

La tecnologia delle reti

Aspetti evolutivi delle reti

Una **rete informatica** è un insieme di sistemi per l'elaborazione delle informazioni messi in comunicazione tra loro.

Il termine **telematica** indica il settore applicativo dell'informatica in rete ed evidenzia l'integrazione tra tecnologie informatiche e tecnologie delle telecomunicazioni.

Le reti di calcolatori sono nate come mezzo con cui collegare terminali periferici ad elaboratori centrali detti **mainframe**. Di anno in anno sono diventate sempre più importanti, ma soltanto con la comparsa del personal computer, hanno assunto un ruolo insostituibile. Infatti, un insieme di personal computer anche se molto potenti, non può da solo rimpiazzare un mainframe, in quanto questo è anche il sistema del quale gli utenti condividono le informazioni e le risorse hardware e software.

Nel tempo si è passati dal modello **mainframe-terminali**, in cui la potenza di calcolo è concentrata in un unico grande elaboratore, a cui si accede per mezzo di un certo numero di terminali, a quello attuale, in cui vi è un grande numero di elaboratori *autonomi e interconnessi*:

- **autonomi**, significa che non deve esserci tra essi una relazione tipo master/slave (per esempio, l'uno non può forzare lo spegnimento dell'altro);
- **interconnessi**, significa che devono essere capaci di scambiare informazioni.

I servizi per gli utenti e le aziende

Gli utenti delle reti non hanno più la necessità di condividere il processore e la memoria centrale, come accadeva nel modello mainframe-terminali; le esigenze importanti diventano:

- condividere il software;
- consultare archivi comuni;
- comunicare dati tra i sistemi;
- condividere dispositivi e connettersi a Internet.

Le reti di computer si sono diffuse perché offrono numerosi servizi sia al singolo individuo che all'organizzazione aziendale. Per questi ultimi, in particolari, i vantaggi sono stati:

- *condivisione delle risorse software e dei dati*: si possono rendere disponibili a chiunque programmi e informazioni, anche a distanza di migliaia di chilometri. Con una rete di A computer si può installare e configurare il software riducendo considerevolmente il tempo necessario per rendere i programmi disponibili per l'intera organizzazione aziendale. Si possono anche porre restrizioni sull'accesso ai programmi per fare in modo, per esempio, che il numero degli utenti di un word processor non ecceda il numero delle copie per cui si possiede la licenza d'uso.
- *condivisione di dispositivi*: le reti di computer possono condividere diversi dispositivi, quali stampanti, modem-fax, scanner, hard disk, floppy disk, ecc.
- *un migliore rapporto prestazioni/costo*: rispetto ai sistemi centralizzati, i piccoli sistemi collegati in una rete offrono un basso costo dell'hardware e, con lo stesso numero di utenti, una velocità di utilizzo superiore, in quanto ogni utente non condivide in modo significativo hardware, ma principalmente dati, e solo talvolta il software.
- *estensione semplificata e graduale dei sistemi hardware*: ogni sistema può essere configurato ed espanso a seconda delle esigenze specifiche dell'azienda, e quindi con hardware differenziati e in tempi differenti; è possibile inoltre aumentare le prestazioni del sistema, aumentando il numero di elaboratori (*scalabilità*).
- *maggior affidabilità del sistema* con l'uso di componenti hardware e software tolleranti rispetto ai guasti: il guasto di un personal computer non blocca il lavoro degli altri, come può succedere quando si ferma un minisistema o un mainframe.

Il modello client/server

Alla fine degli anni '70 il paradigma incentrato sui mainframe e sul concetto di elaborazione time-sharing lasciò il posto al nuovo approccio, che si avvale del paradigma **client/server**, basato sui messaggi. Le nuove tecnologie di elaborazione possono essere infatti considerate come diverse strategie di implementazione, che sfruttano comunque questo paradigma.

L'elaborazione client/server viene realizzata per mezzo di un programma utente (**client** o chiedente) che richiede certi servizi, come l'accesso ai dati, la stampa o l'esecuzione di altri processi, a un programma

server (servente) che li fornisce.

