non è strettamente valido nel caso delle cosiddette Intranet aziendali, che non hanno accesso diretto alla rete Internet mondiale.

In questo caso si può utilizzare teoricamente qualsiasi indirizzo, anche se nella pratica anche per le Intranet esistono delle raccomandazioni.

Oggi l'IP può essere dinamico o statico (fisso): dinamico vuol dire che al computer viene assegnato un indirizzo IP differente ogni volta che si collega al server ISP; mentre l'IP fisso o statico si ottiene tramite un accordo speciale con il proprio fornitore di servizi internet, il quale acconsente, in cambio di un compenso aggiuntivo, ad assicurare che il nome dell'utente che si collega venga associato a un indirizzo IP specifico.

Rappresentazione binaria e dot notation

L'indirizzo IP è notoriamente rappresentato nella forma:

x.y.z.w

dove x, y, z e w sono dei numeri compresi tra 0 e 255.

In pratica un indirizzo IP è composto dalla sequenza di 32 bit (quindi un long integer) anche se esso viene sempre rappresentato nella cosiddetta *dot notation*.

Per esempio il numero 2.130.706.433

in rappresentazione binaria è

01111111 00000000 00000000 00000001

ed in dot notation è

127.0.0.1

che è il cosiddetto indirizzo di loopback.

Come si vede dall'esempio precedente l'indirizzo in dot notation rappresenta il valore di ciascuno dei quattro byte componenti l'indirizzo IP.

In teoria, essendo l'indirizzo IP codificato su 32 bit, è possibile avere fino a $2^{32} - 1$ possibili indirizzi (4 miliardi e rotti).

In pratica non è così, in quanto esistono degli indirizzi riservati.

Classi di indirizzi IP

Gli indirizzi IP si suddividono in cinque classi di appartenenza, identificate con le lettere da A ad E.

Classe A	Un indirizzo di classe A riserva il primo byte all'indirizzamento della rete, e gli altri tre byte all'indirizzamento degli host. Il primo bit del byte che individua la rete è imposto essere 0, di conseguenza è possibile indirizzare teoricamente solo 127 reti, ciascuna con un massimo teorico di 16.777.216 host. In pratica le reti sono 126 in quanto la rete 127 è riservata per gli indirizzi di loopback, e gli host sono 16.777.214, in quanto gli indirizzi x.0.0.0 e x.255.255.255 non sono ammessi.
Classe B	Gli indirizzi di classe B riservano 16 bit alla rete e 16 agli host. Dei 16 bit riservati alla rete, i primi due sono obbligatoriamente 1 e 0, e questo di conseguenza implica che il primo byte possa avere valore tra 128 e 191. Il numero di reti ammesso in questa classe è quindi 16.384, e gli host 65.534 (non essendo ammessi anche in questo caso gli host x.y.0.0 e x.y.255.255).
Classe C	La classe C riserva i primi tre byte alla rete ed il quarto agli host. Poiché i primi tre bit del primo byte sono fissati a 110, il numero di reti univoche in questa classe è superiore a 2 milioni. Il primo byte può assumere valore tra 192 e 223. Il numero di host per ciascuna rete è di 254.
Classe D	Sono riservati ai gruppi multicast e non possono essere utilizzati per singoli host. I primi quattro bit sono obbligatoriamente 1110.
Classe E	Riservato per usi futuri. I primi cinque bit sono obbligatoriamente 11110.

Maschere di sottorete

Le maschere di sottorete servono ad individuare quali byte dell'indirizzo IP indirizzano la rete e quali gli host.

Le maschere di sottorete utilizzate dalle varie classi sono:

Classe A 255.0.0.0
 Classe B 255.255.0.0
 Classe C 255.255.255.0

Indirizzi di rete locale

La RFC 1918 ha riservato una serie di indirizzi IP come dedicati all'uso su reti locali, all'interno di firewall e server proxy.

Questi indirizzi sono

Da	A
10.0.0.0	10.255.255.255
172.16.0.0	172.31.255.255
192.168.0.0	192.168. 255.255

Questa raccomandazione intende fornire alle reti non collegate ad Internet un set di indirizzi che non siano in conflitto con quelli utilizzati su Internet.

