Andrea Piccione

Applied Information Technology

I. La Rete

Marzo 2002

Premessa:

Una rete di elaboratori non è necessariamente Internet. Internet non è solo WWW ed e-mail. Prima di iniziare ad essere autori nella Rete è bene conoscere, anche solo a grandi linee, il contesto in cui si opera. In questo modo si possono sfuttare al meglio le potenzilità del mezzo a propria disposizione.

Obiettivi:

- conoscere le origini della Rete;
- conoscere aspetti della complessità hardware e software della Rete;
- conoscere i diversi livelli di comunicazione.

Contenuti:

1. Introduzione
   - 1.1 Usi delle rete di elaboratori
   - 1.2 Aspetti hardware delle rete
   - 1.3 Aspetti software delle rete
2. Il livello uno (Fisico)
3. Il livello due (DataLink)
4. Il livello tre (Network)
   - 4.1 Internet Protocol
   - 4.2 Indirizzi IP
5. Il livello quattro (Transport)
6. Il livello cinque (Application)
   - 6.1 Il DNS
   - 6.2 La posta elettronica

Riferimenti:

- A. Tanenbaum, *Reti di calcolatori*, UTET Libreria
- C. Hunt, *TCP/IP Network Administration*, O'Reilly
- G. Govanus, *TCP/IP*, McGraw-Hill
- G. Bongiovanni, *Appunti del corso di Sistemi di Elaborazione: Reti 1*, Università La Sapienza - Roma

1 Introduzione

Definizione:

Una rete di elaboratori (computer network) è un grande numero di elaboratori autonomi, interconnessi fra loro:

- autonomi: significa che non deve esserci fra loro una relazione tipo master/slave (ad es., l'uno non può forzare lo spegnimento dell'altro);
- interconnessi: significa che devono essere capaci di scambiare informazioni (sfruttando un opportuno mezzo fisico).

Origini:

1876 Alexander Bell a Thomas Watson iniziano i loro esperimenti per la trasmissione del suono all'interno di cavi elettrici.

1964 L'istituto RAND effettuta ricerche top secret per mettere gli Stati Uniti in condizione di affrontare una guerra termonucleare. Un problema cruciale è quello di disporre di una rete di comunicazione in grado di sopravvivere ad un eventuale bombardamento. La proposta presentata dalla RAND è quella di utilizzare una rete decentralizzata, fata di nodi tutti uguali tra loro e in cui i messaggi viaggiano spezzettati in tanti "pacchetti", che possono percorrere strade diverse ed essere ricomposti una volta giunti a destinazione. Questa tecnica innovativa di inoltro dei messaggi prende il nome di packet switching.

1969 In USA il pentagono affida all'ARPA (Advanced Research Project Agency) il compito di installare il primo nodo di una rete militare basata sul packet switching. Questa rete prenderà il nome di Arpanet e sarà l'embrione della futura Internet.

1973 Nasce in Francia il primo personal computer, chiamato Micral, ad opera di Truong Tromng Hi, un francese di origini vietnamite che lavora per la Realisations Etudes Electroniques S. A.

1977 Su Arpanet viene adottato il protocollo TCP/IP, un sistema di collegamento che dà la possibilità ad Arpanet di scambiare informazioni con altre reti locali diverse tra loro.

1981 Compare sul mercato il primo personal computer IBM.

1983 In USA da un insieme variegato di reti locali nasce Internet.

1984 Il personal computer Apple Macintosh è il primo a fare uso di "finestre","cartelle" e "icone" per rappresentare graficamente in maniera intuitiva le informazioni contenute all'interno del calcolatore. Nasce a Baltimora FidoNet la prima rete di BBS (Bulletin Board System).

1990 Nasce al CERN di Ginevra il World Wide Web per consentire una agevole cooperazione fra i gruppi di ricerca di fisica sparsi nel mondo. E' un'architettura software volta a fornire l'accesso e la navigazione a un enorme insieme di documenti collegati fra loro e sparsi su milioni di elaboratori.

1996 Negli USA Clinton approva il Telecomunications Reform Bill, una legge che deregolamenta il settore delle comunicazioni per permetere lo sviluppo delle "autostrade informatiche". Cadono le barriere imposte dall'antitrust e nascono nuove alleanze tra giganti dell'informatica, della tv via cavo e dell'intrattenimento televisivo e cinematografico.

2000 Dimensioni (stimate) della rete all'inizio del 2000: 800 milioni di documenti disponibili (raddoppiano ogni 18 mesi) 5 milioni di siti web, 200 milioni di utenti.

1.1 Usi delle reti di elaboratori

Moltissimi sono gli usi delle reti di elaboratori, sia per le organizzazioni che per i singoli individui.

- Per le organizzazioni:
  - condivisione risorse: si possono rendere disponibili a chiunque programmi e informazioni anche distanti migliaia di km;
  - affidabilità: si ottiene mettendo in rete sorgenti alternative delle risorse (ad es. duplicando le applicazioni e i dati su più computer). E' importante in sistemi che devono funzionare a tutti i costi (traffico aereo, centrali nucleari, sistemi militari, ecc.);
  - comunicazione fra persone: è possibile inviarsi messaggi, scambiarsi file, ecc.
- Per i singoli individui: (di solito da casa propria tramite "fornitori di accesso" o provider):
  - accesso ad informazioni remote, ad es.:
    - accesso a servizi bancari;
    - acquisti da casa;
    - navigazione sul World Wide Web;
  - comunicazioni fra persone:
    - posta elettronica;
    - videoconferenza;
    - gruppi di discussione;
  - divertimento:
    - video on demand (selezione e ricezione via rete di un qualunque spettacolo tratto da un catalogo);
    - giochi interattivi (contro macchine o avversari umani).