Le reti di computer si basano sul modello client/server. In questo modello la comunicazione generalmente ha la forma di un messaggio (insieme di caratteri e di dati) a un servente da parte di un cliente, che chiede di eseguire un certo lavoro. Il servente esegue il lavoro e restituisce la risposta.

A seconda dei ruoli dei computer connessi, le reti si classificano in tre tipi:

- **reti client/server**, in cui sono presenti computer che funzionano da client uno o più computer che funzionano da server;
- **reti peer to peer**, che non hanno server dedicati e usano la rete per condividere risorse;
- **reti ibride**, che sono reti client/server, ma possono condividere risorse come reti peer to peer.

Reti client/server

Le reti client/server o a *server dedicato* sono caratterizzate dalla presenza di un elaboratore chiamato server che ha molti compiti, tra i quali i più importanti sono: fornire risorse software, permettere la condivisione di dispositivi e garantire la sicurezza nell'accesso alla rete. Questo computer è indicato spesso anche con il termine **host**. Le reti a server dedicato funzionano secondo il modello client/server e generalmente il computer più potente ha il compito di server per far fronte alle molteplici richieste provenienti dai client in modo istantaneo.

Le reti a server dedicato hanno accessi riservati, quindi è necessario possedere i permessi necessari per avere accesso al sistema. Quando l'utente fornisce la propria e la propria **password** viene accettato dal sistema, ma il permesso vale per l'uso delle sole directory stabilite dal server all'atto della creazione del profilo dell'utente. Solo l'**amministratore della rete** possiede i permessi di accesso a tutte le directory di sistema per la gestione dell'intera rete.

I vantaggi delle reti basate su server dedicato sono:

- forte sicurezza centralizzata;
- memorizzazione degli archivi a livello centrale, che permette a tutti gli utenti di lavorare sugli stessi dati;
- maggiore sicurezza rispetto alle intrusioni, perché ciascun utente ha aree personalizzate di azione, chiamate *domini di directory*, con un metodo di autenticazione centralizzato;
- facilità di gestione di un grande numero di utenti.

A fronte dei vantaggi elencati, gli svantaggi sono legati al costo elevato dell'hardware e del software di rete e alla necessità di una figura di amministratore della rete.

Reti peer to peer

Una rete peer to peer è definita dall'assenza di un controllo centrale, per cui non ci sono dedicati e ogni utente può condividere semplicemente le risorse di rete e lo spazio su disco.

Le reti peer to peer sono organizzate in gruppi di lavoro (*workgroup*) che hanno un livello di sicurezza minimo. Generalmente l'accesso alla rete mediante login e password permette all'utente di utilizzare indiscriminatamente dispositivi e directory.

Le reti peer to peer non sono ottimizzate per la condivisione delle risorse, per cui spesso si può avere un degrado nelle prestazioni. I vantaggi sono relativi al risparmio, sia in termini di costi specifici, sia di risorse umane, perché non è richiesto l'amministratore di rete. Gli svantaggi sono relativi al basso livello di sicurezza e alla dimensione della rete che, per funzionare in modo efficiente, deve avere meno di dieci utenti.

Reti ibride

Le reti ibride hanno computer che svolgono tutti i ruoli, client, server e peer: hanno generalmente sia domini di directory, controllati centralmente, sia gruppi di lavoro. Questo significa che mentre la maggior parte delle risorse è gestita dai server, gli utenti della rete hanno accesso ad alcune risorse condivise nel gruppo di lavoro in modo paritetico (peer).

I vantaggi e gli svantaggi di questo tipo di rete sono quelli delle reti a server dedicato e buona parte di quelle delle reti paritetiche.

Classificazione delle reti

Un aspetto particolarmente rilevante per lo studio delle reti è rappresentato dalla distanza che separa i calcolatori in rete.

Reti Locali LAN (Loca/ Area Network)

Queste reti coprono un'area corrispondente a dimensioni varianti tra quelle di un ufficio e quelle di una struttura aziendale composta da più edifici. Sono utilizzate soprattutto per consentire ad un gruppo di utenti di condividere i dati o i programmi software utili ad un certo tipo di attività aziendale.

