1 Lezione del 24-09-25

1.1 Comunicazione dati su Internet

Abbiamo visto la struttura a livello fisico della rete Internet. Vediamo adesso i meccanismi secondo cui la trasmissione di dati avviene. La rete Internet è una rete a **commutazione di pacchetto**, il compito degli host è di:

- Ottenere messaggi dalle applicazioni;
- Dividere quei messaggi in frammenti più piccoli, detti **pacchetti**, di dimensione L bit;
- Trasmettere quei pacchetti nella rete di accesso ad una *frequenza di trasmissione* (o **bit-rate**) R.

Il tempo T_{packet} necessario a trasmettere un pacchetto da L bit su una linea da R bit al secondo di bitrate sarà quindi semplicemente calcolato come:

$$T_{\text{packet}} = \frac{L}{R}$$

Il bitrate, detto anche frequenza di *link*, dipende appunto dal **link** (o *mezzo*) della trasmissione, cioè l'infrastruttura fisica che sta fra trasmettitore e ricevitore. Possiamo classificare 2 tipi di link:

- Mezzi **guidati**: segnali che si propagano in mezzi solidi (rame, fibra, cavi coassiali, ecc...).
  - Un celebre esempio di mezzo trasmissivo guidato è il classico **doppino telefonico**, o trasmettitore a *Twisted Pair* (**TP**). Questo è formato da due fili di rame isolati e avvolti l'uno sull'altro, che permettono la trasmissione differenziale e quindi la riduzione dei rumori in *common-mode*;
  - Un altro tipo di mezzo trasmissivo guidato è il **cavo coassiale**, formato da due conduttori concentrici in rame separati da un dielettrico. Il segnale è trasferito come campo magnetico fra i due conduttori: questo permette bitrate più alti rispetto al normale doppino telefonico e una migliore schermatura dalle interferenze;
  - Infine, possiamo parlare della **fibra ottica**, formata da fibra di vetro che porta impulsi luminosi ad altissima velocità. Questa tecnologia presenta velocità di trasmissione estremamente alte, e vista la natura luminosa del segnale, non è suscettibile ad interferenze (alta affidabilità, cioè piccola frequenza di errore sui bit).
- Mezzi **non guidati**: segnali che si propagano nell'etere (segnali radio, ecc...). Questi sono solitamente meno sicuri ma significativamente più comodi per l'utente finale (possibilità di spostarsi, mancanza di cavi, ecc...).

  La trasmissione wireless è suscettibile a fenomeni fisici come *riflessi*, *interferenze* e *ostruzione* da parte di oggetti fisici.
  - Il **WiFi** è un esempio di mezzo di trasmissione non guidato che può raggiungere centinaia di Mbps su regioni locali;

  - Reti wireless più ampie possono essere quelle **cellulari**, usate nella telefonia mobile;
  - Infine si può parlare delle **reti satellitari**, usate per l'interconnessione di regioni geografiche fra di loro anche molto distanti.

1.1.1 Commutazione di circuito

Prima della commutazione di pacchetto si usava la tecnica della **commutazione di circuito** (ad esempio sulle linee telefoniche). Questo prevede di dedicare completamente una certa linea di trasmissione alla comunicazione fra due host, invalidandone quindi l'uso da parte di altri host.

Chiaramente, la commutazione di pacchetto permette un carico migliore della linea, dove più pacchetti provenienti da diverse fonti possono viaggiare a istanti temporali molto vicini fra di loro.

In particolare, la commutazione di paccheto è utile per dati trasmessi in *burst*, mentre la rete a commutazione di circuito assicura minima congestione possibile a costo di occupazione completa della linea.

1.1.2 Commutazione di pacchetto

La tecnica della **commutazione di pacchetto** o *packet-switching* permette ad una rete di **router** interconnessi di ricevere ed instradare (letteralmente, *"routing"*) pacchetti provenienti da più fonti in modo che raggiungano la loro destinazione.

Questo inserisce chiaramente un ritardo nel sistema, in quanto il router deve:

- Ricevere il pacchetto *completamente* ed memorizzarlo: questo richiede L/R secondi;
- Leggere l'header del pacchetto per capire il prossimo passo dell'instradamento;
- Trasmettere il pacchetto verso la sua nuova destinazione (un altro router o l'host finale), impeigando ancora L/R secondi.