Telnet

Il protocollo *telnet* permette di aprire una sessione di comunicazione bidirezionale tra due host.

Il sistema chiamato deve fornire un servizio di server telnet in esecuzione su TCP/IP.

Una volta connesso si può lavorare sulla macchina remota come se fosse direttamente collegata al proprio computer.

La sessione aperta dal telnet consente di utilizzare solo comandi richiamabili da linea di comando, o che comunque non necessitano di interfaccia grafica.

Risulta molto più arduo lavorare su macchine Windows che supportano sempre meno interazione da linee di comando, privilegiando interazione grafica.

Su macchine Windows, tra l'altro, non è previsto un servizio di server telnet, che deve quindi essere procurato ed installato a propria cura.

FTP

Il *File Transfer Protocol* è sicuramente uno dei protocolli più usati in assoluto. L'FTP consente di trasferire dati, tra host remoti in una rete TCP/IP.

Il protocollo FTP utilizza due processi distinti per il trasferimento dei dati:

- il Data Transfer Process (DTP) utilizzato per la vera e propria trasmissione dei dati ed
- il Protocol Interpreter (PI) che è usato per trasmettere i comandi tra client e server.
- Una sessione FTP è in effetti costituita da due sessioni separate:
- la prima si effettua tra i servizi PI di client e server, e serve a stabilire tutti i meccanismi della connessione (quali il nome dell'utente, la verifica della password). In questa fase viene anche concordata le modalità per effettuare la connessione dati tra il DTP del server e quello del client.
- nella seconda avviene l'effettivo scambio dei dati tra i due processi DTP.

FTP e FTP anonimo

Generalmente per poter collegarsi ad un client FTP bisogna essere autorizzati, cioè possedere una *username* ed una *password* per poter accedere al servizio FTP fornito dal server.

Questo tipo di connessione, che potrebbe a prima vista apparire sicuro, presenta l'inconveniente di far viaggiare sulla rete, in chiaro la password di accesso al server, con la conseguenza che uno *sniffer* presente sulla rete potrebbe catturarla ed usarla in maniera abusiva.

L'inconveniente maggiore deriva dal fatto che su macchine UNIX le username e le password per accedere al server FTP sono quelle effettivamente utilizzate per accedere al sistema in una sessione di lavoro normale. Questo significa che un eventuale sniffer può accedere al sistema con tutti i privilegi che quella username possiede.

Lo stesso inconveniente si verifica per server FTP poste su stazioni di lavoro Windows NT. Un modo per evitare questi inconvenienti, è quello di utilizzare, laddove il server FTP lo metta a disposizione, l'FTP anonimo.

Con l'FTP anonimo si utilizzano due username standard, *anonymous* o *ftp*, e la password da fornire è il proprio indirizzo di posta elettronica.

Questo metodo di accesso evita di far viaggiare in rete delle password in chiaro, ed ha l'ulteriore vantaggio che, nel caso di macchine UNIX, il server FTP prima di consentire l'accesso effettua una *chroot*, impedendo di fatto all'utente collegato di vedere la reale strutturazione del file system della macchina.

HTTP

L'HyperText Transfer Protocol (HTTP) è il protocollo utilizzato per la navigazione sul World Wide Web (WWW).

Il protocollo HTTP è il classico protocollo *connection oriented*, o a *richiesta e risposta*.

Il protocollo è stato specificato inizialmente dalla RFC 1945, resa obsoleta dalla 2068. Attualmente è alla sua terza revisione. La RFC di riferimento attuale per l'HTTP/1.1 è la 2616.

Posta elettronica (Electronic Mail – E Mail)

Un'altra applicazione molto usata sulle reti TCP/IP è la *posta elettronica*. Con l'esplosione di Internet l'uso della posta elettronica è diventato di uso comune per una serie di motivi:

- è il mezzo più economico e sicuro* per lo scambio di informazioni;
- è il mezzo più veloce, essendo in genere la consegna garantita in tempi dell'ordine dei minuti;
- il destinatario può ricevere la posta anche se non è al proprio posto abituale di lavoro;
- si possono utilizzare dei meccanismi per garantire la trasmissione in maniera protetta delle informazioni.

