

Reti e protocolli

1 Rete di calcolatori

Una **rete di calcolatori** è un sistema informatico costituito da due o più calcolatori collegati attraverso un sistema di comunicazione.

L'esistenza delle reti di calcolatori è importante in quanto rende possibile le **applicazioni distribuite**: applicazioni composte da diversi elementi, che sono messi in esecuzione su macchine diverse, e che cooperano comunicando tra loro attraverso una rete di calcolatori.

2 Quadro storico

1. All'inizio degli anni '60, cominciano a esistere i calcolatori commerciali, ma sono troppo costosi per permetterne un uso individuale (anzi, solo le aziende di medie-grandi dimensioni si possono permettere un calcolatore). In compenso, però, le esigenze dei singoli utenti sono abbastanza limitate da permettere la condivisione della potenza di calcolo fornita da un singolo elaboratore.

Allora, in questo periodo, si hanno “reti” composte da un unico elaboratore centrale a cui sono connessi, mediante linee telefoniche, dei terminali remoti “stupidi”, senza capacità di elaborazione, che sono in grado di effettuare l'input e l'output da parte del calcolatore.

Le connessioni sono tipicamente organizzate secondo un'architettura a *stella*: ogni terminale è connesso esclusivamente con l'elaboratore centrale. Infatti, lo scopo principale della comunicazione è quello di poter disporre i terminali remotamente rispetto all'elaboratore.

2. Tra la fine degli anni '60 e gli anni '70, grazie ai circuiti integrati, il costo dell'hardware diminuisce più velocemente di quello delle linee telefoniche. Ciò ha principalmente due conseguenze:
 - Si diffondono i *terminali intelligenti*, in grado di effettuare una qualche pre-elaborazione dei dati in locale.
 - Risulta conveniente utilizzare la stessa linea telefonica per più terminali geograficamente vicini, mediante un'architettura con *concentratori* o con *linee multipunto*.

3. Negli anni '80, il costo dell'hardware continua a scendere, si diffondono i personal computer, e diventa quindi sempre più importante condividere applicazioni e dati (anche perché i primi personal computer avevano una capacità di memorizzazione molto limitata, quindi, soprattutto nelle aziende, era necessario poter accedere a dati situati su calcolatori centrali più potenti). Perciò, nascono le reti locali e geografiche di micro/personal computer.

3 Tecnologia di rete

Le reti di calcolatori possono essere caratterizzate rispetto a vari aspetti. Tutti questi, complessivamente, costituiscono la **tecnologia di rete**, che può essere riassunta da alcuni parametri:

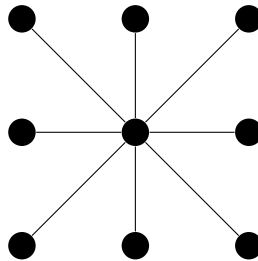
- la *distanza* di trasmissione possibile;
- la *velocità* di trasmissione (in bit al secondo);
- il *costo*.

3.1 Topologia di rete

Il termine **topologia di rete** indica la disposizione fisica dei componenti che realizzano la rete, e in particolare come sono connessi tra loro.

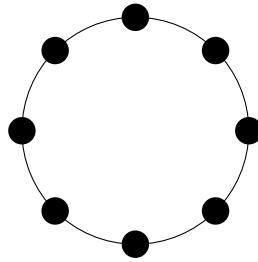
Esistono varie topologie “standard”, tra cui:

- Rete a **stella**:



qualunque comunicazione tra i calcolatori periferici deve necessariamente passare attraverso il calcolatore centrale.

- Rete ad **anello**:

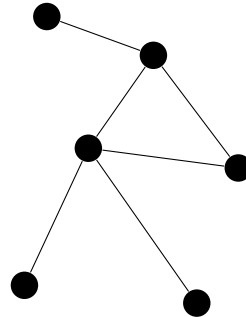


ogni computer è collegato ad altri due, in modo circolare; tutti si possono parlare, ma la comunicazione con i vicini risulta privilegiata rispetto a quella con i calcolatori più lontani.

- Rete a **bus**:



- Topologia irregolare:



i computer sono organizzati come opportuno, ad esempio in base ai diversi ruoli di ciascuno di essi.

