ALOHA e CSMA/CD

L'ALOHA è una tecnica di accesso a canale condiviso. Il principio di funzionamento è estremamente semplice: quando una stazione presente sul canale deve trasmettere una frame lo fa senza attendere alcuna autorizzazione. Se la trasmissione ha successo, cioè la frame non *collide* con trasmissioni di altre stazioni, bene; altrimenti la stazione rileva la collisione e provvede, in un secondo momento, a ritentare la trasmissione della frame, fino a quando non ha successo. Da questa breve descrizione è già possibile intuire che si tratta di una tecnica assai poco efficiente nell'uso del canale ed in generale poco affidabile. Ciononostante, è interessante studiarla per diverse ragioni:

- anzitutto, l'ALOHA, messa a punto negli anni '70 nell'Università delle Hawaii, è stata la prima tecnica dinamica per l'accesso a canale comune, e ha dunque interesse storico;
- sebbene l'accesso a canale mediante contesa sia poco efficiente rispetto all'accesso mediante prenotazione, esso risulta necessario proprio nelle tecniche con prenotazione ogni volta che il meccanismo di passaggio del token si inceppa (ad esempio nelle fasi di inizializzazione dell'anello);
- introducendo alcuni semplici correttivi all'ALOHA si ottiene una tecnica di accesso semplice ed efficiente (la CSMA/CD) che costituisce il cuore dello standard IEEE 802.3 e del prodotto commerciale Ethernet, la rete locale di gran lunga più diffusa nell'office automation.

Modello

Per condurre una trattazione analitica dell'ALOHA, e quindi determinarne le prestazioni, è necessario specificare chiaramente il modello del canale, il funzionamento delle stazioni, il traffico. Cominciamo qui a fissare alcune caratteristiche generali.

1. <u>Canale comune</u>. Esiste un singolo canale ad accesso comune, al quale tutte le stazione della rete sono collegate; le stazioni sono sempre in ascolto, anche se l'entità dello strato MAC di una certa stazione passa allo strato di rete solo le frame indirizzate alla stazione stessa.
2. <u>Feedback immediato</u> (0,1,e). Ogni stazione, ascoltando il canale, è in grado di determinarne lo stato corrente e scoprire quindi se il canale è inattivo (0) impegnato da una singola trasmissione (1) o impegnato da più trasmissioni, nel qual caso si ha una collisione o errore (e); questa ipotesi è generalmente soddisfatta eccetto che nelle reti satellitari, nelle quali le stazioni conoscono lo stato del canale solo dopo un significativo tempo di propagazione.
3. <u>Numero stazioni e buffering</u>. L'ipotesi corretta è che ci sia un numero finito m di stazioni e che ognuna abbia un buffer in cui conservare un certo numero di frame (pacchetti) passate dall'entità dello strato di rete in attesa di trasmetterle sul canale. Questa ipotesi

però non permette una trattazione analitica semplice, perchè si tratterebbe di analizzare il comportamento congiunto di m code (una per ogni stazione sul canale). Useremo invece le seguenti ipotesi in alternativa l'una all'altra;

- infinite stazioni: per ogni frame in arrivo dallo strato di rete c'è una stazione dedicata in esclusiva alla sua trasmissione, e quindi ci sono potenzialmente infinite stazioni presenti sulla rete; questa è un'ipotesi pessimistica poichè tutte le frame in arrivo competono subito per il canale anche se alcune frame realisticamente andrebbero in una coda ordinata;
- m stazioni senza buffering: in questo caso eliminiamo l'accodamento all'interno delle stazioni semplicemente rigettando ogni nuova frame in arrivo a stazioni già impegnate nella trasmissione di una frame; questa ipotesi, al contrario della precedente, è ottimistica perchè parte del traffico viene bloccato in ingresso e quindi non grava sul canale.

4. Arrivi Poisson. Si fa l'ipotesi che il traffico in arrivo dallo strato di rete segua statistiche di Poisson e che, inoltre, il traffico aggregato composto dagli arrivi nuovi dallo strato di rete più le ritrasmissioni (interne al MAC) sia complessivamente Poisson. Nel caso di m stazioni questa è necessariamente un'approssimazione. Inoltre nelle reti reali, non sempre questa ipotesi è ragionevole.

5. Tempo continuo o slotted. Per semplicità si fa l'ipotesi che tutte le frame abbiano la stessa durata T_f: la rete è dunque in grado di servire al più $\mu = 1/T_f$ frame al secondo. Considerando un tasso di arrivi di λ ed un throughput di γ frame al secondo rispettivamente, e normalizzando rispetto a μ si avrà un'intensità di traffico offerto di $\rho = \lambda/\mu$ ed un'intensità di traffico servito di $\epsilon = \gamma/\mu$. Per semplicità, e senza perdita di generalità, porremo $T_f = 1$ (e quindi $\mu = 1$) e useremo λ e γ con il duplice significato di tasso e intensità di traffico. Nell'ALOHA puro si farà riferimento ad un asse dei tempi continuo, mentre nell'ALOHA slotted il tempo sarà discretizzato in intervalli di durata T_f.

