

Reti di calcolatori

Luca Tagliavini

25 febbraio 2022

Capitolo 1

Introduzione

1.1 Definizione: Rete di calcolatori

Una *rete di calcolatori* è un insieme di dispositivi di calcolo interconnessi con la particolarità che essi siano **autonomi**.

1.1.1 Classificazione delle Reti

Le reti possono essere suddivise in categorie in base alla dimensione fisica:

1. Reti personali (PAN ovvero *Personal Area Network*), ristrette ad una scrivania o una stanza o perfino locali ad un dispositivo.
2. Reti locali (LAN ovvero *Local Area Network*), come il wifi di casa o la rete di un piccolo ufficio.
3. Reti metropolitane (MAN ovvero *Metropolitan Area Network*) sono spesso unione di più reti locali ad esempio in quartieri o la stessa ALMAWIFI.
4. Reti geografiche (WAN ovvero *Wide Area Network*) che realizzano piani geografici su area nazionale, geografica o internazionale.
5. Infine si ha l'**Internet** che è l'unica rete che connette tutte le reti con complessi protocolli e regole di comunicazione.

1.1.2 Le basi

Nozioni di base sulle reti sono la *banda*, il *ritardo/latency* e il relativo *jitter*.

I nodi di una rete possono essere collegati in vario modo:

- Collegandoli in modo ciclico, o in modo che ci sia un collegamento ad ogni nodi ma si usino il minimo numero di connessioni (minimamente connessa).

- Ogni nodo può essere collegato ad ogni altro dando la massima banda e di conseguenza le migliori performance, a discapito dell'altissimo costo.

Una importante differenza da notare è quella tra la struttura e la topologia. Entrambe rappresentano le connessioni tra i nodi, tuttavia l'ordinamento topologico descrive le connessioni in modo riproducibile e regolato.

1.1.3 Topologia ad anello

La topologia ad anello è una configurazione e connessione dei nodi che connette ogni nodo a due altri nodi, chiamati precedente e successivo fino a formare una struttura ad anello chiuso. In tal modo la rete può supportare anche la perdita di un nodo (se il protocollo è propriamente implementato) senza perdere funzionalità, degradando tuttavia le performance.

1.1.4 Topologia a stella

La topologia a stella è di gran lunga la più usata al giorno d'oggi in tutte le principali reti locali come Router e Switch. La vulnerabilità principale di questa rete è il fallimento dello switch, che causerebbe la disconnessione di tutti i dispositivi: questa vulnerabilità è conosciuta come *single point of failure*.

1.1.5 Topologia a bus

Questa topologia ricorda quella dei calcolatori moderni proposta da Von Neumann, ed era originariamente ampiamente usata quando i calcolatori usavano ancora un clock condiviso per comunicare tra di loro. Attualmente è caduta per la maggior parte in disuso.

1.1.6 Topologia ad albero

Quest'ultima topologia ha una struttura che ricorda quella dell'albero informatico e offre un collegamento simile a quello a stella. Tuttavia il maggior carico di lavoro viene svolto dalla radice. La peculiarità di questa topologia (che è anche minima), è di non *presentare mai cicli*.

1.1.7 Dispositivo o scheda di rete

Lo scopo della scheda di rete è quello di trasmettere/ricevere i dati e svolgere il buffering, ossia mantenimento di dati e codifica/decodifica dei messaggi. Le regole di operazione sono stabilite dagli standard che sono seguiti dai produttori; a loro volta gli standard vengono definiti dai protocolli usati per la comunicazione. At GitHub Satellite India 2021, we announced the GitHub India Grants for Open Source to support the open source community by providing monetary grants to contributors/maintainers, with a total monetary fund of INR 1 crore.