La LAN permette anche di condividere dispositivi hardware, quali stampanti laser oppure archivi di dati e

connessioni a internet. Le ridotte dimensioni consentono una notevole velocità nella trasmissione e ricezione dei dati, nonché una semplificazione delle procedure, per garantire i necessari livelli di sicurezza nella gestione dei dati. Una rete locale è un mezzo di trasporto equamente condiviso tra tutte le stazioni che vi si collegano, ad alta velocità e basso tasso di errore.

Reti metropolitane MAN (Metropolitan Area Network)

Le MAN sono estensioni delle reti locali in ambito urbano.

Reti geografiche WAN (Wide Area Network)

Queste reti hanno dimensioni geograficamente estese, da una città all'intero pianeta. Una rete con tali estensioni presenta ovviamente un limite fisico alla velocità di funzionamento, soprattutto perché si utilizza prevalentemente l'infrastruttura esistente, cioè quella realizzata per fornire il servizio telefonico.

La diffusione dei dati sul canale trasmissivo

Dal punto di vista delle modalità di diffusione dei dati, esistono due tipi di reti:

- reti a diffusione globale;
- reti punto a punto.

Nelle **reti a diffusione globale** ogni trasmissione viene ricevuta da tutti gli *host* connessi alla rete. Ciò significa che il mezzo trasmissivo è condiviso e questo avviene tipicamente nelle LAN. Le reti che condividono lo stesso mezzo trasmissivo prendono il nome di **broadcast**.

Nelle **reti punto a punto** i due dispositivi che vogliono scambiarsi i dati sono direttamente collegati l'uno all'altro.

Nelle reti WAN gli *host*, chiamati anche **end system**, non sono connessi direttamente tra loro, ma sono connessi attraverso nodi intermedi, chiamati **intermediate system** e generalmente rappresentati da **router** che svolgono le funzioni di indirizzamento.

Regole per il trasferimento dei dati

Esistono tre tipi di regole per il trasferimento dei dati che danno origine a linee diverse o ad usi diversi della stessa linea:

- linea **simplex**, per la quale il trasferimento dei dati è monodirezionale, cioè il sistema che riceve la comunicazione non è in grado di rispondere (es. trasmissione radiotelevisiva);
- linea **half-duplex**, per la quale la comunicazione è possibile in entrambe le direzioni, ma uno solo per volta dei due elementi posti in comunicazione può trasmettere (es. ricetrasmittitori);
- linea **full-duplex**, per la quale la comunicazione è possibile in entrambe le direzioni e gli elementi interconnessi possono fruire del canale di comunicazione in contemporanea (es. comunicazione telefonica).

Le tecniche di commutazione

Per mettere in comunicazione due utenti esistono fondamentalmente due tecniche:

- **commutazione di circuito**, di derivazione dal sistema telefonico;
- **commutazione di pacchetto**, di derivazione dall'informatica.

Le linee commutate, cioè le normali linee telefoniche, sono state basate per lungo tempo su una tecnica di commutazione detta *commutazione di circuito*. Le prime centrali telefoniche funzionavano manualmente. Le operatrici collegavano a richiesta la linea dell'utente, effettuava la chiamata, con l'utente desiderato. La commutazione di circuito crea quindi un vero collegamento fisico tra i due utenti, ed esso resta stabile e riservato a loro per tutta la durata della comunicazione. Questo comporta in media un basso utilizzo del canale trasmissivo, risultando occupato da una comunicazione anche quando i due interlocutori non parlano.

L'utilizzo della commutazione di circuito prevede tre fasi distinte:

- **attivazione del circuito**, in cui si stabilisce la connessione fisica tra l'unità chiamante e l'unità chiamata; nel caso della comunicazione telefonica corrisponde all'attivazione del numero e all'attesa che il numero chiamato risponda.
- **utilizzo del canale trasmissivo**, è la fase in cui i dati possono essere trasmessi senza necessità di particolari controlli; nel caso della comunicazione telefonica è la parte di comunicazione tra i due interlocutori.