3.2 Mezzi trasmissivi

Un'altra caratterizzazione delle reti è data dal tipo di mezzo trasmissivo utilizzato. Ne esistono molti, tra cui

- doppino telefonico,
- cavo coassiale,
- fibra ottica,
- onde elettromagnetiche,
- porte a infrarossi,
- ecc.

e, in generale, anche all'interno di una stessa rete, è possibile usare combinazioni di mezzi diversi.

I mezzi trasmissivi determinano il costo e la **banda** (cioè la quantità di informazione che può essere trasmessa nell'unità di tempo – in pratica, la velocità) della rete.

3.3 Senso di trasmissione

La trasmissione tra due nodi può essere:

simplex: il senso di trasmissione è fisso, unico;

half-duplex: la trasmissione è possibile *alternativamente* nei due sensi;

full-duplex: la trasmissione è possibile *contemporaneamente* nei due sensi.

3.4 Commutazione

In generale, in una rete non c'è comunicazione diretta tra tutti i nodi. Allora, per collegare due nodi, occorre stabilire tra di essi un collegamento. Esistono fondamentalmente due modi di farlo:

commutazione di circuito: si crea un collegamento “fisico” (come se ci fosse un filo elettrico connesso tra i due nodi);

commutazione di pacchetto: il collegamento è “virtuale”.

3.4.1 Commutazione di circuito

La connessione a commutazione di circuito era quella tipica delle comunicazioni telefoniche “classiche”:¹ componendo un numero si dava a una serie di centraline l'istruzione di realizzare un circuito fisico tra il telefono chiamante e il telefono chiamato.

Una connessione di questo tipo garantisce una determinata banda, ma, di conseguenza, tale banda risulta occupata anche mentre non è concretamente sfruttata dalla comunicazione. Perciò, la commutazione di circuito comporta costi elevati.

¹Al giorno d'oggi, con la tecnologia VOIP, le telefonate viaggiano su Internet, quindi anche per il telefono usa la commutazione di pacchetto.

3.4.2 Commutazione di pacchetto

Con una connessione a commutazione di pacchetto, il traffico viene diviso in tanti piccoli messaggi, chiamati appunto **pacchetti**, ciascuno di poche centinaia di byte. Ogni tratto della rete viene occupato da un pacchetto solo per il tempo strettamente necessario. Allora:

- quando un calcolatore collegato in rete non utilizza la banda, questa può essere utilizzata da altri calcolatori, consentendo così più comunicazioni simultanee;
- pacchetti diversi possono seguire percorsi diversi per arrivare alla stessa destinazione;
- si ha lo svantaggio che la comunicazione è frammentata: i dati da inviare devono prima essere divisi in pacchetti, e il destinatario deve poi riuscire a ricostruirli dai pacchetti ricevuti.

Le reti a commutazione di pacchetto forniscono due tipi di servizi:

Servizi a circuito virtuale: si stabilisce (appunto) un circuito virtuale tra mittente e destinatario, e viene mantenuto l'ordinamento tra messaggi diversi inviati lungo tale circuito (ma essi possono comunque compiere strade diverse nella rete per raggiungere il destinatario).

Servizi a datagramma: non viene creato alcun circuito tra mittente e destinatario. Ogni singolo messaggio viene gestito indipendentemente dai precedenti e dai successivi, e perciò l'ordine in cui arrivano può essere diverso da quello in cui sono stati inviati.

4 Modello di interazione client-server

Il principale modello di interazione utilizzato nelle applicazioni di rete è il modello **client-server**.

- Un **server** è un programma che offre un servizio, in questo caso tramite la rete: esso si “affaccia” alla rete a un indirizzo ben noto, e rimane passivamente in attesa di richieste da parte dei client.
- Un **client** è un programma che vuole usufruire del servizio offerto dal server: quando ne ha bisogno, manda una richiesta all'indirizzo del server, e ottiene da questo una risposta.

In generale, nulla vieta che uno stesso programma si comporti da cliente nei confronti del server A e da server nei confronti del client B . Però, quando ci sono solo due programmi, uno che fa il client e uno che fa il server, si parla di *architettura* client-server, che non deve allora essere confusa con il caso generale dell'*interazione* client-server.