Le schede di rete in alcune configurazioni topologiche (come quella ad anello o a bus) ricevono *tutti* i messaggi, anche quelli non destinati al calcolatore di cui

fanno parte. Per identificare il mittente/ricevente si assegna un indirizzo unico ad ogni scheda di rete chiamato **MAC** (medium access control). L'indirizzo MAC è mostrato in codifica esadecimale con coppie di "numeri" 0 – F separate da due punti. L'indirizzo MAC di broadcast (ovvero un indirizzo che invia un dato pacchetto a *tutti* i nick connessi alla rete) è quell'indirizzo con tutti i bit impostati a 1, e viene dunque mostrato come **FF:FF:FF:FF**.

La definizione del protocollo MAC varia in base al medium con cui si trasmette il segnale. Ad esempio una scheda ethernet ha un protocollo molto più semplice rispetto ad una per la wireless LAN.

1.1.8 Reti di comunicazione

Un canale logico è un'astrazione dei collegamenti fisici, dove un collegamento tra due o più macchine può essere suddiviso in più canali (logici). Esistono due tipi di canali logici: *punto-a-punto* o *broadcast*. Nei canali punto a punto (estremamente costosi) la coordinazione di chi parla è semplicissima, viene solo scelto un tempo nel quale parlare e ci si alterna. Nei canali punto-a-punto gli identificatori MAC sono pressochè superflui. Nei canali broadcast, che sono i più comuni, tutti i messaggi devono passare. Ci si deve dunque coordinare tramite protocolli per quando inviare su determinati canali logici, e si devono usare indirizzi MAC più complessi per tenere traccia del destinatario e mittente.

1.1.9 Reti a commutazione di circuito

Messaggi scambiati tra macchine molto distanti possono saltare attraverso altri calcolatori per raggiungere il destinatario. Un modo di fare ciò è tramite una rete punto-a-punto riservata che collega tutte le macchine l'una all'altra da A a Z in modo che l'unico compito delle macchine B,C,...,Y sia ritrasmettere il segnale sulla rete. In questo modo la latenza è molto bassa, tuttavia si ha un grande spreco per mantenere la connessione (e iniziarla, che richiede tempo) anche se non si inviano dati. Infatti quando la macchina A richiede di parlare con la Z si deve costruire una serie di collegamenti virtuali (istruire le macchine intermedie a reindirizzare i messaggi) e mantenere queste regole ha un costo alto in quanto leva banda agli altri utenti della rete.

1.1.10 Reti a commutazione di pacchetto

Questo approccio è ampiamente usato sulle reti a broadcast. I dati vengono suddivisi in *pacchetti* indipendenti che vengono inviati separatamente. Ogni pacchetto deve contenere l'indirizzo MAC del destinatario (dunque i dati aumentano in dimensione) e attraversando una serie di canali broadcast la comunicazione richiederà sicuramente più tempo. I pacchetti vengono suddivisi in flussi separati (virtuali), per poter consentire una comunicazione in contemporanea di più calcolatori con altri. Nella pratica si usa sempre un timing in modo che una sola macchina occupi il mezzo fisico ad un dato istante, e i calcolatori che vogliono comunicare si coordinano tra di loro.

Il grande vantaggio è che un canale può potenzialmente essere saturato ed il costo non dipende dalla lunghezza della connessione ma dall'utilizzo della banda. Lo svantaggio è la latenza che si crea per le computazioni necessarie allo smistamento.

1.1.11 Servizi orientati alla connessione e non

I pacchetti quando passano per un nodo intermedio seguono la procedura *store and forward*, secondo la quale essi vengono memorizzati sulla macchina intermedia e poi distribuiti ai destinatari necessari. In base alla loro priorità essi possono essere inviati secondo due politiche:

1. Servizi **orientati alla connessione** (connection-oriented): garantiscono l'inizio *ordinato* e l'eventuale reinvio di pacchetti persi.
2. Servizi **non orientati alla connessione** (connectionless): i pacchetti possono seguire strade diverse, arrivare nell'ordine non originale o non arrivare mai.

I servizi che **non perdono e non disordinano i pacchetti** sono detti *servizi orientati alla connessione*.