- **svincolo**, e la fase in cui la connessione viene chiusa; nella comunicazione telefonica, corrisponde a rimettere nella sede il ricevitore del telefono con la riattivazione del servizio da parte della centrale telefonica.

L'alternativa alla commutazione di circuito, per un migliore utilizzo dei canali trasmissivi, è rappresentata dalla *commutazione di pacchetto*, basata su sistemi digitali sia per l'instradamento che per la trasmissione dei dati.

Il **pacchetto** è costituito da due parti: la parte di dati (*payload*) e la parte di intestazione (*header*) con le informazioni riguardanti la connessione.

Nell'intestazione sono specificati l'indirizzo del mittente e l'indirizzo del destinatario, oltre a un numero progressivo attribuito al pacchetto specifico nella sequenza dei pacchetti che costituiscono il messaggio da trasmettere.

Questi attributi permettono al pacchetto di essere svincolato dal percorso fisico dei dati, perché, anche se i pacchetti di una stessa sequenza fanno percorsi diversi per giungere a destinazione, il destinatario ha comunque gli elementi per ricostruire la sequenza.

I pacchetti vengono instradati dai **nodi intermedi** su percorsi differenti. Essi svolgono la funzione di **instradamento**, cioè in base all'indirizzo di destinazione del pacchetto decidono su quale canale deve essere instradato per farlo giungere a destinazione nel tempo più breve possibile. I nodi finali sono gli elaboratori connessi alla rete. Quando un elaboratore riceve un pacchetto, esamina l'indirizzo di destinazione; se questo coincide con il proprio indirizzo, il pacchetto viene copiato sul computer locale, altrimenti viene ignorato.

Architettura di rete

Per ridurre la complessità di progetto, le architetture di rete sono organizzate a **livelli**, ciascuno dei quali fornisce al livello superiore i servizi richiesti, mascherando le modalità con le quali sono ottenuti. Le regole e le convenzioni usate nel dialogo tra livelli omologhi sono generalmente conosciute come **protocolli**.

Un insieme di livelli e protocolli è chiamato **architettura di rete**.

Il livello n su un host porta avanti una conversazione con il 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** .

Le entità (*processi*) che effettuano tale conversazione si chiamano **peer entity** (entità di pari livello). Il dialogo tra due *peer entity* di livello n viene realizzato attraverso lo scambio di **PDU** (Protocol/ Data Unit), composta dalla parte di dati e dall'intestazione specifica del livello. In realtà non c'è trasferimento diretto di dati dal livello n di *host1* al livello n di *host2*. Ogni livello di *host1* passa i dati e le informazioni di controllo, al livello sottostante. Al livello 1 c'è il mezzo fisico, attraverso il quale i dati vengono trasferiti da *host1* ad *host2*. Quando arrivano a *host2*, i dati vengono passati da ogni livello (a partire dal livello 1) a quello superiore, fino a raggiungere il livello delle applicazioni.

Ogni livello n comunica con quello direttamente superiore $n+1$ attraverso un'**interfaccia**, che caratterizza le operazioni primitive che possono essere richieste al livello sottostante.

SDU (*Service Data Unit*) è il termine con cui si indicano i dati scambiati attraverso l'interfaccia.

I vantaggi di una buona progettazione delle interfacce sono:

- rendere minimo il numero di informazioni da trasferire;
- possibilità di modificare l'implementazione del livello (per esempio, laddove le linee telefoniche venissero sostituite da canali satellitari) con una più attuale, che offra gli stessi servizi.

Il sistema ISO-OSI

Tutto il sistema hardware e software che ruota attorno al mondo delle reti è regolato da norme precise che fanno riferimento al modello OSI (Open System Interconnection), realizzato nel 1984 dall'ISO (International Standard Organization). Questo modello fornisce uno standard per le interconnessioni in rete dei vari computer. Alla base del modello vi è una suddivisione delle funzioni che un sistema di rete deve svolgere affinché il processo di trasmissione dei dati venga portato a termine.