ALOHA puro

Nell'ALOHA puro ogni stazione trasmette esattamente quando vuole, senza alcun vincolo: se, a causa di collisioni, la trasmissione non va a buon fine, verrà ritentata in un secondo momento. Per inquadrare più precisamente la situazione facciamo riferimento allo scambio di pacchetti e frame fra il sottostrato MAC e quelli contigui, cioè lo strato fisico (PHY) verso il basso e lo strato di rete (NET) verso l'alto (per semplicità consideriamo trasparente il sottostrato LLC). Lo strato di rete passa dei pacchetti al MAC chiedendone la trasmissione; il MAC poi li ingloba in frame (aggiungendo campi di indirizzamento, checksum, ecc.) che passa a sua volta allo strato fisico per la trasmissione sul canale. Va sottolineato che il traffico fra MAC e PHY è più intenso di quello fra NET e MAC perchè alcune frame dovranno essere ritrasmesse (eventualmente più volte) a causa delle collisioni sul canale. La situazione dunque è quella mostrata in Fig.1: c'è un traffico di γ_{in} pacchetti per unità di tempo (diciamo pacchetti/s) in ingresso al PHY, e ogni pacchetto corrisponde ad una frame da trasmettere; a questo flusso di frame "nuove" si aggiunge un flusso di frame "vecchie", che sono state già trasmesse ma hanno sofferto una collisione. Complessivamente λ frame/s sono passate allo strato fisico, ed altrettante ne riemergono: detta P_C la probabilità che una frame trasmessa collida con altre, λP_C frame/s dovranno essere ritrasmesse, e sono perciò quelle che contribuiscono al traffico in ingresso al MAC, mentre $\lambda(1 - P_C)$ frame/s sono ricevute correttamente e formano il flusso uscente di $\gamma_{out} = \lambda(1 - P_C)$ pacchetti/s verso il NET. In condizioni di equilibrio, naturalmente, $\gamma_{in} = \gamma_{out} = \gamma$.

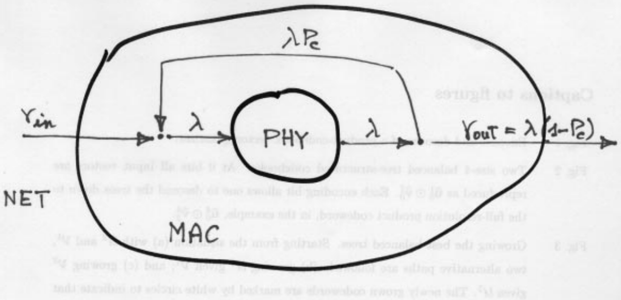


Fig. 1

Per poter studiare questo sistema, bisogna valutare la P_C. Nelle ipotesi di traffico in ingresso Poisson e durata costante delle frame questo è un calcolo elementare. Con riferimento alla Fig.2, sia t_k l'istante di arrivo della k-esima frame sul canale (passaggio a PHY), e sia τ_k il suo tempo di interarrivo ($\tau_k = t_k - t_{k-1}$); sia infine, come già detto, $T_f = 1$ il tempo di trasmissione della generica frame. La k-esima frame non soffrirà una collisione solo se $\tau_k > 1$ (la frame precedente è

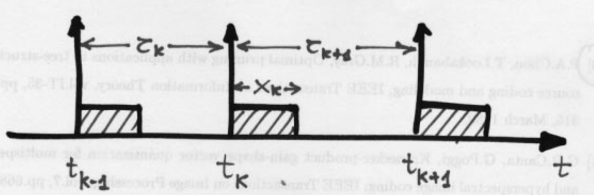


Fig. 2

arrivata almeno un secondo prima) e contemporaneamente $\tau_{k+1} > 1$ (la frame successiva arriverà fra almeno un secondo). Quindi, ricordando che gli interarrivi sono tutti indipendenti ed hanno pdf esponenziale di parametro pari al tasso di arrivi λ, risulta

$$1 - P_C = \Pr\{ \ \tau_k > 1, \tau_{k+1} > 1 \ \} = \Pr\{ \ \tau_k > 1 \ \}\Pr\{ \ \tau_{k+1} > 1 \ \} = e^{-2\lambda}$$

In ipotesi di equilibrio possiamo perciò dire che il throughput del sistema, in frame/s, è di

$$\gamma = \lambda e^{-2\lambda}$$

E' immediato verificare che il throughput massimo si ottiene per $\lambda = 0.5$ e vale $\gamma_{max} = 1/2e \simeq 0.18$. Per comprendere questo risultato, notiamo che l'ideale sarebbe quello di trasmettere una frame/s (di più non è possibile a causa delle collisioni). Tuttavia, il meglio che si riesce a fare con questa semplice tecnica è di trasmettere con successo circa 0.18 frame/s (con un tasso di trasmissioni *tentate* sul canale di 0.5 frame/s) sfruttando il canale solo per il 18% delle sue possibilità, un'efficienza estremamente limitata.