1.2 Aspetti hardware delle reti

Due parametri sono utili per definire le caratteristiche di una rete, anche se non esiste una tassonomia universalmente accettata:

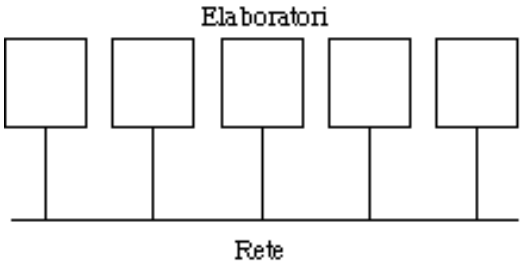

Figura 1. *Una rete broadcast*

- tecnologia trasmissiva;
- scala dimensionale.

Tecnologia trasmissiva

Ci sono due tipologie per quanto riguarda la tecnologia trasmissiva:

- reti broadcast;
- reti punto a punto.

Le reti broadcast sono dotate di un unico "canale" di comunicazione che è condiviso da tutti gli elaboratori. Brevi messaggi (spesso chiamati pacchetti) inviati da un elaboratore sono ricevuti da tutti gli altri elaboratori. Un indirizzo all'interno del pacchetto specifica il destinatario. Quando un elaboratore riceve un pacchetto, esamina l'indirizzo di destinazione; se questo coincide col proprio indirizzo il pacchetto viene elaborato, altrimenti viene ignorato. Le reti broadcast, in genere, consentono anche di inviare un pacchetto a tutti gli altri elaboratori, usando un opportuno indirizzo (broadcasting). In tal caso tutti prendono in considerazione il pacchetto.

Le reti punto a punto consistono invece di un insieme di connessioni fra coppie di elaboratori. Per arrivare dalla sorgente alla destinazione, un pacchetto può dover attraversare uno o più elaboratori intermedi. Spesso esistono più cammini alternativi, per cui gli algoritmi di instradamento (routing) hanno un ruolo molto importante.

In generale (ma con molte eccezioni):

- le reti geograficamente localizzate tendono ad essere broadcast;

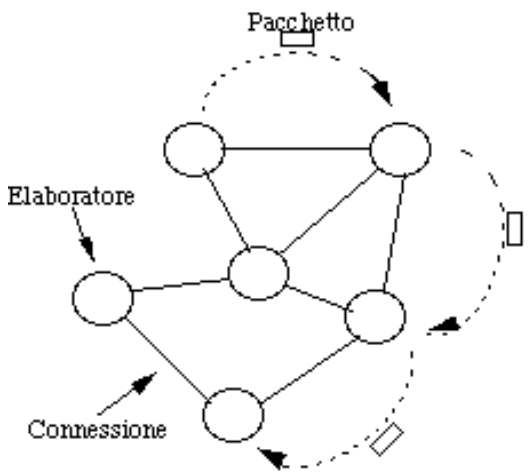

Figura 2. *Una rete punto a punto*

- le reti geograficamente molto estese tendono ad essere punto a punto.

Alcune eccezioni:

- rete geografica realizzata via satellite (e quindi broadcast);
- rete locale basata su ATM: Asynchronous Transfer Mode (Modalità di Trasmissione Asincrona) è un metodo per trasmettere dati, video e suono su reti locali (LAN) ad alta velocità. I dati vengono trasmessi utilizzando raffiche continue di celle (pacchetti di lunghezza fissa). (Questa modalità è punto a punto).

Scala dimensionale

Un criterio alternativo di classificazione è la scala dimensionale delle reti. In questo contesto si distingue fra reti locali, reti metropolitane e reti geografiche. La distanza è un fattore molto importante, poiché a differenti scale dimensionali si usano differenti tecniche.

Distanza fra processori	Ambito	Tipo di rete
10 m.	Stanza	Rete locale
100 m.	Edificio	Rete locale
1 km.	Campus	Rete locale
10 km.	Città	Rete metropolitana
100 km.	Nazione	Rete geografica
1000 km.	Continente	Rete geografica
10.000 km.	Pianeta	Internet (Rete geografica)

Reti locali

Le reti locali (Local Area Network, LAN), in genere:

- sono possedute da una organizzazione (reti private);
- hanno un'estensione che arriva fino a qualche km;
- si distendono nell'ambito di un singolo edificio o campus;
- sono usatissime per connettere PC o workstation.

Esse si distinguono dagli altri tipi di rete per tre caratteristiche:

- dimensione: la dimensione non può andare oltre un certo limite, per cui è noto a priori il tempo di trasmissione nel caso peggiore. Questa conoscenza permette di utilizzare delle tecniche particolari per la gestione del canale di comunicazione;
- tecnologia trasmissiva: come già accennato, le LAN sono in generale reti broadcast. Velocità di trasmissione tipiche sono da 10 a 100 Mbps (megabit al secondo, cioé milioni di bit al secondo), con basso ritardo di propagazione del segnale da un capo all'altro del canale (qualche decina di microsecondi) e basso tasso di errore.

Reti metropolitane

Le reti metropolitane (Metropolitan Area Network, MAN)hanno un'estensione tipicamente urbana (quindi anche molto superiore a quella di una LAN) e sono generalmente pubbliche (cioé un'azienda, ad es. Telecom Italia, mette la rete a disposizione di chiunque desideri, previo pagamento di una opportuna tariffa).