Il **modello ISO-OSI** è stato creato al fine di produrre uno standard a livello mondiale per guidare sia l'attività di progettazione delle reti di comunicazione, che l'attività di programmazione delle applicazioni di rete.

Esso doveva, infatti, servire come base comune per coordinare gli sforzi degli sviluppatori, per esempio standardizzando la terminologia e definendo quali sono le funzionalità di una rete.

Per gestire la complessità dei problemi, l'OSI ha adottato un approccio a livelli (**layers**): l'intero problema della comunicazione tra due applicazioni è stato scomposto in un insieme di sette livelli, ciascuno dei quali esegue funzioni specifiche. Ogni livello può in generale comunicare solamente con il livello inferiore e fornisce servizi solo a quello superiore. I moduli di pari livello in nodi differenti comunicano tra loro secondo le regole fissate dai protocolli.

1. Livello di collegamento fisico (physical layer)

A questo livello spetta la definizione delle funzioni basilari della connessione fisica, dalla struttura elettronica che realizza il collegamento, agli aspetti tecnici delle interconnessioni.

Questo è il livello che gestisce le caratteristiche hardware.

Il livello fisico si occupa della trasmissione dei singoli bit da un estremo all'altro dei vari mezzi di comunicazione che possono essere le fibre ottiche, le onde radio, i satelliti ecc.

Problemi tipici di questo livello sono: quanti volt devono essere utilizzati per rappresentare un 1 e uno 0, quanti microsecondi richiede la trasmissione di un bit, come si stabilisce la connessione iniziale, in quale direzione vengono trasmessi i dati.

La connessione tipica di una rete locale utilizza sistemi **ethernet**.

Gli elementi che si trovano a livello 1 sono i seguenti:

- Le schede di rete o NIC che vengono installate all'interno del computer nell'alloggiamento della scheda madre dedicato alle schede di espansione. Generalmente sono schede per bus standard PCI per Personal Computer.
- Gli hub o ripetitori sono semplici apparecchiature di livello 1 che collegano tra loro gruppi di utenti.

2. Livello di collegamento dati (data link layer)

Il livello del collegamento dati riguarda i dispositivi che gestiscono il collegamento da un computer a un altro della stessa rete.

Controlla la correttezza delle sequenze di bit trasmesse e ne richiede eventualmente la trasmissione.

Provvede alla formattazione delle informazioni e alla sincronizzazione dei frame, nonché alla correzione e al recupero dei messaggi errati. Un frame (in italiano trama) contiene, a livello del collegamento dati, l'indirizzo di destinazione e, se richiesto da un livello superiore, anche l'indirizzo di origine, e un codice per la correzione e rilevazione degli errori. Il frame costituisce l'entità di dati che è trasmessa senza interruzioni. A livello successivo (livello 3) prende il nome di pacchetto.

Gli elementi di interconnessione della rete a livello 2 sono i seguenti:

- Gli switch sono dispositivi più intelligenti degli hub e si caratterizzano anch'essi per il numero di porte disponibili. Uno switch invia i pacchetti di dati alle porte specifiche dei destinatari, sulla base delle informazioni contenute nelle header di ogni pacchetto. Per isolare la trasmissione dalle altre porte, lo switch stabilisce una connessione temporanea tra la sorgente e il punto di destinazione, chiudendola al termine del collegamento.
- I bridge sono dispositivi del tutto analoghi agli switch, ma hanno solo due porte e quindi sono gli elementi di interconnessione tipici di due LAN.

3. Livello di controllo della rete (network layer)

A questo livello appartengono funzioni tipicamente di rete. Nel livello di rete i messaggi vengono suddivisi in pacchetti che, una volta giunti a destinazione, vengono riassemblati nella loro forma originaria. Il livello di rete si fa carico di scegliere una strada tra quelle disponibili, tramite i router che instradano i pacchetti verso il computer di destinazione. Il protocollo di rete più utilizzato nel livello 3 è il **protocollo IP**. Il principale apparato di interconnessione della rete a livello 3 è il **router**.