Abbiamo quindi che il ritardo end-end immesso dal router è necessariamente di almeno $2L/R$ secondi, tralasciando il tempo necessario all'instradamento stesso.

Nel caso generale si abbiano N router (quindi $N+1$ link fra i router) e P pacchetti da inviare, dovremo considerareche il primo pacchetto arriva in $(N+1)\frac{L}{R}$ (deve attraversare tutti i link) e i successivi $P-1$ pacchetti arrivano in $(P-1)\frac{L}{R}$, per cui il tempo complessivo è:

$$T_{end-to-end} = (N+P)\frac{L}{R}$$

Se troppi pacchetti arrivano in un breve lasso di tempo, cioè se la frequenza di arrivo supera quella di trasmissione:

- I pacchetti verrano messi in coda finché non sarà possibile trasmetterli;
- I pacchetti possono essere persi se il buffer di memoria dedicato alla loro memorizzazione nel ruoter si riempie.

1.2 Prestazioni della commutazione di pacchetto

Facciamo qualche considerazione ulteriore sulle prestazioni delle linee a commutazione di pacchetto. Abbiamo ottenuto il valore $T_{\text{packet}} = \frac{L}{R}$ per la trasmissione di un singolo pacchetto da L bit su una linea con bitrate R, e $T_{\text{end-to-end}} = (N + P)\frac{L}{R}$ per più pacchetti su un numero arbitrario di router.

Da quanto abbiamo detto nella scorsa sezione, un modello più sofisticato del packet-switching terrà conto di 4 sorgenti di ritardo:

- Ritardo di **trasmissione** T_{trans}, dato dalle caratteristiche del link. Come abbiamo già detto, questo vale:
$$T_{\text{trans}} = \frac{L}{R}$$
con L lunghezza del pacchetto in bit e R bitrate del link;
- Ritardo di **propagazione** T_{prop}, dato dalle proprietà fisiche del mezzo di trasmissione. In particolare, questo è il tempo fisico di trasmissione del segnale su un link, dato da:
$$T_{\text{prop}} = \frac{d}{s}$$
con d distanza del link e s velocità del mezzo di trasmissione. Chiaramente, per i nostri scopi s sara una frazione significativa della velocità della luce $c \approx 3 \cdot 10^8$;
- Ritardo di **laborazione** (instradamento) T_{proc}, dipende dalle caratteristiche del router ed è perlopiù costante;
- Ritardo di **accodamento** dato dalla presenza di code T_{queue}. Questo è il più complicato da trattare, in quanto dipende dal numero di pacchetti presenti nel buffer del router. Come vedremo fra poco, una buona euristica per la valutazione di questo ritardo (che è comunque trattabile solo in maniera statistica) è l'*intensità di traffico* sulla linea di trasmissione.

Sommando queste sorgenti di ritardo potremo ottenere una stima del ritardo complessivo su un router (nodo) T_{node}:

$$T_{\text{node}} = T_{\text{trans}} + T_{\text{prop}} + T_{\text{proc}} + T_{\text{queue}}$$

Dati N nodi, il ritardo end-to-end potrà quindi essere calcolato semplicemente come:

$$T_{\text{end-to-end}} = NT_{\text{node}}$$

1.2.1 Ritardo di accodamento

Come anticipatp, possiamo valutare statisticamente il ritardo di accodamento T_{queue} calcolando l'*intensità di traffico* su una linea:

$$I_{\text{traffic}} = \frac{L\alpha}{R}$$

presa α come la frequenza media di trasmissione di pacchetti sull'istante temporale.

Semplicemente tracciando la funzione si nota che se $I_{\text{traffic}} >= 1$ il tempo di di accodamento tende a infinito (si ha necessariamente perdita di pacchetti), per cui vorremo mantenere $I_{\text{traffic}} < 1$, e idealmente $I_{\text{traffic}} << 1$.

1.2.2 Traceroute

Per fare diagnostica in situazioni reali, si possono usare software di **traceroute**, che inviano pacchetti con lo scopo di tracciare gli host incontrati e misurare il tempo impiegato nella trasmissione *round-trip*, o almeno capire se i pacchetti sono stati inviati o persi.