Per contro esistono anche dei dati negativi, quali:

- non è possibile leggere la posta elettronica senza avere accesso ad un computer;
- l'assenza del destinatario molto raramente può essere conosciuta;
- la facilità di accesso alla posta elettronica ha portato ad abusi del tipo catena di S. Antonio.

* per scambio sicuro si intende la sicurezza della consegna e dell'integrità dei dati, non la sicurezza intesa come riservatezza dei dati.

SMTP

L'SMTP (*Simple Mail Transfer Protocol* – RFC 821) è il protocollo utilizzato per trasmettere messaggi di posta elettronica, utilizzando il protocollo TCP per il trasporto.

Un server SMTP è un programma sempre attivo (su Unix non è sempre vero nel caso si utilizzi *inetd*) in ascolto sulla porta 25.

POP3

Il POP3 (Post Office Protocol version 3 – RFC 1939) è il protocollo più comunemente usato per prelevare i messaggi di posta elettronica.

In una sessione POP3 si seguono i seguenti passi:

- Il client si connette alla porta 110 del server.
- Il server invia un messaggio di saluto.
- Si inizia la sessione vera che consiste di una fase di AUTHORIZATION e di una successiva di TRANSACTION.
- Allo stato di TRANSACTION si passa solo dopo aver superato con successo lo stato di AUTHORIZATION, fornendo la propria identificazione.

Uuencode

Uuencode è uno dei metodi usati per convertire dati binari in testo per poter spedire documenti come allegati al messaggio di posta elettronica.

Come dice il nome stesso (Unix to Unix ENCODE) il metodo è stato sviluppato originariamente su macchine Unix per poter trasferire dati in maniera sicura da una macchina all'altra.

La macchina che riceve i dati deve ovviamente avere un programma analogo (*Uudecode*) per poter decodificare i dati.

Come è facilmente intuibile con questo metodo la mole di dati da trasferire aumenta, ma aumenta di pari passo la sicurezza che i dati vengano trasferiti senza perdita di informazione.

Altri metodi alternativi per il trasferimento di allegati sono il *BinHex* (sviluppato dalla Apple per i sistemi Macintosh) ed il metodo *MIME* (*Multipurpose Internet Mail Extensions* – RFC 2045) sviluppato appositamente per permettere la formattazione di messaggi non ASCII su Internet.

Quest'ultimo viene utilizzato anche per il trasferimento di file attraverso il protocollo HTTP.

World Wide Web

Il *World Wide Web* ha iniziato ad avere diffusione all'inizio degli anni 90 sulla spinta del protocollo HTTP.

Attualmente è noto come WWW, W3 o semplicemente Web: in ogni caso sono sinonimi del World Wide Web.

Infatti il WWW non è altro che una vasta rete di server HTTP in grado di comunicare tra di loro grazie ad Internet.

Il *Web non è Internet*: è solo uno dei servizi che è possibile trovare su Internet.

Attualmente il termine usato dagli utenti del Web per indicare che si consultano documenti sulla rete è *navigare* (in inglese *surfing*).

I Browser

Per poter accedere al Web bisogna utilizzare delle applicazioni sviluppate ad hoc, chiamati *browser*.

Il progenitore di tutti i browser esistenti è stato Mosaic sviluppato all'NCSA (National Center for Supercomputing Applications) presso l'università dell'Illinois.

I browser più diffusi sono

- Netscape Navigator (o Communicator)
- Microsoft Internet Explorer
- Opera

Per poter navigare sul Web non è ovviamente necessario avere un browser, potendo ovviamente farlo anche in modalità *linea di comando*.

I link

Come detto non è necessario avere un browser per poter navigare sul Web, ma senza un browser risulta oltremodo scomodo poter consultare i collegamenti ipertestuali (*link*) presenti nei documenti acceduti.

I link sono l'aspetto più importante legato al Web, in quanto permettono di collegare tra di loro file situati ovunque sul Web.