* Caso di T_f variabile

Nell'analisi si è supposto T_f costante e pari ad uno. Ora invece supponiamo che la durata della k-esima frame sia anch'essa una variabile aleatoria esponenziale, $X_k \sim \mathrm{Exp}(\mu)$, con $\mu = 1$ in modo che la durata media continui ad essere unitaria. Con riferimento alla Fig.2bis, possiamo condurre un ragionamento analogo a quello fatto sopra, con la differenza che le X_k adesso sono aleatorie; questo comporta che, anche se la $(k-1)$-esima frame non causa collisione, quelle precedenti possono farlo, se la loro durata è molto grande; in formule:

$$1 - P_C = \Pr\{ \ \tau_{k+1} > X_k, \ \tau_k > X_{k-1}, \ (\tau_k + \tau_{k-1}) > X_{k-2}, \dots \ \}$$

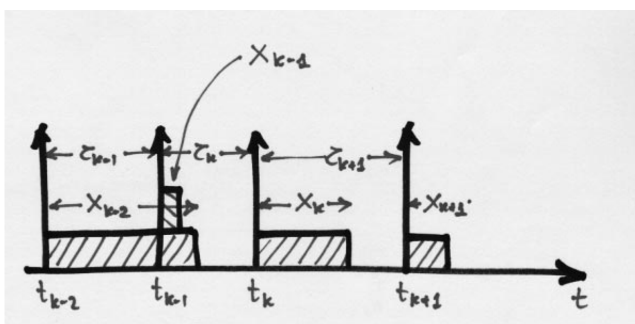


Fig. 2 bis

Come prima possiamo individuare due eventi indipendenti e calcolare il prodotto delle loro probabilità:

$$1 - P_C = \Pr\{ \ \tau_{k+1} > X_k \ \}\Pr\{ \ \tau_k > X_{k-1}, \ \tau_k + \tau_{k-1} > X_{k-2}, \dots \ \}$$

Per la prima abbiamo

$$\Pr\{ \ \tau_{k+1} > X_k \ \} = \int \mu e^{-\mu t} u(t) e^{-\lambda t} dt = \frac{\mu}{\mu + \lambda}$$

Per la seconda, invece, basta rendersi conto che stiamo calcolando la probabilità che nessuna delle frame arrivate in precedenza sia attualmente in trasmissione; dato che gli arrivi (istanti di inizio trasmissione) sono Poisson, i tempi di servizio (durata delle frame) sono esponenziali indipendenti fra loro e dagli arrivi, e tutti i clienti (frame) possono essere serviti (trasmessi) contemporaneamente, possiamo modellare il sistema come una coda $M/M/\infty$. In un generico istante, la probabilità che vi siano n clienti nel sistema sarà

$$p_n = \frac{e^{-\rho}\rho^n}{n!}, \quad \rho = \lambda/\mu$$

e quindi la probabilità cercata (sistema vuoto) semplicemente $e^{-\lambda}$. In definitiva il throughput varrà

$$\gamma = \frac{\lambda}{1+\lambda}e^{-\lambda}$$

che è massimo per $\lambda \simeq 0.618$ ed in tal caso vale circa 0.205.

Analisi della stabilità

In Fig.3 sono riportate alcune curve che permettono di studiare in maggior dettaglio il comportamento di una rete che lavora in ALOHA. In particolare, in funzione di λ vengono riportati:

- il tasso di trasmissioni tentate (pari a λ stesso);
- il tasso di arrivi dal NET, γ_{in}, supposto costante;
- il tasso di trasmissioni riuscite, $\gamma_{out} = \lambda e^{-2\lambda}$.

Per un fissato valore di λ si ricava subito anche il tasso di trasmissioni tentate ma fallite, pari a $\lambda - \gamma_{out}$. Quando λ è molto piccolo, cioè il traffico sul canale è ridotto, quasi tutte le trasmissioni tentate hanno successo, mentre al crescere di λ aumenta la probabilità di collisione fino a che quasi tutte le trasmissioni tentate falliscono ed il throughput si riduce a 0.