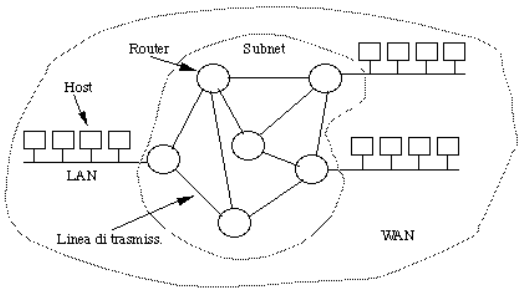

Figura 3. *Struttura tipica di una WAN*

Reti geografiche

Le reti geografiche (Wide Area Network, WAN) si estendono a livello di una nazione, di un continente o dell'intero pianeta. Una WAN è tipicamente costituita da due componenti distinte:

- un insieme di elaboratori (host oppure end system) sui quali girano i programmi usati dagli utenti;
- una comunication subnet (o subnet), che connette gli end system fra loro. Il suo compito è trasportare messaggi da un end system all'altro, cos come il sistema telefonico trasporta parole da chi parla a chi ascolta.

Di norma la subnet consiste, a sua volta, di due componenti:

- linee di trasmissione (dette anche circuiti, canali, trunk):
- elementi di commutazione (switching element): gli elementi di commutazione sono elaboratori specializzati utilizzati per connettere fra loro due o più linee di trasmissione. Quando arrivano dati su una linea, l'elemento di commutazione deve scegliere una linea in uscita sul quale instradarli. Non esiste una terminologia standard per identificare gli elementi di commutazione. Termini usati sono: o sistemi intermedi; o nodi di commutazione pacchetti; o router (quello che utilizzeremo noi).

Una tipica WAN è utilizzata per connettere più LAN fra loro come illustrato in Fig. 3. In generale una WAN contiene numerose linee (spesso telefoniche) che congiungono coppie di router. Ogni router, in generale, deve:

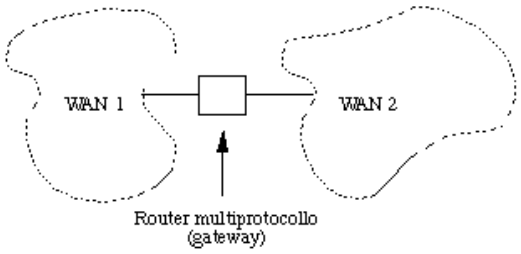

Figura 4. *Interconnessione di reti*

1. ricevere un pacchetto da una linea in ingresso;
2. memorizzarlo per intero in un buffer interno;
3. appena la necessaria linea in uscita è libera, instradare il pacchetto su essa.

Una subnet basata su questo principio si chiama:

- punto a punto;
- store and forward;
- a commutazione di pacchetto (packet switched).

Una WAN può essere anche realizzata in maniera mista: in parte cablata, in parte basata su radio o satellite.

Interconnessione di reti (Internetwork)

Una internetwork è formata quando reti diverse (sia LAN che MAN o WAN) sono collegate fra loro. A prima vista, almeno in alcuni casi, la cosa è apparentemente uguale alla definizione di WAN vista precedentemente (se al posto di subnet si scrive WAN, abbiamo una internetwork costituita da una WAN e quattro LAN). Alcuni problemi però sorgono quando si vogliono connettere fra di loro reti progettualmente diverse (spesso incompatibili fra loro). In questo caso si deve ricorrere a speciali attrezzature, dette gateway (o router multiprotocollo), che oltre ad instradare i pacchetti da una rete all'altra, effettuano le operazioni necessarie per rendere possibili tali trasferimenti. Solitamente si usano i termini:

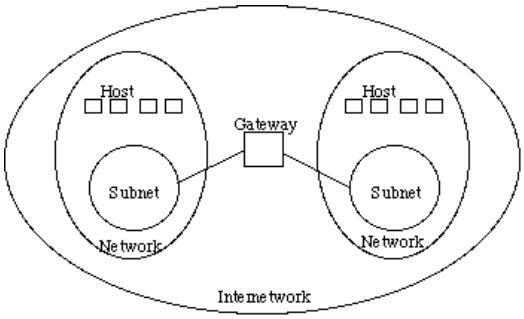

Figura 5. *Relazioni fra subnet, network e internetwork*

- internet come sinonimo di internetwork, cioé la interconnessione di più reti generiche;
- Internet (con la I maiuscola) per riferirci alla specifica internetwork, basata su TCP/IP, che ormai tutti conoscono.

1.3 Aspetti software delle reti

Le prime reti furono progettate cominciando dallo hardware (HW) e sviluppando il software (SW) solo successivamente, quasi come se esso fosse un'accessoria appendice dello hardware. Questo approccio non funziona più. Il SW di rete è oggi altamente strutturato. Esaminiamo ora, a grandi linee, tale strutturazione.

Gerarchie di protocollo

Per ridurre la complessità di progetto, le reti sono in generale organizzate a livelli, ciascuno costruito sopra il precedente. Fra un tipo di rete ed un altra, possono essere diversi:

- il numero di livelli;
- i nomi dei livelli;
- il contenuto dei livelli;
- le funzioni dei livelli.

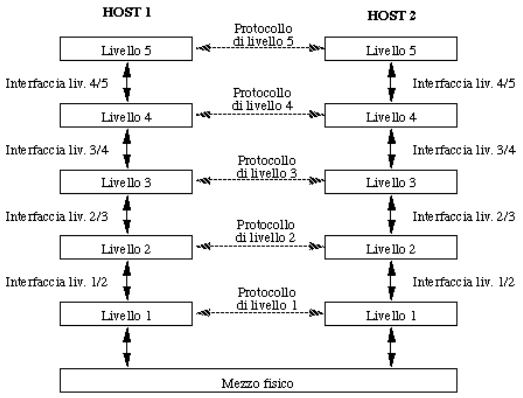

Figura 6. *Dialogo fra peer entity*

Un principio generale è sempre rispettato: lo scopo di un livello è offrire certi servizi ai livelli più alti, nascondendo i dettagli sul come tali servizi siano implementati. Ogni livello deve solo sapere come passare i dati ai livelli adiacenti.