La possibilità di avere questi collegamenti permette di creare documenti come unione di più documenti, ciascuno curato da una persona (o gruppo di persone) diversa, e avere in linea sempre la versione aggiornata senza bisogno di alcun lavoro di sincronizzazione.

Formati dei file sul Web

I formati dei file accessibili sul Web possono essere i più svariati, i principali sono:

- pagine HTML, scritte utilizzando il linguaggio HTML (dette anche pagine Web)
- immagini (generalmente GIF o JPEG, ma anche icone, bitmap etc)
- file testo (i file .txt su Windows)

In aggiunta a questi esistono i cosiddetti file multimediali, ovvero file contenenti filmati e suoni.

Per poter accedere ad una pagina Web è necessario che essa sia *pubblicata sul Web* ovvero che essa sia data ad un server HTTP che ne gestisca l'accesso.

La pagina Web può essere resa accessibile anche mediante altro protocollo (per es. FTP), ma in questo caso si perde il significato intrinseco di pagina Web.

In genere i collegamenti di tipo FTP sono utilizzati per mettere a disposizione file che devono essere trasferiti sulla macchina dell'utente Web per poter poi essere utilizzati (per es. archivi compressi da applicazioni tipo Winzip o Gzip, documenti da leggere utilizzando appositi Word Processor quali MsWord o Frame Maker, etc).

Pagina Web

La pagina Web è generalmente una pagina HTML contenente immagini e link ad altre pagine Web.

Ogni pagina Web è identificata da un *URL (Uniform Resource Locator)* che non è altro che l'indirizzo univoco che identifica quella sull'intero Web.

Gli Indirizzi Internet (URL)

Il formato di una URL è:

protocollo:indirizzo

il protocollo può essere:

- http
- ftp
- file
- gopher
- telnet
- news
- mailto

il protocollo è separato dall'indirizzo dal carattere *due punti* (:); nel caso dei primi quattro il carattere due punti è seguito da due caratteri *slash* (//).

Alla fine dell'indirizzo, un ulteriore carattere *slash* indica l'inizio del file Web che si richiede al server HTTP.

Es.:

<http://www.altavista.com/index.html>

URL con numeri di porta

Il protocollo HTTP, cui la stragrande maggioranza delle pagine Web fa riferimento, utilizza come porta di default per le comunicazioni, la porta numero 80.

È possibile però impostare in server HTTP in modo da renderlo su una porta diversa da quella di default.

In tal caso il client (il browser utilizzato per la navigazione) deve essere informato del fatto, per evitare che spedisca le proprie richieste alla porta 80, porta su cui potrebbe non rispondere nessuno.

Per far ciò, l'URL deve essere scritta come segue:

http://www.nome.it:80000/

In pratica, alla fine dell'indirizzo e prima dello *slash* finale, va aggiunto il carattere due punti seguito dal numero della porta su cui il server è in attesa per poter soddisfare le richieste.

HTML

Abbiamo detto che le pagine Web sono scritte in HTML: ma cos'è l'HTML?

HTML è l'acronimo per HyperText Markup Language, quindi l'HTML è un linguaggio per il contrassegno di file ipertestuali.

HTML è quindi un linguaggio di *contrassegno* non di *programmazione*.

Con HTML è possibile presentare i propri documenti formattati in maniera personale, ma non è assolutamente possibile scrivere un programma in HTML.

HTML Un po' di storia

L'HTML, così come il WWW, sono nati alla fine degli anni 80 presso i laboratori di fisica nucleare del CERN di Ginevra.

Nel 1990 Tim Berners-Lee tenne ufficialmente a battesimo il World Wide Web che doveva servire a tutte le comunità di ricercatori e scienziati sparsi per il mondo (da cui Worldwide) per poter scambiare nel più breve tempo possibile e nella maniera più completa le proprie esperienze sulle ricerche in corso.

Tim Berners-Lee insieme a Robert Cailliau scrisse il primo client WWW (un browser-editor che girava sotto NeXTStep) ed il primo server insieme alla maggior parte del software di comunicazione che definiva URL, HTTP ed HTML.