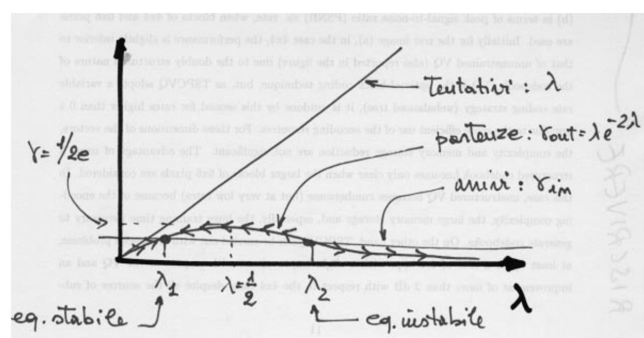


Fig. 3

Se $\gamma_{in} < 1/2e$ ci sono due intersezioni fra le curve dei tassi di arrivi e partenze dal MAC, e cioè due punti di equilibrio. Se $\gamma_{in} = 1/2e$ questi collassano in un unico punto. Bisogna ricordare, tuttavia, che tutte le quantità considerate sono medie statistiche: nella singola realizzazione si avranno delle funzioni del tempo e si osserveranno variazioni temporali dovute all'aleatorietà delle grandezze in gioco. Ad esempio, si può avere una raffica di arrivi per cui $\gamma_{in}(t)$ aumenta, o una sequenza di collisioni per cui $\gamma_{out}(t)$ diminuisce; entrambi i fenomeni influenzeranno $\lambda(t)$, nel caso specifico facendolo aumentare. E' quindi necessario analizzare anche la dinamica del sistema.

Sia $m(t)$ il numero di stazioni attive all'istante t (per ipotesi tale numero può crescere indefinitamente, poichè corrisponde al numero di pacchetti da trasmettere presenti nel MAC), e sia q il tasso di tentate trasmissioni per la singola stazione, pari alla probabilità di trasmettere la singola frame presente, di modo che $\lambda(t) = m(t)q$. Dopo un piccolo intervallo Δt saranno attive

$$m(t + \Delta t) = m(t) + \gamma_{in}\Delta t - \gamma_{out}\Delta t$$

stazioni, cioè quelle precedenti, più le stazioni che si sono attivate a causa di nuovi arrivi, meno quelle che hanno trasmesso con successo la loro frame e si sono quindi disattivate. Passando al limite per Δt che tende a 0 si ottiene subito $m'(t) = \gamma_{in} - \gamma_{out}$ e quindi

$$\lambda'(t) = q[\gamma_{in} - \gamma_{out}]$$

Questa relazione indica, come era ragionevole aspettarsi, che il traffico sul canale $\lambda(t)$ tende ad aumentare se il tasso di nuovi arrivi eccede quello delle partenze, e tende a diminuire nella situazione opposta. Sul grafico di Fig.3, allora, detti λ_1 e λ_2 i tassi per i quali si ha equilibrio, si individuano tre distinte regioni:

- per $\lambda(t) < \lambda_1$ il traffico sul canale è molto scarso e, anche se quasi tutte le trasmissioni tentate hanno successo, i nuovi arrivi sono più frequenti, e $\lambda(t)$ tende ad aumentare;
- per $\lambda_1 < \lambda(t) < \lambda_2$ ci troviamo in condizioni quasi ottimali ed il traffico servito è maggiore di quello in arrivo dal NET, per cui $\lambda(t)$ tende a diminuire;
- quando $\lambda(t) > \lambda_2$ il traffico sul canale è troppo intenso e quasi tutte le trasmissioni tentate soffrono una collisione, il traffico servito è minore di quello entrante per cui $\lambda(t)$ tende a crescere ulteriormente.

Da questa analisi è chiaro che, mentre λ_1 è un punto di equilibrio *stabile*, λ_2 è un punto di equilibrio *instabile*, superato il quale la rete tende ad essere sovraccarica e ad avere un traffico sempre più intenso ed un throughput sempre più piccolo, fino al blocco totale. Di conseguenza, l'efficienza dell'ALOHA puro è, nella pratica, ancora minore del 18%, poichè conviene lavorare con un carico minore di $1/2e$ in modo da distanziare i due punti di equilibrio e minimizzare la probabilità di spostarsi casualmente oltre il secondo, nella zona di instabilità, con il conseguente blocco della rete.

In realtà, questa analisi è alquanto pessimistica poichè prevede che il traffico possa effettivamente aumentare senza limite e, inoltre, che il traffico accettato non vari in funzione dello stato della rete. Più avanti vedremo, per il caso slotted, un'analisi più realistica.

ALOHA slotted

Nell'ALOHA slotted, l'asse dei tempi è suddiviso in intervalli, o "slot", di durata pari al tempo di trasmissione di una frame, e le stazioni non possono iniziare la trasmissione in istanti qualsiasi ma devono aspettare l'inizio di uno slot. Questo semplice espediente permette di aumentare significativamente il throughput massimo rispetto al caso dell'ALOHA puro. Continuiamo a supporre che il traffico generato dallo strato MAC sia Poisson, e quindi con arrivi in istanti generici, solo che adesso la trasmissione non comincia nell'istante dell'arrivo ma a partire dal primo slot successivo, come mostrato in Fig.4.