Isolare le funzioni di comunicazione in livelli differenti minimizza l'impatto dei cambiamenti tecnologici o di parti dell'insieme dei protocolli. Il livello n su un host porta avanti una conversazione col livello n su di un'altro host. Le regole e le convenzioni che governano la conversazione sono collettivamente indicate col termine di protocollo di livello n.

I protocolli sono regole formali di comportamento. Nelle relazioni internazionali i protocolli minimizzano i problemi causati dalle differenti tradizioni culturali quando diverse nazioni lavorano insieme e sono costituiti da regole comuni stabilite indipendentemente dalla cultura di origine. Quando consideriamo diversi computer collegati in rete tra loro o l'interconnessione di reti si pone una analoga esigenza di stabilire regole di comunicazione indipendenti dai dettagli strutturali di calcolatori e reti.

Le entità (processi) che effettuano tale conversazione si chiamano peer entitiy (entità di pari livello). Il dialogo fra due peer entity di livello n viene materialmente realizzato tramite i servizi offerti dal livello (n-1).

- In realtà non c'è un trasferimento diretto dal livello n di host 1 al livello n di host 2. Ogni livello di host 1 passa i dati, assieme a delle informazioni di controllo, al livello sottostante.

- Al di sotto del livello 1 c'è il mezzo fisico, attraverso il quale i dati vengono trasferiti da host 1 ad host 2.
- Quando arrivano a host 2, i dati vengono passati da ogni livello (a partire dal livello 1) a quello superiore, fino a raggiungere il livello n.

Fra ogni coppia di livelli adiacenti è definita una interfaccia, che caratterizza:

- le operazioni primitive che possono essere richieste al livello sottostante;
- i servizi che possono essere offerti al livello sovrastante.

Architettura di rete

L'insieme dei livelli e dei relativi protocolli è detto architettura di rete. La specifica dell' architettura deve essere abbastanza dettagliata da consentire la realizzazione di SW e/o HW che, per ogni livello, rispetti il relativo protocollo. Viceversa, i dettagli implementativi di ogni livello e le interfacce fra livelli non sono parte dell'architettura, in quanto sono nascosti all'interno di un singolo host. É quindi possibile che sui vari host della rete ci siano implementazioni che differiscono fra di loro anche in termini di interfacce fra livelli, purché ogni host implementi correttamente i protocolli previsti dall'architettura. In questo caso possono dialogare fra loro anche host aventi caratteristiche (processore, sistema operativo, costruttore) diverse.

Un modello di architettura sviluppato dalla International Standard Organization(ISO) solitamente usato per descrivere la struttura e le funzioni dei protocolli di comunicazione di dati prevede sette diversi livelli.

L'architettura TCP/IP è la più diffusa ed è lo standard utilizzato in Internet. Notate ancora che è essenziale definire standard comuni di comunicazione. Pensate ad esempio al trasporto dei container. Se ci fossero standard diversi tra nazioni e tra il trasporto su strada, su rotaie, navi e aerei, il trasporto di merci sarebbe estremamente complesso.

Non esiste un accordo comune di come descrivere l'architettura TCP/IP con un modello a livelli. Nella nostra breve descrizione noi useremo 5 livelli.

Funzionamento del software di rete

Per comprendere i meccanismi basilari di funzionamento del software di rete si può pensare alla seguente analogia umana, nella quale un filosofo indiano vuole conversare con uno stregone africano (vedi Fig. 7).

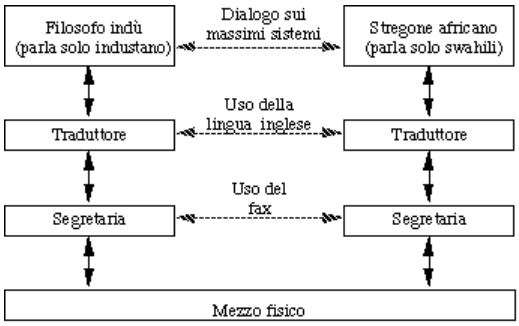

Figura 7. *Dialogo tra sapienti*

Nel caso delle reti, la comunicazione fra le due entità di livello superiore avviene con una modalità che, almeno in linea di principio, è uguale in tutte le architetture di rete (vedi Fig. 8). Vediamo cosa accade:

1. il programma applicativo (livello 5) deve mandare un messaggio M alla sua peer entity;
2. il livello 5 consegna M al livello 4 per la trasmissione;
3. il livello 4 aggiunge un suo header in testa al messaggio (talvolta si dice che il messaggio è inserito nella busta di livello 4); questo header contiene informazioni di controllo.
4. il livello 4 consegna il risultato al livello 3;
5. il livello 3 può trovarsi nella necessità di frammentare i dati da trasmettere in unità più piccole, (pacchetti) a ciascuna delle quali aggiunge il suo header;
6. il livello 3 passa i pacchetti al livello 2;
7. il livello 2 aggiunge ad ogni pacchetto il proprio header (e magari un trailer) e lo spedisce sul canale fisico;
8. nella macchina di destinazione i pacchetti fanno il percorso inverso, con ogni livello che elimina (elaborandoli) l'header ed il trailer di propria competenza, e passa il resto al livello superiore.