Per prevenire il disordinamento dei pacchetti alcuni protocolli richiedono che ogni pacchetto contenga un numero d'ordine in modo da poterli ordinare in base al MAC del mittente e al numero d'ordine. Nella pratica in moltissimi protocolli il destinatario, quando riceve un pacchetto, invia al mittente un segnale **ACK** di *acknowledgement*. In caso questo segnale non sia stato ricevuto entro il *timeout* il pacchetto viene rinviato. Il timeout viene scelto in funzione della latenza della rete.

Questi problemi non hanno una soluzione matematica in quanto dipendono fortemente da variabili di ambiente che sono incontrollabili. Di conseguenza la trasmissione delle informazioni tramite internet viene fatta "facendo il proprio meglio", o in gergo *best-effort*.

1.1.12 Protocolli

I protocolli sono regole (semantica) di gestione della comunicazione e formati (sintassi) in cui scambiare l'informazione in modo non ambiguo. Questi protocolli consentono la comunicazione tra dispositivi ampiamente diversi i quali parlano tutti lo stessi "linguaggio" (protocollo).

Il lavoro è suddiviso tra vari *livelli di protocolli* dove ogni livello protocollo gestisce uno specifico problema. I protocolli di livello superiore possono usufruire delle strutture offerte dai protocolli inferiori e si ha dunque una comunicazione in entrambi i sensi tra tutti i livelli. I protocolli comunicano tra di loro tramite interfacce, che sono plasmate dalle richieste e dai servizi scambiati.

1.1.13 Artichettura standard di rete

Lo standard *ISO/OSI RM* definisce una serie di 7 livelli, che *devono* essere implementanti da tutti i dispositivi che vogliono comunicare su internet tra di loro. Ogni livello gestisce una classe di problemi e quello superiore riceve un ambiente semplificato in cui non si deve più curare dei problemi dei livelli sottostanti. Il dialogo tra i livelli avviene tramite una interfaccia comune. I protocolli dello stesso livello comunicano tra di loro nella pratica passando tramite livelli sottostanti, ma concettualmente e durante lo sviluppo in modo astratto e diretto tra di loro.

Ecco la nomenclatura dei 7 livelli dal più basso (1) al più alto (7):

1. Livello Applicazione
2. Livello Presentazione
3. Livello Trasporto/Sessione
4. Livello Rete
5. Livello MAC/LLC (collegamento)
6. Livello fisico

Il livello fisico si occupa di codificare i bit in segnali analogici. Il livello MAC/LLC si occupa di inviare ai giusti dispositivi nelle reti locali i messaggi e "mandare avanti" pacchetti non destinati ai dispositivi che li ricevono. Al livello successivo (quello di rete) si collocano i protocolli IP (Internet Protocol) che consentono di smistare i pacchetti a livello globale assegnando a ogni dispositivo un *indirizzo logico* che rende lo smistamento veloce con grandi reti. Ancora sopra si ha il livello di trasporto, che si occupa della comunicazione *end to end* e garantisce l'invio e la ricezione (occupandosi dell'eventuale rispedizione) dei pacchetti attraverso reti. Si noti che i protocolli sottostanti ritentano l'invio in caso di errore (un fissato numero di volte) ma non garantiscono l'effettiva ricezione dei dati e non comunicano con il mittente originale per richiedere il rinvio in caso di errori in mezzo alla catena. Al livello sessione (disuso) si hanno dei protocolli che si occupano di mantenere traccia dei messaggi precedentemente scambiati durante l'attuale sessione di comunicazione. Il livello di presentazione (disuso) si occupa della formattazione dei dati, eventuale traduzione, ad esempio tra dati big endian e little endian. Infine il livello applicazione è quello più importante, dove vivono i protocolli più ad alto livello come FTP, HTTP, etc. Questi protocolli sono usati dalle applicazioni vere e proprie, che non vanno confuse con i protocolli al livello applicazione.

1.1.14 Incapsulamento

I dati provenienti dai protocolli ad alto livello (livello 7 solitamente) attraversano tutti i protocolli sottostanti prima di essere trasmessi. I protocolli sottostanti incapsulano l'informazione assieme ad altri dati necessari per le loro controparti

alla ricezione o per arricchire l'informazione con bit di controllo o i dati del destinatario per esempio.