Un client che vuole richiedere il servizio deve, per prima cosa, *connettersi* al server. A tale scopo:

- Il server deve essersi dichiarato disposto ad accettare richieste di connessione da parte di client, cioè apre la connessione in modo passivo (e, solitamente, non conosce a priori l'identità dei suoi client).
- Un client chiede in modo attivo l'apertura della connessione con il server; per farlo, deve conoscere l'identità (cioè l'indirizzo) del server.

Si osserva che la connessione è un'operazione asimmetrica.

Un'altra caratteristica tipica di questo modello è che, normalmente, la vita di un server si prolunga oltre il tempo dell'interazione con il singolo client.

5 Protocollo di comunicazione

Il termine **protocollo di comunicazione** indica l'insieme di regole di comunicazione che devono essere seguite da due interlocutori affinché essi possano comprendersi.

I protocolli per le reti di calcolatori sono organizzati secondo una *pila* (stack), una gerarchia a strati: salendo nella gerarchia, cresce il livello di astrazione dei servizi offerti dai protocolli, e ogni protocollo si appoggia ai protocolli di più basso livello per fornire un servizio di qualità superiore.

6 Modello ISO/OSI

L'**Open Systems Interconnection**, meglio noto come modello o stack **ISO/OSI**, è uno standard *de iure* per reti di calcolatori, stabilito nel 1978 dall'International Organization for Standardization (ISO).

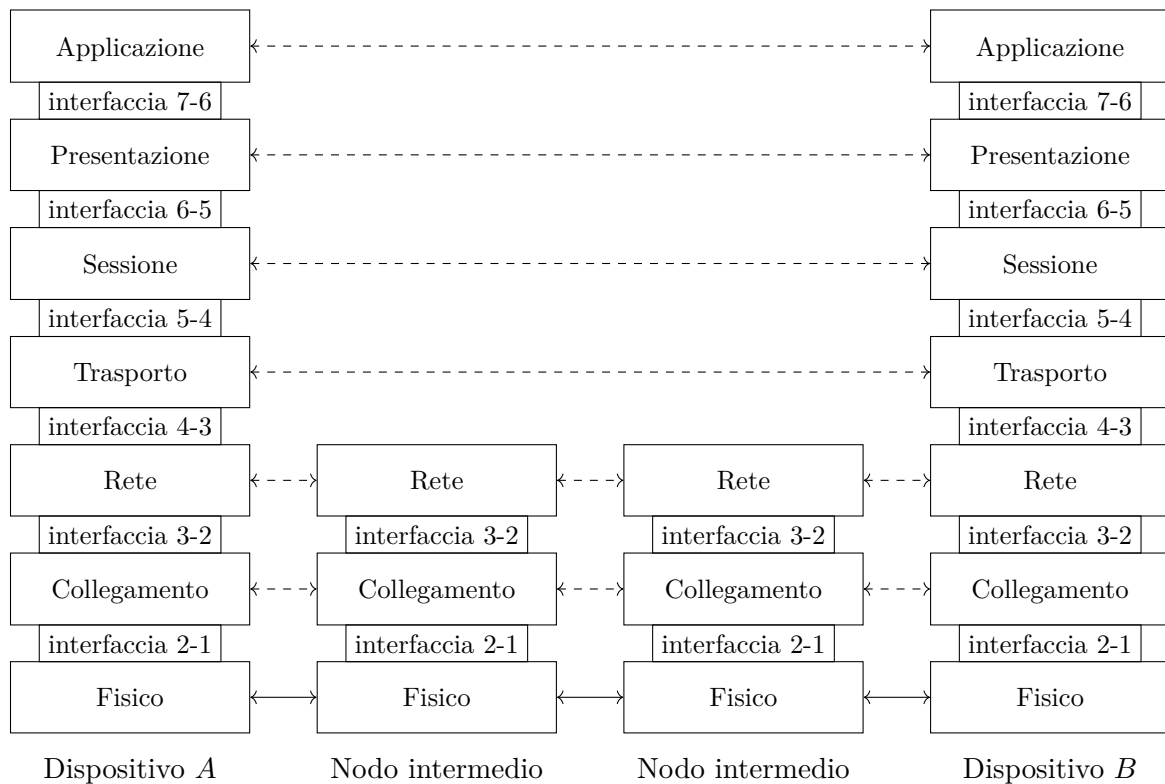
Questo modello prevede una pila di protocolli suddivisa in 7 livelli (layer):