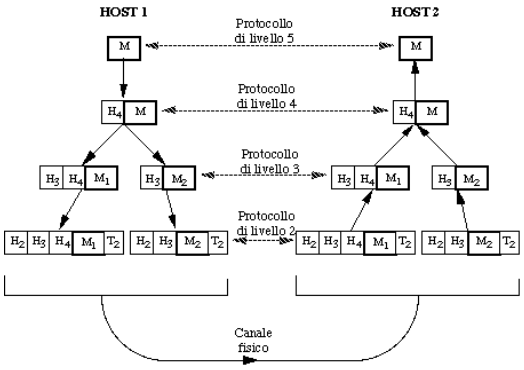

Figura 8. *Flusso dell'informazione fra peer entity*

Aspetti importanti sono i seguenti:

- le peer entity pensano concettualmente ad una comunicazione orizzontale fra loro, basata sul protocollo del proprio livello, mentre in realtà comunicano ciascuna col livello sottostante attraverso l'interfaccia fra i due livelli;
- spesso i livelli bassi sono implementati in hardware per ragioni di efficienza.

2 Il livello uno (Fisico)

L'informazione può essere trasmessa a distanza variando opportunamente una qualche caratteristica fisica del mezzo scelto per la trasmissione. Tale variazione si propaga, con una certa velocità, lungo il mezzo di trasmissione e dopo un certo tempo arriva all'altra estremità del mezzo, dove può venir rilevata. Ad esempio, se il mezzo è un cavo metallico, si può variare la tensione applicata ad un'estremità. Tale variazione di tensione verrà successivamente rilevata all'altra estremità. I mezzi trasmissivi sono sostanzialmente di tre tipi:

- mezzi elettrici (cavi): in essi il fenomeno fisico utilizzato è l'energia elettrica;

- mezzi wireless (onde radio): il fenomeno fisico è l'onda elettromagnetica, una combinazione di campo elettrico e campo magnetico variabili, che si propaga nello spazio e che induce a distanza una corrente elettrica in un dispositivo ricevente (antenna);
- mezzi ottici (LED, laser e fibre ottiche): in essi il fenomeno utilizzato è la luce. Si tratta dei mezzi più recenti, che hanno rivoluzionato il settore.

Rappresentando il valore nel tempo del fenomeno fisico utilizzato come una funzione f(t), si può studiare matematicamente il segnale risultante.

In linea di principio, la trasmissione può avvenire con due modalità differenti: trasmissione di segnale analogico e trasmissione di segnale digitale. La differenza fondamentale fra un segnale analogico e uno digitale è che:

- il primo può variare gradualmente in un intervallo costituito da un numero infinito di possibili valori;
- il secondo può variare solamente passando bruscamente da uno all'altro di un insieme molto piccolo di valori (da due a qualche decina).

Si tenga presente però che il fenomeno fisico utilizzato non è digitale ma analogico. Un segnale quindi non può passare istantaneamente da un valore ad un altro, ma impiegherà un certo tempo per effettuare la transizione. La conseguenza è che un mezzo fisico farà del suo meglio per trasportare un segnale digitale, ma non riuscirà a farlo arrivare esattamente com'è partito.

3 Il livello due (DataLink)

Competenze del livello data link Questo livello ha le seguenti incombenze principali:

- offrire servizi al livello network con un'interfaccia ben definita;
- determinare come i bit del livello fisico sono raggruppati in frame (framing);
- gestire gli errori di trasmissione;
- regolare il flusso della trasmissione fra sorgente e destinatario.

Servizi offerti al livello Network

Il servizio principale è trasferire dati dal livello network della macchina di origine al livello network della macchina di destinazione. Come sappiamo, la trasmissione reale passa attraverso il livello fisico, ma noi ragioniamo in termini di dialogo fra due processi a livello data link che usano un protocollo di livello data link. Vediamo ora un tipico esempio di funzionamento. Consideriamo un router con alcune linee in ingresso ed alcune in uscita. Ricordiamo che il routing avviene a livello tre (network), quindi il router gestisce i livelli uno, due e tre.

1. Quando al router arrivano dei bit da una linea fisica, lo hardware apposito se ne accorge (siamo a livello 1) e li passa al corrispondente SW/HW di livello due.

2. Il SW/HW di livello due (data link), che in genere è contenuto in un chip sulla scheda (o adattatore) di rete (tipico esempio è una scheda Ethernet o un Modem+Scheda di Accesso Remoto) fa i controlli opportuni:

   - framing;
   - controllo errori di trasmissione;
   - controllo numero di frame (se necessario).

3. Se tutto è OK, il SW/HW di livello due genera un interrupt alla CPU, che chiama in causa il SW di livello tre (network):

   - questo è tipicamente un processo di sistema, al quale viene passato il pacchetto contenuto nel frame di livello due per l'ulteriore elaborazione;
   - l'elaborazione consiste nel determinare, sulla base delle caratteristiche del pacchetto (in particolare dell'indirizzo di destinazione), su quale linea in uscita instradarlo.

4. Presa questa decisione, il SW di livello tre consegna il pacchetto al corrispondente SW/HW di livello due, che lo imbusta in un nuovo frame e lo consegna al sottostante livello fisico (ossia quello relativo alla linea in uscita prescelta).

Il livello uno accetta un flusso di bit grezzi e cerca di farli arrivare a destinazione. Però:

- il flusso non è esente da errori;
- possono arrivare più o meno bit di quanti sono stati inviati.

E' compito del livello due rilevare, e se possibile correggere, tali errori. L'approccio usuale del livello due è il seguente.

- In trasmissione:
  - spezza il flusso di bit che arriva dal livello tre in una serie di frame;
  - calcola un'apposita funzione (checksum) per ciascun frame;
  - inserisce il checksum nel frame;
  - consegna il frame al livello uno, il quale lo spedisce come sequenza di bit.
- In ricezione:
  - riceve una sequenza di bit dal livello uno;
  - ricostruisce da essa un frame dopo l'altro;
  - per ciascun frame ricalcola il checksum;
  - se esso è uguale a quello contenuto nel frame questo viene accettato, altrimenti viene considerato errato e quindi scartato.