Una volta che il pacchetto sarà ricevuto il pacchetto svolgerà il percorso di decodifica passando in ordine inverso attraverso tutti i protocolli, e dunque verrà decapsulato e i metadati saranno utilizzati per vari controlli e smistamenti.

1.1.15 Costruzione di segmenti di reti locali

Le reti locali sono composte dai seguenti componenti, che ne consentono il funzionamento sia al livello fisico che al livello MAC/LLC:

- Un **Repeater** agisce al livello 1 (fisico) e ha il semplice compito di ripropagare i segnali elettrici che riceve, senza realizzare qualunque computazione sui dati. I segnali vengono ricevuti da un cavo e propagati in un altro in uscita.
- Una **Hub** è un *repeater multiporta* che riceve un segnale da un singolo cavo e lo propaga su una serie di cavi in uscita. È la parte centrale delle connessioni a stella.
- Un **Bridge** (oggi in disuso) è un ponte di traduzione che adatta i livelli MAC differenti degli utenti sulla rete.
- Un **Switch** agisce al livello 2 e funge sia come un bridge (traduzione) e *smistatore* (indirizzamento sul giusto canale) su una rete di molti utenti. Molti switch sono *buffered* ovvero hanno una capacità di memoria in cui possono tenere i pacchetti in coda per essere inviati nel caso di unione di reti a velocità diversa.

1.1.16 Reti di reti

Quando le reti diventano grandi vengono create sottoreti collegate all'internet tramite router, e vengono poi collegati questi router tra di loro ($n_{router} \lll n_{utenti}$). A questo punto i client non comunicano più tramite gli indirizzi MAC (ci vorrebbe una tabella immensa per lo smistamento) ma bensì tramite gli indirizzi IP (Internet Protocol), che sono molto più concisi e facili da smistare.

Lo scopo dei router è quello di smistare nel modo più efficiente possibile pacchetti tra reti diverse e capire se essi devono uscire o entrare nella rete locale. Queste informazioni vengono ottenute guardando l'indirizzo IP (Internet Protocol).

Internet Protocol

Il protocollo che garantisce lo smistamento dei pacchetti a livello globale viene gestito tramite il protocollo IP. La struttura della rete prende la forma ad albero nel caso ideale (ma spesso anche in quello pratico) quindi le tabelle di indirizzamento sono estremamente semplici e lo scambio di informazioni altamente veloce.

L'indirizzamento IP usato per internet viene assegnato all'interno di una rete in modo dinamico o statico. Nella modalità statica l'indirizzo viene deciso dal client e rimane fisso, mentre in quella dinamica l'ip viene assegnato dal router. I router tendono ad assegnare sempre lo stesso IP allo stesso MAC (quando si usa la configurazione dinamica).

1.1.17 Struttura IPv4

Un indirizzo IPv4 è composto da 32 bit ovvero 4 byte. Viene spesso formattato come i quattro byte separati da punti. Se nel primo byte partiamo con il primo bit a 0, sappiamo che siamo in una rete di classe A. Nelle reti di classe A si usano i restanti tre byte per indirizzamenti successivi, dando vita a 2^{24} combinazioni possibili. Le reti di classe A saranno dunque $2^{8-1} = 126$.

Se i primi due bit sono impostati a 10 siamo invece in una rete a classe B che usa i primi due byte per l'indirizzamento tra reti di classe B, e i restanti 2^{16} per indirizzamenti successivi. Esistono dunque $(2^{16-2} = 16384)$ reti di classe B.

Se i primi tre bit sono a 110 siamo in una rete di classe C, dove i primi tre byte vengono usati per l'indirizzamento tra simili e i restanti 2^8 bit per indirizzamenti interni. Ci sono dunque $2^{24-3} = 16777216$ reti di classe C con $2^8 = 256$ host ad esse interni.