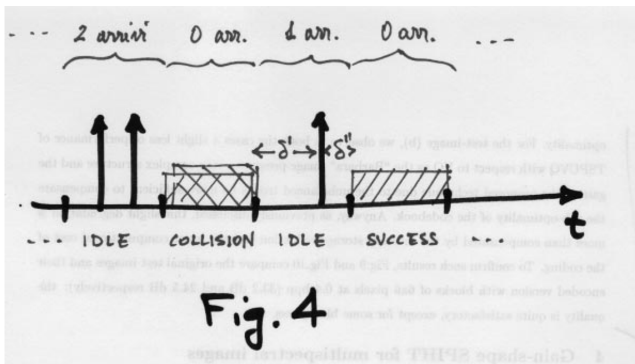


Fig.4

In questa situazione, la frame che tentiamo di trasmettere andrà incontro ad una collisione solo se nello stesso slot in cui essa arriva c'è almeno un'altro arrivo. Detto δ' l'intervallo fra inizio slot e arrivo della frame e δ'' l'intervallo fra arrivo della frame e fine slot, in analogia a quanto fatto prima possiamo scrivere

$$1 - P_C = \Pr\{\ \tau_k > \delta', \tau_{k+1} > \delta''\ \} = \Pr\{\ \tau_k > \delta'\ \}\Pr\{\ \tau_{k+1} > \delta''\ \} = e^{-\lambda(\delta'+\delta'')} = e^{-\lambda}$$

Di conseguenza il throughput sarà

$$\gamma = \lambda e^{-\lambda}$$

con un massimo di $1/e \simeq 0.37$ frame/s ottenuto per $\lambda = 1$. Rispetto all'ALOHA puro, l'efficienza è dunque raddoppiata.

Il motivo di questo miglioramento risiede nella minore sensibilità del sistema alle collisioni; infatti, l'intervallo di vulnerabilità, quello in cui non ci devono essere altri arrivi, si è ridotto ad un singolo tempo di frame (uno slot) invece di due.

E' interessante notare che la P_C fin qui considerata si può riguardare come una probabilità condizionata, vale a dire la probabilità che, dato che sia arrivata una frame in un certo istante, ne arrivino ancora una o più in un dato intervallo di vulnerabilità. Nell'ALOHA slotted è facile seguire un procedimento alternativo che fa uso solo di probabilità assolute. Ricordiamo che il throughput del sistema è pari al numero medio di frame trasmesse con successo per unità di tempo, e calcoliamo tale media. In funzione del numero di arrivi nello slot precedente (0, 1 o più), in uno slot possiamo avere tre situazioni

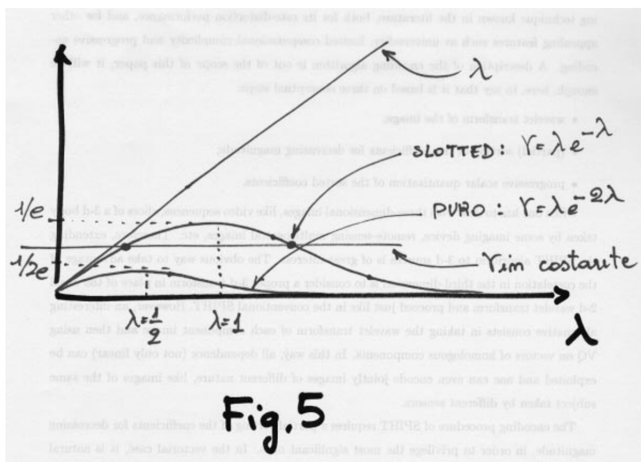
- 0 tentate trasmissioni (idle) e quindi 0 trasmissioni utili;
- 1 tentata trasmissione (success) e quindi 1 trasmissione utile;
- più di 1 tentata trasmissione (collision) e quindi di nuovo 0 trasmissioni utili.

Ricordando che gli arrivi seguono un processo di Poisson è immediato calcolare la probabilità di avere n arrivi in uno slot di durata unitaria (e perciò n tentate trasmissioni nel successivo):

$$\Pr\{\ n\ \} = \frac{e^{-\lambda T_f}(\lambda T_f)^n}{n!} = \frac{e^{-\lambda}\lambda^n}{n!}$$

Il numero medio di trasmissioni utili per slot sarà dunque

$$\gamma = 1 \times \Pr\{\ \text{success}\ \} = \lambda e^{-\lambda}$$



Per quanto riguarda l'analisi dinamica (vedi Fig.5), la situazione non è cambiata rispetto al caso dell'ALOHA puro se non per l'aumento del throughput: abbiamo ancora due punti di equilibrio, stabile il primo, instabile il secondo, per cui il sistema è sempre tendenzialmente instabile. Come già detto, però, questo risultato discende dalle ipotesi pessimistiche fatte, e l'analisi più accurata fatta qui di seguito rivela come la situazione sia in realtà meno critica.