Rilevamento e correzione errori

Ci sono molti fattori che possono provocare errori, sopratutto sul local loop e nelle trasmissioni wireless. Viceversa, essi sono piuttosto rari nei mezzi più moderni quali le fibre ottiche. Gli errori sono dovuti in generale a:

- rumore di fondo;
- disturbi (ad es. fulmini) improvvisi;
- interferenze (ad es. motori elettrici).

Ci sono due approcci al trattamento degli errori:

- includere abbastanza informazione aggiuntiva in modo da poter ricostruire il messaggio originario (correzione dell'errore);
- includere meno informazione aggiuntiva, in modo da accorgersi che c'è stato un errore, senza per questo essere in grado di correggerlo (rilevazione dell'errore).

4 Il livello tre (Network)

Il livello network è incaricato di muovere i pacchetti dalla sorgente fino alla destinazione finale, attraversando tanti sistemi intermedi (router o gateway) della subnet di comunicazione quanti è necessario. Ciò è molto diverso dal compito del livello data link, che è di muovere informazioni solo da un capo all'altro di un singolo canale di comunicazione wire-like. Le incombenze principali di questo livello sono:

- conoscere la topologia della rete;
- scegliere di volta in volta il cammino migliore (routing);
- gestire il flusso dei dati e le congestioni (flow control e congestion control);
- gestire le problematiche derivanti dalla presenza di più reti diverse (internetworking).

Nel progetto e nella realizzazione del livello network di una architettura di rete si devono prendere decisioni importanti in merito a:

- servizi offerti al livello transport;
- organizzazione interna della subnet di comunicazione.

4.1 Internet Protocol

Internet Protocol (IP) è il building block della Rete. Le sue funzioni includono:

- definire il datagram (vedi sotto), che è l'unità base della trasmissione in Internet;

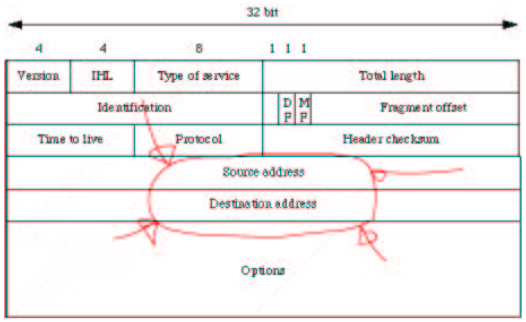

Figura 9. *Header IP*

- definire lo schema di indirizzamento;
- mouvere i dati dal livello DataLink al livello Transport;
- guidare i dati verso host remoti;
- frammentare e riassemblare i datagrams.

IP è un protocollo senza connessione, nel senso che non si occupa di controllare che la connessione sia stata attivata (questo compito spetta al livello quattro). Un datagram è il pacchetto di informazioni definito dal protocollo IP.

4.2 Indirizzi IP

Un indirizzo IP è formato da 32 bit e codifica due cose:

- network number, cioé il numero assegnato alla rete IP (detta network) su cui si trova l'elaboratore; in questo contesto una network è caratterizzata dal fatto di essere costituita da un unico canale di comunicazione cui sono connessi tutti gli host della network stessa (e quindi, ad esempio, una LAN oppure una linea punto punto fra due router);
- host number, cioé il numero assegnato all'elaboratore.

La combinazione è unica: non possono esistere nell'intera rete Internet due indirizzi IP uguali. Si noti che solitamente si ritiene che ogni host sulla rete abbia un singolo indirizzo IP.

5 Il livello quattro (Transport)

Il livello transport è il cuore di tutta la gerarchia di protocolli. Il suo compito è di fornire un trasporto affidabile ed efficace dall'host di origine a quello di destinazione, indipendentemente dalla rete utilizzata. Il software di livello transport è presente solo sugli host, e non nei router della subnet di comunicazione. Questi alcuni compiti che spettano al livello Transport:

- aprire e chiudere una connessione (messaggi di conferma);
- gestire il flusso di dati (diverse applicazioni possono avere bisogno della connessione contemporaneamente).

5.1 Transmission Control Protocol

Il livello transport di Internet è basato su due protocolli:

- TCP (Transmission Control Protocol)
- UDP (User Data Protocol)

Il secondo è di fatto IP con l'aggiunta di un breve header, e fornisce un servizio di trasporto datagram (quindi non affidabile).

Il protocollo TCP è stato progettato per fornire un flusso di byte affidabile, da sorgente a destinazione, su una rete non affidabile. TCP è sviluppato indipendentemente da ogni harware specifico, dal sistema operativo e dalla struttura fisica della rete. TCP offre un servizio reliable e connection oriented, e si occupa di:

- accettare dati dal livello application;
- spezzarli in segment;
- consegnarli al livello network, eventualmente ritrasmettendoli;
- ricevere segmenti dal livello network;
- rimetterli in ordine, eliminando buchi e doppioni;
- consegnare i dati, in ordine, al livello application.