	Livelli	Unità di dato	
Livelli degli host	7	Applicazione	Dati
	6	Presentazione	Dati
	5	Sessione	Dati
	4	Trasporto	Segmenti
Livelli dei mezzi	3	Rete	Pacchetti
	2	Collegamento	Trame
	1	Fisico	Bit

Per ciascuno di questi livelli, sono definiti diversi protocolli. Ad esempio:

Livello	Esempi di protocolli
7 Applicazione	HTTP, FTP, DNS, SNMP, Telnet
6 Presentazione	SSL, TLS
5 Sessione	NetBIOS, PPTP
4 Trasporto	TCP, UDP
3 Rete	IP, ARP, ICMP, IPSec
2 Collegamento	PPP, ATM, Ethernet
1 Fisico	Ethernet, USB, Bluetooth

Nella comunicazione tra due nodi A e B , il livello n del nodo A può scambiare informazioni con il livello n del nodo B , e non con gli altri. In realtà, per realizzare questa comunicazione, il livello n di A utilizza il livello sottostante ($n - 1$), che a sua volta utilizza quello ancora sottostante ($n - 2$), e così via, fino ad arrivare al livello fisico (l'unico al quale avviene concretamente la comunicazione); quest'ultimo parla con il livello fisico di B , e, da qui, si "risale" fino al livello n . Inoltre, spesso il percorso di comunicazione da A a B passa da dei nodi intermedi: anche per questi nodi, si ha lo stesso meccanismo dei livelli.



Oltre a consentire la comunicazione nelle situazioni “normali”, i vari protocolli gestiscono una serie di possibili problemi:

- malfunzionamenti hardware;
- congestione della rete: una macchina congestionata può sopprimere ulteriore traffico;
- ritardo o perdita di pacchetti: il software deve riconoscere e adattarsi a ritardi lunghi e di durata variabile;
- alterazione dei dati dovuta a possibili errori di trasmissione;
- duplicazione dei dati, o errori nella sequenza di trasmissione, ad esempio dovuti a reti che offrono più percorsi alternativi.

Perciò, i programmatori delle applicazioni distribuite non si devono in genere preoccupare di gestire tutti questi problemi.

6.1 Livello fisico

Si occupa della gestione fisica (cioè, tipicamente, meccanica ed elettrica/ottica) dell'interfaccia con il mezzo fisico usato per il collegamento.

I protocolli a livello fisico definiscono le regole per l'interpretazione dei segnali scambiati attraverso il mezzo trasmissivo (tensioni, modulazione, ecc.).

6.2 Livello di collegamento (data link)

Si occupa dello spostamento di una stringa di bit da un nodo all'altro, con un certo grado di affidabilità. Esso svolge tre principali funzioni: distingue il segnale dal rumore, riconosce certi tipi di errori, e li corregge (almeno in parte).

6.3 Livello di rete

Si occupa dell'indirizzamento dei messaggi lungo la rete, implementando gli opportuni meccanismi di commutazione. Esso permette quindi di stabilire connessioni tra due nodi (*host*) della rete.

Il servizio fornito è *funzionalmente* indipendente dal particolare tipo di rete adottata, ma reti diverse possono comunque comportare differenze dal punto di vista delle prestazioni (ad esempio per quanto riguarda l'affidabilità).

6.4 Livello di trasporto

Permette di stabilire connessioni tra applicazioni diverse su host diversi. Tali connessioni possono eventualmente essere *affidabili*: chi invia un pacchetto viene poi informato sul fatto che esso sia arrivato o meno, così da poter compiere delle azioni correttive se il pacchetto si perde.

Grazie a questi meccanismi, il livello di trasporto, rendere indistinguibili tipi di reti diversi (cosa già fatta, dal punto di vista funzionale, dal livello di rete) anche per quanto riguarda le prestazioni (compresa l'affidabilità):

- fornisce connessioni con una qualità di servizio richiesta;
- gestisce la correttezza delle informazioni trasmesse e il loro ordinamento.

Questo livello è il più importante per la programmazione di applicazioni distribuite: di fatto, i servizi offerti dai livelli superiori (sessione, presentazione e applicazione) sono implementati da programmi che si appoggiano al livello di trasporto.