ALOHA slotted con m stazioni senza buffering

Finora abbiamo supposto che per ogni pacchetto passato dal NET si attivasse una nuova stazione nel MAC, senza limite superiore per il numero di stazioni. In realtà, è ragionevole supporre che sulla rete ci sia solo un numero finito m di stazioni, che sono suddivise in due insiemi:

- n stazioni sono "occupate", cioè hanno in memoria, nello strato MAC, una frame che hanno già tentato senza successo di trasmettere, e che adesso devono ritrasmettere; in uno slot, queste stazioni tentano di trasmettere con probabilità q_r;

- $m-n$ stazioni sono invece "libere", non hanno nulla in memoria, e in uno slot trasmettono con probabilità q_t, pari alla probabilità che ci sia stato un arrivo dal NET nello slot precedente.

E' facile rendersi conto che il tasso di arrivi dal NET (γ_{in}), quello delle trasmissioni tentate sul canale (λ), e quello delle trasmissione riuscite (γ_{out}), dipendono tutti dallo stato della rete, e cioè dal numero di stazioni occupate n. In particolare $\lambda(n)$ è pari al numero medio di trasmissioni tentate in uno slot per cui, supponendo che le stazioni siano indipendenti l'una dall'altra, si ha

$$\lambda(n) = nq_r + (m-n)q_t \ = \ mq_t + n(q_r - q_t)$$

Si individuano chiaramente i due contributi al traffico offerto al canale, cioè le ritrasmissioni (nq_r) e il traffico entrante nel MAC:

$$\gamma_{in}(n) = (m-n)q_t$$

Bisogna infine calcolare $\gamma_{out}(n)$, che sarà ancora pari alla probabilità di successo nello slot, e cioè alla probabilità che una sola stazione, occupata o libera che sia, tenti la trasmissione per cui

$$\begin{aligned} \gamma_{out}(n) &= nq_r(1-q_r)^{n-1}(1-q_t)^{m-n} + (m-n)q_t(1-q_t)^{m-n-1}(1-q_r)^n \\ &= \left[n\tfrac{q_r}{1-q_r} + (m-n)\tfrac{q_t}{1-q_t}\right](1-q_r)^n(1-q_t)^{m-n} \\ &= \left[n\tfrac{q_r}{1-q_r} + (m-n)\tfrac{q_t}{1-q_t}\right] e^{[n\log(1-q_r)+(m-n)\log(1-q_t)]} \end{aligned}$$

Nell'ipotesi che q_r e q_t siano entrambe molto minori di 1 si può approssimare $q/(1-q)$ con q e $\log(1-q)$ con $-q$ per cui risulta infine

$$\gamma_{out}(n) \simeq [nq_r + (m-n)q_t]e^{-[nq_r+(m-n)q_t]} = \lambda(n)e^{-\lambda(n)}$$

esattamente come nell'ALOHA slotted con infinite stazioni. A questo punto possiamo considerare un grafico simile a quello di Fig.5 per studiare le prestazioni e la stabilità dell'ALOHA slotted nelle nuove condizioni qui considerate.

Dalla Fig.6 si vede che, mentre la curva per γ_{out} è invariata rispetto al caso con infinite stazioni, il traffico in ingresso γ_{in} non è più costante ma varia (ed in particolare diminuisce) al crescere di λ. Infatti, ricavando n in funzione di λ (per $q_r \neq q_t$)

$$n = \frac{\lambda - mq_t}{q_r - q_t}$$

si ottiene subito

$$\gamma_{in}(\lambda) = (m - \frac{\lambda - mq_t}{q_r - q_t})q_t$$

Quindi γ_{in} varia linearmente con λ, ed in particolare decresce all'aumentare di λ se $q_r > q_t$. Quest'ultima ipotesi è generalmente verificata, infatti è ragionevole che una stazione con una frame da ritrasmettere cerchi di liberarsene abbastanza rapidamente, in modo da poter tornare libera ed accettare nuovo traffico dal NET. Quindi il parametro q_r è un grado di libertà del