Nel 1994 Tim Berners-Lee e gli altri partecipanti al progetto fondarono il W3C (World Wide Web Consortium), che è attualmente il responsabile degli standard HTTP, HTML e delle tecnologie per il Web.

SGML – Il padre dell'HTML

L'SGML (Standard Generalized Markup Language) è una famiglia di linguaggi di Markup di cui HTML è uno dei membri.

In pratica l'SGML è una specifica di linguaggi di contrassegno che permette ad un utente di definire il proprio linguaggio di contrassegno.

L'HTML è quindi una istanza dell'SGML.

Evoluzione dell'HTML

La prima versione dell'HTML, nota semplicemente come HTML senza specifica della versione, non si differenzia in maniera sostanziale dalla versione corrente che è la 4.0.

I tag definiti nella versione iniziale del linguaggio sono rimasti anche nella versione attuale, e questo comporta che una vecchia pagina scritta la versione 1.0 dell'HTML è ancora correttamente visualizzabile da un browser basato sull'ultima versione.

Non è ovviamente vero il viceversa, per cui una pagina scritta secondo gli standard della 4.0 difficilmente porterà alla visualizzazione di qualcosa di gradevole in vecchio browser (come per es. Netscape 2.0 o IE 3.0).

Dalla versione 1.0 si è passati alla versione HTML+ e successivamente alla HTML 2.0.

È a questo punto che i produttori dei browser maggiormente diffusi incominciano a far sentire la propria influenza sullo sviluppo del linguaggio. Infatti la release HTML 3.0 non sarà mai supportata da nessun browser commerciale (solo Arena, sviluppato dagli stessi membri del consorzio W3C, supporterà questa versione) e di conseguenza la successiva

versione (HTML 3.2) è di fatto l'acquisizione da parte del linguaggio di una serie di estensioni apportate dai produttori.

L'ultima versione del linguaggio è la 4.0 ed aggiunge una serie di estensioni al linguaggio. In tale versione vengono accolte alcune estensioni ad HTML 3.2 quali i frame. Viene inoltre raccomandato l'utilizzo dei fogli di stile (CSS – Cascading Style Sheet) per la formattazione del documento.

HTML – funzioni principali

La funzione principale è quella di essere un linguaggio universale per la classificazione delle varie parti di un documento.

Essendo un linguaggio di Markup, non è legato a nessuna piattaforma in particolare, in quanto una pagina HTML viene sempre distribuita in formato sorgente ed è compito del browser visualizzare correttamente tutti i *tag* contenuti nella pagina.

Teoricamente la visualizzazione di una pagina HTML dovrebbe essere indipendente dal browser, ma in effetti non lo è per una serie di motivi, primo fra tutti è quello che non esiste alcuno standard che imponga al browser di visualizzare un *tag* con un font anziché con un altro, oppure utilizzando una dimensione anziché un'altra.

I tag

I tag sono l'insieme dei simboli definiti in HTML e che hanno un significato speciale.

Un tag inizia sempre con il simbolo minore (<), è seguito da una parola riservata, ed è chiuso dal carattere maggiore (>).

Le parole riservate che compongono i tag possono essere sia parole vere e proprie (per es. BODY), o abbreviazioni in genere significative. Per esempio HR è una parola chiave per un tag che disegna una linea orizzontale nella pagina e sta per Horizontal Rule (letteralmente linea orizzontale), B sta per Bold e serve per evidenziare in grassetto una parte di testo.

Un tag *dovrebbe* essere chiuso da un tag analogo a quello di apertura, con la sola differenza del carattere slash che precede la parola chiave: es. <BODY> ... </BODY>.

Il condizionale *dovrebbe* è stato usato in quanto non esiste alcun parser che possa validare la correttezza di una pagina HTML, ed i browser nel caso incontrino un tag aperto ma non chiuso possono cercare di dedurre quale sia la soluzione migliore per poter comunque visualizzare la pagina.

Le differenti interpretazioni da parte dei browser portano quindi a risultati diversi nella visualizzazione.