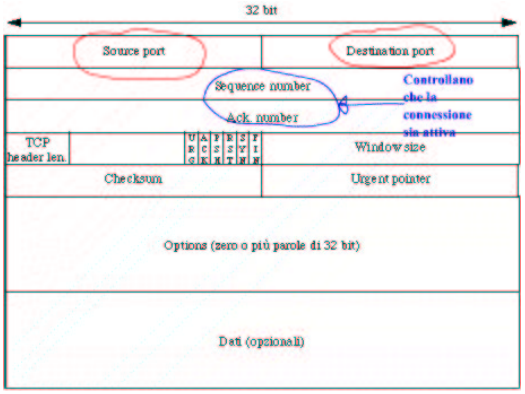

Figura 10. *Formato di un segmento TCP*

I servizi di TCP si ottengono creando connessione di livello transport identificata da una coppia di punti d'accesso detti *socket*. Ogni socket ha un socket number che consiste della coppia:

`IP address: Port number`

I port number hanno 16 bit. Quelli minori di 256 sono i cosidetti well-known port, riservati per i servizi standard. Ad esempio:

Port number	Servizio
7	Echo
20	Ftp (control)
21	Ftp (data)
23	Telnet
25	Smtp
80	Http
110	Pop versione 3

6 Il livello cinque (Application)

Nel nostro modello di architettura (e anche nell'architettura TCP/IP) sopra il livello transport c'è il livello application, nel quale viene effettivamente

svolto il lavoro utile per l'utente. In questo livello si trovano diverse tipologie di oggetti:

- protocolli di supporto a tutte le applicazioni, come per esempio il DNS (Domain Name System);
- protocolli di supporto ad applicazioni di tipo standardizzato, come ad esempio:
  - SNMP (Simple Network Management Protocol) per la gestione della rete;
  - FTP (File Transfer Protocol per il trasferimento di file;
  - SMTP e POP3 (Simple Mail Transfer Protocol e Post Office Protocol per la posta elettronica;
  - HTTP (HyperText Transfer Protocol alla base del World Wide Web);
- applicazioni scritte in conformità ai protocolli di cui sopra;
- applicazioni proprietarie, basate su regole di dialogo private (ad esempio, un'applicazione di tipo client/server per la gestione remota di un magazzino).

6.1 Il DNS

Poiché riferirsi a una risorsa (sia essa un host oppure l'indirizzo di posta elettronica di un utente) utilizzando un indirizzo IP numerico (della forma x.y.z.w) è estremamente scomodo, si è creato un meccanismo tramite il quale tali risorse possono essere identificate tramite un nome logico, cioé una stringa di caratteri (molto pù comprensibile per un essere umano) quale ad esempio:

- `fisica.unige.it` (riferimento ad un host);
- `john@cern.ch` (indirizzo di posta elettronica).

La corrispondenza fra gli indirizzi IP numerici ed i nomi logici si effettua mediante l'uso del DNS. Esso consiste di:

1. uno schema gerarchico di nominazione, basato sul concetto di dominio (domain);

com	aziende
edu	università
gov	istituzioni governative
mil	istituzioni militari
net	fornitori d'accesso
org	organizzazioni non-profit

Tabella 1. *Top-level domanin negli USA*

2. un database distribuito che implementa lo schema di nominazione;
3. un protocollo per il mantenimento e la distribuzione delle informazioni sulle corrispondenze.

Il funzionamento, in breve, è il seguente:

- quando un'applicazione deve collegarsi ad una risorsa di cui conosce il nome logico (ad es. fisica.unige.it), invia una richiesta al DNS server locale (l'applicazione chiama per questo una apposita procedura di libreria detta resolver);
- il DNS server locale, se conosce la risposta, la invia direttamente al richiedente. Altrimenti interroga a sua volta un DNS server di livello superiore, e cos via. Quando finalmente arriva la risposta, il DNS server locale la passa al richiedente;
- quando l'applicazione riceve la risposta (costituita del numero IP della risorsa in questione) crea una connessione TCP con la (o spedisce segmenti UDP alla) destinazione, usando l'indirizzo IP ricevuto.

Lo spazio dei nomi DNS è uno spazio gerarchico, organizzato in dominii, ciascuno dei quali può avere dei sottodominii. Esiste un insieme di dominii di massimo livello (top-level domain), i più alti nella gerarchia. Nel caso di un host, la forma del nome logico è costituita da un certo numero di sottostringhe separate da punti, come nell'esempio seguente:

```
host.subdomain3.subdomain2.subdomain1.topleveldomain
```

dove:

au	Australia
ch	Svizzera
fr	Francia
it	Italia
jp	Giappone
uk	Inghilterra

Tabella 2. *Top-level domanin fuori dagli USA*

cern.ch	CERN di Ginevra, Svizzera
cnr.it	Consiglio Nazionale delle Ricerche, Italia
mit.edu	Massachussetts Institute of Technology, USA
nasa.gov	NASA, USA
unige.ch	Università di Ginevra, Svizzera
unige.it	Università di Genova, Italia
fisica.unige.it	Dip. di Fisica, Università di Genova

Tabella 3. *Estensioni di dominio*

- la prima sottostringa (quella più a sinistra) identifica il nome dell'host;
- le altre sottostringhe (tranne quella più a destra) identificano ciascuna un sottodominio del dominio di cui alla sottostringa seguente;
- l'ultima sottostringa (quella più a destra) identifica il top-level domain di appartenenza.

Per gli USA sono definiti, fra gli altri, i top-level domain mostrati in Tab. Fuori degli USA, ogni nazione ha un suo top-level domain (vedi Tab. 2). 1.

Ogni dominio è responsabile della creazione dei suoi sottodominii, che devono essere registrati presso una apposità autorità. Esempi di sottodominii sono riportati in Tab. 3 L'estensione di un dominio è del tutto indipendente da quella delle reti e sottoreti IP. Ogni dominio ha la responsabilità di fornire il servizio DNS per quanto di propria competenza. Ossia, deve poter rispondere a interrogazioni riguardanti tutti gli host contenuti nel dominio stesso.