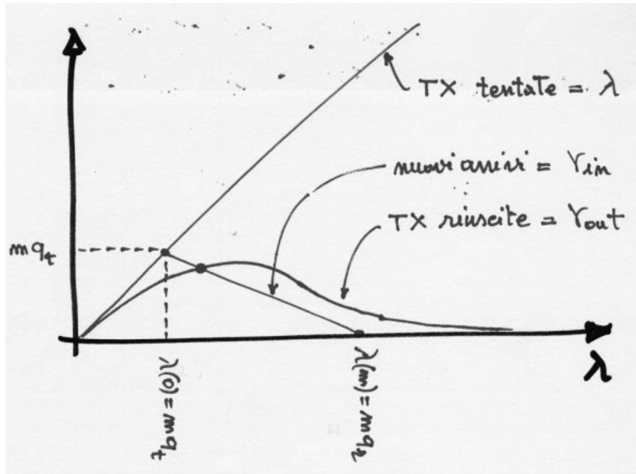


Fig. 6

protocollo, che può essere usato per massimizzare le prestazioni. Ammesso quindi che q_r sia maggiore di q_t risulta che il minimo traffico sulla rete si ha quando tutte le stazioni sono libere, poichè in tal caso $\lambda = \gamma_{in} = mq_t$, mentre il massimo traffico si ha quando tutte le stazioni sono occupate, poichè in questa condizione $\lambda = mq_r$ e $\gamma_{in} = 0$. Già da questo risultato è evidente che il sistema non può essere instabile dato che quando tutte le stazioni sono occupate non viene più accettato traffico nuovo. Naturalmente, questa è una conseguenza del fatto che il numero di stazioni è limitato; inoltre poichè una stazione occupata non può accettare traffico nuovo dovrà presto o tardi ridiventare libera e quindi ridurre il traffico complessivo offerto alla rete. In particolare, quando γ_{out} supera γ_{in} il numero di stazioni occupate tende a decrescere e con esso il tasso di tentate trasmissioni. Nel caso riportato in figura esiste un solo punto di equilibrio, stabile. La posizione di questo punto, e quindi il corrispondente throughput, può essere variata cambiando il valore di q_r. Se però q_r aumenta troppo (si ricordi che nell'analisi si è posto $q_r \ll 1$) compaiono altri due punti di equilibrio, uno instabile ed un altro stabile ma con throughput molto basso e quindi indesiderato.

CSMA-CD

Le tecniche fin qui descritte sono raramente usate nella pratica a causa della loro limitata efficienza, sempre inferiore al 37%. L'ALOHA, tuttavia, è il concetto su cui si basa una semplice ed efficiente tecnica di accesso a canale comune, il CSMA-CD, che a sua volta è il cuore dello standard IEEE 802.3 e del prodotto commerciale Ethernet che realizza tale standard e domina il mercato delle reti locali.

Nel CSMA-CD (Carrier-Sensing Multiple Access with Collision Detection) si apportano alcuni semplici miglioramenti all'ALOHA, basati sulla capacità di ogni stazione di "ascoltare" continuamente il canale (carrier sensing) e percepirne lo stato di libero, occupato, o in collisione (collision detection). In effetti, se una stazione è in grado di percepire lo stato istantaneo del canale (a volte questo non accade, come nel caso delle reti satellitari) e questo è occupato, non ha senso iniziare una trasmissione che darà luogo certamente ad una collisione: la stazione attenderà che il canale sia libero prima di tentare di trasmettere. E' evidente come il carrier sensing, elimininando gran parte delle collisioni, aumenti notevolmente l'efficienza della tecnica rispetto al semplice ALOHA. Non tutte le collisioni però possono essere prevenute. Ad esempio, considerando il tempo diviso in slot, due stazioni potrebbero percepire il canale libero in un certo slot ed iniziare entrambe la trasmissione di una frame in quello successivo. Anche lavorando a tempo continuo, può accadere che due stazioni percepiscano il canale libero e comincino a trasmettere quasi contemporaneamente (l'esatta contemporaneità avviene con probabilità zero): anche se la seconda comincia in leggero ritardo, sarà in grado di percepire la trasmissione della prima solo dopo un certo tempo, il tempo necessario perchè il segnale generato dalla prima stazione si propaghi fino alla seconda lungo il canale.

Se le collisioni non si possono tutte prevenire, è però possibile rivelarle (collision detection), dopo di che, tutte le stazioni coinvolte nella collisione smettono immediatamente di trasmettere evitando di sprecare risorse di canale per un tentativo ormai pregiudicato.

Queste due (ovvie) variazioni rispetto all'ALOHA permettono di portare l'efficienza molto vicino al 100%. Per l'analisi, consideriamo un tempo fisso τ necessario perchè ogni stazione sia sicura dello stato in cui si trova il canale, e sincronizziamo su questo tempo tutte le operazioni. In altri termini consideriamo il tempo slotted con slot di durata τ (essendo $\tau \ll T_f$ parleremo di "minislot"): partendo dall'istante t, all'istante $t + \tau$ tutte le stazioni conoscono lo stato della rete e possono prendere i provvedimenti del caso, cioè cominciare a trasmettere, smettere di trasmettere o rimanere nello stato precedente.