7 Reti locali

Gli standard utilizzati per le reti locali (cioè a livello data link) sono diversi:

- Ethernet (IEEE 802.3), Fast Ethernet, e versioni successive;
- Token Ring (introdotto da IBM);
- ecc.

La comunicazione tra calcolatori deve però essere possibile anche fra reti di tipi diversi: quello che si vuole ottenere è un'unica rete logica, una “rete di reti” che renda trasparenti i dettagli delle reti locali nelle quali si trovano i calcolatori connessi.

8 Internet

Una rete di reti è detta **internet** (con la i minuscola), ed è definita fornendo i protocolli per trasferire le informazioni tra le varie reti che connette.

Affinché sia possibile la comunicazione tra due reti diverse, è necessaria una macchina, chiamata **gateway**, che sia condivisa tra le due reti, e in grado di connettersi a entrambi i tipi di reti (nel caso siano diversi). Non serve, però, che tutte le reti di una internet siano direttamente connesse l'una con l'altra: i dati possono anche compiere percorsi a più “salti”, passando da gateway e reti intermedi. Ad esempio, nella internet rappresentata in questa figura,



un calcolatore situato nella rete 1 può comunicare con uno nella rete 3 passando tramite il gateway 1, la rete 2 e il gateway 2.

Prende il nome di **Internet** (con la I maiuscola) la più diffusa internet del mondo. Essa è realizzata usando i protocolli **TCP/IP**, che, in generale, sono quelli più comunemente utilizzati per le internet.

Si chiama invece **intranet** una rete privata (tipicamente interna a un'azienda) basata sulle stesse tecnologie di Internet. Infatti, la diffusione di Internet ha favorito lo sviluppo di numerose applicazioni distribuite basate su TCP/IP, e ciò ha reso conveniente l'uso degli stessi protocolli anche in ambito locale. Attualmente, la maggior parte delle reti locali sfruttano quindi TCP/IP come protocolli base.

8.1 Architettura di Internet

A livello logico, tante macchine condividono l'accesso alla rete Internet, e su ciascuna di esse possono essere eseguite applicazioni diverse, che comunicano tra di loro attraverso la rete.

L'architettura fisica è invece composta da reti locali:

- i calcolatori sono connessi alle varie reti locali;
- le reti locali sono connesse tra loro da dei router/gateway.

Inoltre, solitamente, la connessione tra un calcolatore / una rete locale e le altre avviene mediante un Internet provider.

8.2 Storia di Internet

1. Alla fine degli anni '60, ARPA (Advanced Research Project Agency, poi rinominata DARPA, Defense Advanced Research Project Agency) sviluppa ARPANET, che connette laboratori di ricerca, università e reti governative.
2. Alla fine degli anni '70, DARPA finanzia lo sviluppo di protocolli a commutazione di pacchetto. Nasce TCP/IP, al quale ARPANET inizia a “convertirsi” nel 1980.
3. Anni '80:
 - Nel 1983, la conversione di ARPANET a TCP/IP è completa, e l'ufficio del Segretario della Difesa degli Stati Uniti ordina che tutti i computer connessi a reti a lunga distanza usino TCP/IP. Inoltre, nello stesso anno, la rete governativa e militare MILNET si separa da ARPANET.
 - DARPA finanzia lo sviluppo di Berkeley Unix, che comprende un'implementazione di TCP/IP, la quale introduce l'astrazione dei socket.
 - ARPANET diventa un sottoinsieme di Internet.
 - La National Science Foundation (NSF) realizza una rete di supercomputer (NSFNET) che agisce come backbone di Internet.
 - Nel 1986, si stima che Internet connettesse circa 20 000 computer.
4. Negli anni '90, Internet “esplode”, iniziando a crescere con ritmi velocissimi (sia nelle dimensioni che nella quantità di traffico).