1.2.3 Troughput

Introduciamo il concetto di **troughput** come la frequenza (bit al secondo) con cui i bit vengono spediti da trasmettitore a ricevitore. Il throughput può essere:

- **Istantaneo**: calcolato ad un certo istante temporale (per quanto possibile dalla natura discreta della trasmissione);
- **Medio**: calcolato su un periodo temporale più lungo.

Ipotizziamo che un server debba inviare un file di F bit ad un client. La linea di comunicazione fra server e router ha capacità di link di R_S bit/secondo, mentre la linea fra router e client ha capacità di R_C bit/secondo. Chiaramente la capacità totale sarà:

$$R_F = \min(R_S, R_C)$$

cioè link con basse capacità di trasmissione rappresentano *bottleneck* per il throughput di tutto il collegamento fra dispositivi.

Facciamo adesso l'esempio di una rete condivisa fra più coppie client/server. In questo caso, date N coppie, la capacità della rete di interconnessione agli occhi di una singola coppia sarà, preso R come la capacità complessiva:

$$R_i = \frac{R}{N}$$

per cui ogni coppia vedrà una linea di comunicazione con capacità:

$$R_F = \min(R_S, R_C, \frac{R}{N})$$

1.3 Struttura di Internet

Abbiamo detto che Internet è una rete di reti.

- Gli host si collegano a internet attraverso le *reti di accesso* (reti LAN, istituzionali, reti mobili, ecc...);
- Le reti di accesso si collegano agli **ISP** (*Internet Server Provider*);
- Gli ISP devono essere collegati ad altri ISP per permettere la comunicazione fra host su diversi ISP (all'interno di **IXP** (*Internet eXchange Point*, ecc...)).

La struttura di reti interconnesse che si viene a formare è molto complessa, e la sua evoluzione è stata guidata da fattori economici e politici.

Abbiamo quindi una struttura gerarchica:

1. Al livello più alto troviamo i cosiddetti **tier 1 ISP** e i **content provider** (Google, Facebook, ecc...) che si occupano di copertura nazionale e internazionale. In particolare, i content provider preferiscono collegarsi direttamente agli IXP per risparmiare sugli ISP;
2. Seguono gli **IXP**, che collegano più ISP fra di loro, e gli ISP locali (regionali, ecc...);
3. Infine troviamo le reti di **accesso** locale (reti LAN, WLAN, ecc...).

1.4 Sicurezza Internet

La struttura di Internet espone i suoi utenti a diversi pericoli, fra cui:

- **Virus**: programmi maligni che si replicano modificando altri programmi;
- **Worm**: programmi maligni che si replicano con lo scopo di diffondersi in altri computer;
- **Spyware**: programmi maligni che cercano di ottenere informazioni che violano la privacy;

ed altre svariate categorie di *malware*.

1.4.1 Attacchi DoS

Un caso tipico di attacchi informatici in rete è quello degli attacchi **DoS** (*Denial of Service*), dove l'attaccante cerca di sfruttare le risorse di un server (memoria, larghezza di banda) al punto di renderlo inutilizzabile ad altri utente. Questo può essere fatto inviando traffico maligno al server (pacchetti molto grandi o corrotti).

La tecnica può essere espansa a più attaccanti (*botnet*) dando vita a tecniche più sofisticate (come il **DDoS**, *Distributed Denial of Service*).

1.4.2 Intercezione di pacchetti

Una problematica delle reti a commutazione di pacchetto è l'**intercezione** dei pacchetti instradati, o in inglese *packet sniffing*. Questo è più facile su reti wireless piuttosto che cablate, ed espone pericoli ovvi per la sicurezza (password, dati sensibili, ecc...).

Oggi, come vedremo, si usano protocolli che implementano forme di crittografia per mitigare questo tipo di problematiche.

1.4.3 Spoofing IP

Il problema di identità fasulle si presenta in Internet attraverso pacchetti malformati, con indirizzi sorgente sbagliati. Queste tecniche, seppure meno pericolose, possono comunque essere usate per confondere o comunque complicare il traffico sulle reti.