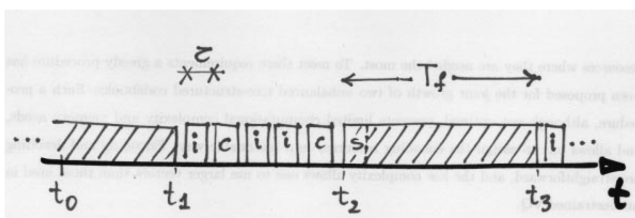


Fig. 7

In Fig.7 è mostrata una tipica successione di eventi sul canale. Supponiamo che all'istante $t_0 = 0$ sia già in atto una trasmissione: tutte le stazioni ne sono al corrente e quindi si astengono dal trasmettere a loro volta. All'istante t_1 la trasmissione termina, ma solo dopo un tempo τ (un minislot) tutte le stazioni ne sono al corrente. A questo punto le varie stazioni presenti si possono contendere il canale con tecnica ALOHA slotted, dove i minislot di contesa prendono il posto dei molto più lunghi slot di trasmissione. Potrà quindi esserci una successione di minislot idle e di collisioni finchè in un minislot non accadrà che ci sia un successo, vale a dire una singola trasmissione: quando ciò accade, poichè tutte le stazioni conoscono lo stato del canale, quella che ha cominciato a trasmettere continuerà indisturbata fino alla fine della frame al tempo t_3. A questo punto comincerà un nuovo intervallo di contesa.

Da questa descrizione risulta chiaro che esiste una rigida alternanza fra la trasmissione di una frame ed un periodo di contesa, per cui l'efficienza d'uso del canale con la tecnica CSMA/CD si può valutare come il rapporto fra il tempo effettivo di trasmissione della frame e quello complessivo necessario alla trasmissione, somma quest'ultimo del tempo di trasmissione stesso e del tempo medio T_c necessario per risolvere la contesa. A sua volta T_c sarà dato dal prodotto del numero medio N_c di minislot impegnati nel periodo di contesa per la durata τ del generico minislot, e quindi

$$\epsilon = \frac{T_f}{T_f + T_c} = \frac{T_f}{T_f + N_c \tau} = \frac{1}{1 + N_c \beta}$$

dove $\beta = \tau / T_f$ è la durata normalizzata di un minislot. Detta p la probabilità di successo, cioè la probabilità che sia tentata una singola trasmissione in un minislot, il numero medio di minislot necessario per avere un successo è $1/p$. Dall'analisi dell'ALOHA slotted, sappiamo che $p \leq 1/e$ e quindi $N_c \geq e$, con uguaglianza solo quando il tasso di tentativi è di uno per minislot. Anche se $\lambda \neq 1$, ci si può aspettare che in condizioni ragionevoli sia comunque $p > 0.1$ e quindi $N_c < 10$.

La durata di un minislot, poi, è legata al tempo di propagazione sul canale (perchè tutte le stazioni, anche le più lontane, devono conoscere lo stato del canale) e quindi possiamo dire in prima approssimazione che $\tau = l/c$ con l lunghezza del canale e c velocità di propagazione del segnale. Se infine scriviamo T_f come rapporto fra dimensione in bit della frame N_f e capacità di trasmissione del canale C, arriviamo per l'efficienza alla seguente formula

$$\epsilon = \frac{1}{1 + N_c \beta}, \text{ con } \beta = \frac{lC}{cN_f}$$

Si vede allora che l'efficienza dipende da lunghezza del canale, capacità di trasmissione, velocità di propagazione e lunghezza media delle frame. Fatta eccezione per la velocità di propagazione, il progettista può agire su questi parametri per garantire un'adeguata efficienza. Consideriamo un esempio realistico per una rete locale, in cui $l = 2.5$ km, $C = 10$ Mbps, $c = 200.000$ km/s, e $N_f = 800$ byte. Allora risulta

$$\beta = \frac{2.5 \times 10^3 \times 1.0 \times 10^7}{2.0 \times 10^8 \times 6.4 \times 10^3} \simeq 2 \times 10^{-2} \ll 1$$

Quindi, anche se prendessimo pessimisticamente $N_c = 10$ avremmo ancora un'efficienza molto elevata, superiore all'80%, mentre con $N_c = 5$ si supererebbe il 90%. In queste condizioni, dunque, il CSMA/CD risulta essere una soluzione semplice ed efficiente al problema dell'accesso al mezzo. Se invece considerassimo una rete metropolitana, in cui tipicamente $l = 100$ km e $C = 100$ Mbps, l'efficienza crollerebbe drasticamente ed il CSMA/CD non sarebbe più un'alternativa interessante.