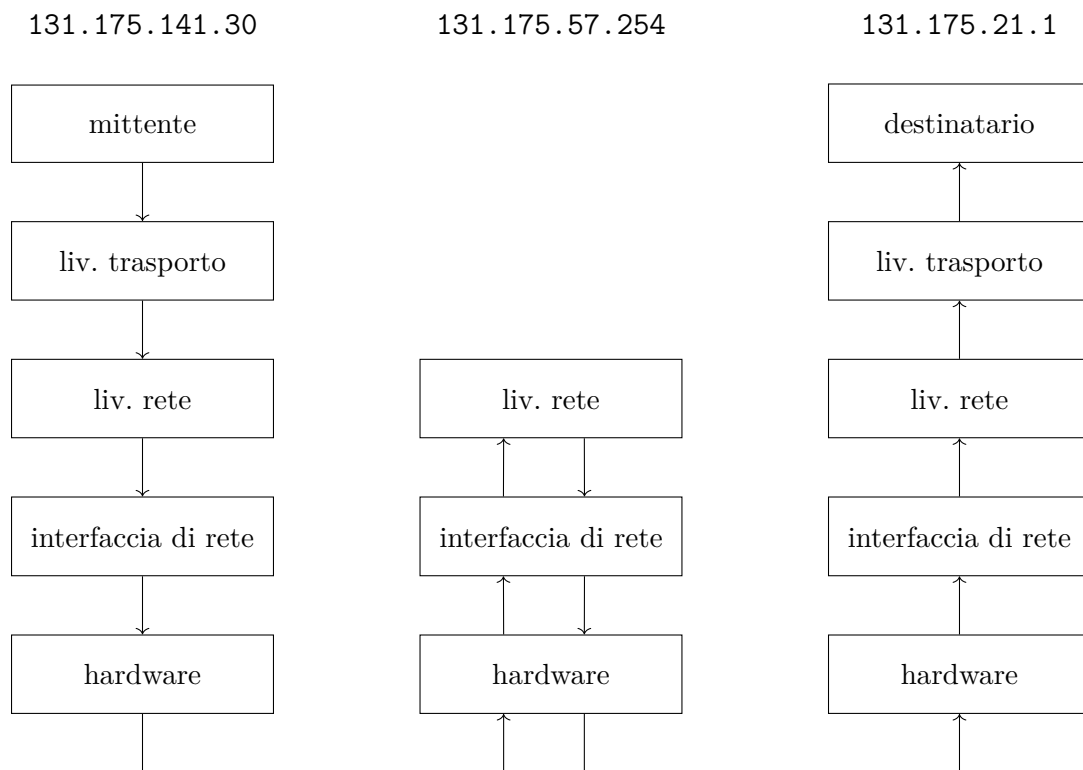
9 Indirizzi IP

Ogni host collegato a una internet (basata su TCP/IP) è identificato da un **indirizzo IP** univoco (non esistono due macchine di una stessa internet che abbiano indirizzo IP uguale). Esso è composto da 32 bit (4 byte), suddivisi² tra *netid* (che identifica la rete) e *hostid* (che individua un singolo host all'interno della rete). Tutte le macchine di una rete hanno lo stesso netid.

Solitamente, un indirizzo IP viene scritto indicando i valori decimali di ciascuno dei 4 byte, separati da punti: ad esempio, 131.175.5.25.

10 Esempio di comunicazione tra due host

Si suppone che l'host identificato dall'indirizzo IP 131.175.141.30 debba comunicare con l'host 131.175.21.1.



²La ripartizione dei 32 bit tra netid e hostid è variabile. Ad esempio, nelle reti locali di piccole-medie dimensioni, si usano spesso 24 bit (i primi 3 byte) per il netid, e i restanti 8 bit per l'hostid.

1. Il mittente prepara il proprio messaggio, e lo affida al protocollo del livello di trasporto. Questo utilizza a sua volta il protocollo del livello di rete, e così via, fino al livello fisico, che trasmette concretamente il messaggio.
2. Tipicamente, la comunicazione non avviene in modo diretto, perché i due host non sono nella stessa rete e/o il mittente non sa come raggiungere il destinatario. Invece, il messaggio viene mandato a una macchina intermedia (nell'esempio, 131.175.57.254), che lo interpreta al livello di rete per leggere l'indirizzo di destinazione, e in base a tale indirizzo determina come inoltrare il messaggio verso il destinatario. Qui si suppone che il destinatario sia connesso direttamente con questa macchina (altrimenti, dovrebbero esserci altri passaggi intermedi).
3. Il messaggio raggiunge il destinatario, che lo interpreta attraverso i protocolli in ordine inverso, "risalendo" la pila a partire dal livello fisico.