Ancora più intelligenti di hub e switch, i router utilizzano un indirizzo IP per determinare il nodo intermedio successivo che deve ricevere il pacchetto. Basandosi su una mappa di rete denominata **tabella di routing**, i router possono fare in modo che i pacchetti raggiungano le loro destinazioni attraverso i percorsi più idonei. Se cade la connessione tra due router, per non bloccare il traffico, il router sorgente può cercare un percorso alternativo. I router definiscono anche i collegamenti tra reti che utilizzano protocolli diversi.

4. Livello del trasporto (transport layer)

Il livello di trasporto gestisce la trasmissione dei pacchetti end-to-end. Ha il compito specifico di assicurare il trasferimento dei dati tra strati di sessione appartenenti a sistemi diversi, geograficamente separati, evitando che sui dati vi siano errori o duplicazioni. È in grado di identificare il destinatario, aprire o chiudere una connessione con il sistema corrispondente, suddividere o riassemblare un testo, controllare e recuperare gli errori, controllare la velocità con cui fluiscono le informazioni.

A questo livello l'esistenza dei livelli inferiori è completamente ignorata: ciò porta a identificare

questo livello come il primo che prescinde dal tipo e dalle caratteristiche della rete utilizzata. Il protocollo standard utilizzato nel livello 4 è il **protocollo TCP** (Transmission Control Protocol). Talvolta viene usato anche il protocollo **UDP** (User Datagram Protocol).

5. Livello di sessione (session layer)

Il livello di sessione gestisce la corretta sincronizzazione della corrispondenza dei dati che verranno poi visualizzati. Instaura una **sessione**, cioè un collegamento logico e diretto tra due interlocutori, organizzandone il dialogo. La modalità del dialogo può essere full-duplex, cioè entrambe le applicazioni in trasmissione e ricezione contemporaneamente, o half-duplex, cioè mentre una stazione trasmette l'altra riceve o viceversa, oppure in simplex, dove una stazione può sempre e solo trasmettere e l'altra sempre e solo ricevere.

Per **sincronizzazione** si intende invece la capacità di sapere sempre fino a che punto la comunicazione sia arrivata a buon fine. Si immagini un trasferimento dati tra due host della durata di alcune ore: la sincronizzazione consiste nel mettere dei punti di controllo nel processo in modo che, se il trasferimento si interrompe, non sia necessario ritrasferire l'intero archivio, ma solo la parte inviata dopo l'ultimo punto di controllo.

6. Livello di presentazione (presentation layer)

Le varie informazioni che viaggiano all'interno della rete subiscono a questo livello una particolare decodifica, che le trasforma in modo da renderle visualizzabili nei normali dispositivi di output a disposizione degli utenti, tipicamente i video e le periferiche di stampa.

La nostra abitudine all'informatica dei Personal Computer può indurci a pensare erroneamente che il codice ASCII utilizzato dai terminali sia uguale per tutti, dimenticando realtà differenti, cioè terminali che utilizzano codifiche diverse, che vengono supportate correttamente dal livello di presentazione.

Il livello di presentazione gestisce quindi i formati di **conversione dei dati**, cioè effettua tutte le opportune conversioni in modo da compensare eventuali differenze di rappresentazione e di formato dei dati in arrivo e in partenza.

7. Livello di applicazione (application layer)

Il livello di applicazione riguarda i cosiddetti programmi applicativi. Questo livello gestisce la visualizzazione dei dati relativa a programmi di login remoto, file transfer, posta elettronica. Per la gestione dei Personal Computer, il problema si presenta quando due sistemi che vogliono comunicare possiedono video o tastiere diverse, e quindi non compatibili. Per esempio, per spostare il cursore a inizio linea o per cancellare lo schermo, ogni scheda ha i suoi comandi specifici: invece di dotare tutti i sistemi di opportuni traduttori per tutti i possibili interlocutori, è evidentemente molto più semplice definire un terminale virtuale di rete come standard unico virtuale, a cui tutti i corrispondenti terminali reali devono adeguarsi per comunicare.