6.2 La posta elettronica

La posta elettronica è uno dei servizi più consolidati ed usati nelle reti. In Internet è in uso da circa 20 anni, e prima del WWW era senza dubbio il servizio più utilizzato. Un servizio di posta elettronica, nel suo complesso, consente di effettuare le seguenti operazioni:

- comporre un messaggio;
- spedire il messaggio (a uno o più destinatari);
- ricevere messaggi da altri utenti;
- leggere i messaggi ricevuti;
- stampare, memorizzare, eliminare i messaggi spediti o ricevuti.

Di norma, un messaggio ha un formato ben preciso. In Internet un messaggio ha un formato costituito da un header e da un body, separati da una linea vuota. Lo header è a sua volta costituito da una serie di linee, ciascuna relativa a una specifica informazione (identificata da una parola chiave che è la prima sulla linea); alcune informazioni sono:

To	indirizzo di uno o più destinatari.
From	indirizzo del mittente.
Cc	indirizzo di uno o più destinatari a cui si invia per conoscenza.
Bcc	blind Cc: gli altri destinatari non sanno che anche lui riceve il messaggio.
Subject	argomento del messaggio.
Sender	chi materialmente effettua l'invio (ad es. nome della segretaria).

Il body contiene il testo del messaggio, in caratteri ASCII. L'ultima riga contiene solo un punto, che identifica la fine del messaggio. Gli indirizzi di posta elettronica in Internet hanno la forma:

`username@hostname`

dove username è una stringa di caratteri che identifica il destinatario, e hostname è un nome DNS oppure un indirizzo IP. Ad esempio, apweb@inwind.it è il mio indirizzo di posta elettronica.

La posta elettronica viene implementata in Internet attraverso la cooperazione di due tipi di sottosistemi:

- Mail User Agent (MUA);
- Mail Transport Agent (MTA).

Il primo permette all'utente finale di:

- comporre messaggi;
- consegnarli a un MTA per la trasmissione;
- ricevere e leggere messaggi;
- salvarli o eliminarli.

Il secondo si occupa di:

- trasportare i messaggi sulla rete, fino alla consegna a un MTA di destinazione;
- rispondere ai MUA dei vari utenti per consegnare loro la posta arrivata; in questa fase l'MTA richiede ad ogni utente una password per consentire l'accesso ai messaggi.

Corrispondentemente, sono definiti due protocolli principali per la posta elettronica:

- SMTP (Simple Mail Transfer Protocol per il trasporto dei messaggi:
  - dal MUA di origine ad un MTA;
  - fra vari MTA, da quello di partenza fino a quello di destinazione;
- POP3 (Post Office Protocol versione 3 per la consegna di un messaggio da parte di un MTA al MUA di destinazione.

Recentemente sono stati introdotti altri protocolli più sofisticati, quali IMAP (Interactive Mail Acces Protocol e DMSP (Distributed Mail System Protocol, il cui supporto però non è ancora molto diffuso nel software disponibile agli utenti.

Come avviene la trasmissione di un messaggio? Supponiamo che l'utente

`pippo@topolinia.wd`

spedisca un messaggio a

`minnie@paperinia.wd`

e immaginiamo che:

- Pippo usi un MUA configurato per consegnare la posta ad un SMTP server in esecuzione sull'host mailer.topolinia.wd;
- Minnie abbia un MUA configurato per farsi consegnare la posta da un POP3 server in esecuzione sull'host mailer.paperinia.wd.

La sequenza di azioni che hanno luogo è la seguente:

1. Pippo compone il messaggio col suo MUA, che tipicamente è un programma in esecuzione su un PC in rete;
2. appena Pippo preme il pulsante SEND, il suo MUA:
   - interroga il DNS per sapere l'indirizzo IP dell'host mailer.topolinia.wd;
   - apre una connessione TCP ed effettua una conversazione SMTP con il server SMTP in esecuzione sull'host mailer.topolinia.wd, per mezzo della quale gli consegna il messaggio;
   - chiude la connessione TCP;
3. Pippo se ne va per i fatti suoi;
4. il server SMTP di mailer.topolinia.wd:
   - chiede al DNS l'indirizzo IP di paperinia.wd;
   - scopre che è quello dell'host mailer.paperinia.wd;
   - apre una connessione TCP e poi una conversazione SMTP con il server SMTP in esecuzione su quell'host e gli consegna il messaggio scritto da Pippo;

5. Minnie lancia il suo MUA;
6. appena Minnie preme il pulsante "check mail", il suo MUA:
   - interroga il DNS per avere l'indirizzo IP dell'host mailer.paperinia.wd;
   - apre una connessione TCP e poi una conversazione POP3 col server POP in esecuzione su mailer.paperinia.wd e preleva il messaggio di Pippo, che viene mostrato a Minnie.

Si noti che topolinia.wd e paperinia.wd in genere corrispondono a un dominio nel suo complesso e non ad un singolo host, al fine di rendere gli indirizzi di posta elettronica indipendenti da variazioni del numero, dei nomi logici e degli indirizzi IP degli host presenti nel dominio. Nel DNS ci sono opportuni record detti di tipo MX (Mail Exchange), che si occupano di indicare quale host effetttivamente fa da server SMTP per un dominio. Nel nostro esempio avremo, nel DNS server di Paperinia, i record:

mailer.paperinia.wd	A	100.10.10.5
paperinia.wd	MX	mailer.paperinia.wd

Il secondo record è un record di tipo MX, e indica che tutta la posta in arrivo per

`chiunque@paperinia.wd`

deve essere consegnata all'host mailer.paperinia.wd, e cioé quello che ha l'indirizzo IP

`100.10.10.5`

Inoltre, non è detto che mailer.topolinia.wd consegni i messagi direttamente a mailer.paperinia.wd. È possibile che le macchine siano configurate in modo da trasferire i messaggi attraverso un certo numero di server SMTP intermedi.