

Gli standard della Famiglia IEEE 802.3

Ethernet

Corso di Reti di calcolatori.

Prof. Orazio Mirabella

Indice degli argomenti trattati.

1. Ethernet..... 4

1.1.	<i>Nascita ed evoluzione del protocollo.....</i>	4
1.2.	<i>Physical Layer di 802.3.....</i>	6
1.3.	<i>Formato della frame Ethernet.....</i>	9
1.4.	<i>Metodo di accesso CSMA/CD.....</i>	12
1.4.1.	L'algoritmo di regressione binaria esponenziale.....	14
1.4.2.	Round Trip Delay.....	15
1.4.3.	Parametri del protocollo.....	17
1.4.4.	Caratteristiche funzionali del CSMA/CD.....	18
1.4.5.	Domini di collisione.....	20
1.4.6.	Prestazioni di 802.3.....	21
1.5.	<i>Principali configurazioni 802.3.....</i>	24
1.5.1.	10Base5.....	24
1.5.2.	10Base2.....	25
1.5.3.	10BaseT.....	26
1.5.4.	10BaseF.....	28
1.6.	<i>Repeater 802.3.....</i>	29
1.6.1.	Definizione di un repeater.....	29
1.6.2.	Funzioni principali di un repeater.....	31
1.7.	<i>Bridge.....</i>	35
1.7.1.	Caratteristiche generali dei bridge.....	35
1.7.2.	Frammentazione.....	37
1.7.3.	Architettura fisica di un bridge.....	37
1.7.4.	Architettura logica di un bridge.....	38
1.7.5.	Tabella di instradamento.....	39
1.7.6.	Funzioni chiave: Filtering, Learning e Forwarding.....	40
1.7.7.	Algoritmo di Spanning Tree.....	42
1.7.8.	Prestazioni di un Bridge 802.3.....	45

2. Fast Ethernet..... 47

2.1.	<i>Architettura di 100BaseT.....</i>	48
2.2.	<i>100BaseT4.....</i>	51
2.3.	<i>100BaseTX.....</i>	52
2.4.	<i>100BaseFX.....</i>	53
2.5.	<i>Diametro della rete.....</i>	54
2.6.	<i>Ripetitori a 100Mbps.....</i>	55
2.7.	<i>Funzionamento full duplex.....</i>	55
2.8.	<i>Auto-Negoziazione.....</i>	56

3. Gigabit Ethernet 58

3.1.	<i>Lo standard 802.3z</i>	58
3.2.	<i>GMII (Gigabit Media Independent Interface)</i>	59
3.3.	<i>Physical Layer</i>	60
3.3.1.	1000BaseX.....	61
3.3.2.	1000BaseT	62
3.4.	<i>MAC Layer</i>	62
3.4.1.	Carrier Extension.....	63
3.4.2.	Packet Bursting	64

Capitolo 1

Ethernet

1.1. Nascita ed evoluzione del protocollo.

Ethernet nacque nei primi anni 70 grazie al consorzio DIX, formato dalle società Digital, Intel e Xerox, come rete operante a 3 Mbps su cavo coassiale. Il consorzio lavorò per circa 10 anni su una prima versione di Ethernet, la 1.0, operante a 10 Mbps e nel 1982 pubblicò le specifiche di Ethernet versione 2.0, con cui nasceva la rete locale per antonomasia, e su cui si sarebbe basato, poi, il comitato costituito dall'Institute of Electrical and Electronics Engineers per lo sviluppo dello standard IEEE 802.3.

Tale standard, pubblicato inizialmente nel 1985 e approvato definitivamente come standard ISO 8802.3 nel 1990, è, come detto, basato su Ethernet V2.0, ma differisce da questo per alcune caratteristiche logiche, riferite al livello Data Link, ed elettroniche (livello Fisico) riferite ai transceiver ed ai repeater.

Le differenze principali tra i due standard consistono nella lunghezza della sequenza di Jamming, trasmessa in caso di collisione, e nel formato della *frame*. In Ethernet 802.3 è infatti presente un campo, denominato PAD, che serve a raggiungere la lunghezza minima di 64 byte, laddove il MAC in Ethernet V2.0 scarta i pacchetti più corti.

Le reti Ethernet e 802.3 sono nate con una topologia a bus basata su cavo coassiale, con velocità trasmissiva di 10 Mbps, e coinvolgono il livello 1 della pila OSI ed il sottolivello MAC del livello 2, come si può vedere nella figura seguente:

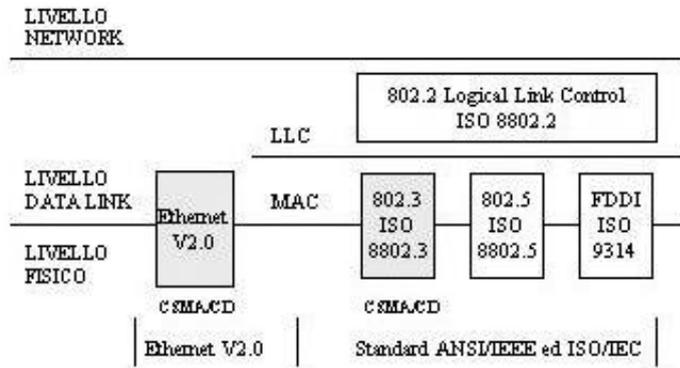


Fig. 1.1. - Relazioni tra i livelli OSI ed Ethernet e 802.3

Il MAC (Media Access Control), cioè il protocollo usato per arbitrare l'utilizzo del canale trasmissivo tra le stazioni della rete, è il **CSMA/CD**, identico in Ethernet e in 802.3. Esso è stato progettato per l'utilizzo del cavo coassiale come mezzo trasmissivo, ma è stato mantenuto inalterato anche in seguito all'introduzione in 802.3 di altri mezzi trasmissivi come la fibra ottica ed il doppino.

In questa architettura, tutti i dispositivi sono collegati ad un mezzo fisico condiviso ed hanno la stessa priorità di accesso ad esso. I segnali vengono inviati serialmente, un bit alla volta, sul canale condiviso a tutte le stazioni collegate. Così, tutti i nodi Ethernet possono "ascoltare" dal canale, ma solo un nodo alla volta può trasmettere. Se due o più nodi trasmettono nello stesso tempo, infatti, le frame da esse trasmesse ne verrebbero danneggiate risultando illeggibili.

Nei seguenti paragrafi tratteremo le caratteristiche dello standard 802.3, descrivendo il livello fisico, il formato della frame Ethernet, i supplementi allo standard relativi ai diversi mezzi trasmissivi ammessi, nonché la descrizione del protocollo di accesso al mezzo CSMA/CD.

1.2. Physical Layer di 802.3

La figura 1.2. illustra la relazione tra il modello di riferimento OSI e l'effettiva implementazione dello standard 802.3 all'interno di un nodo (o stazione) Ethernet, indicato nello standard IEEE anche con il termine DTE (Data Terminal Equipment).

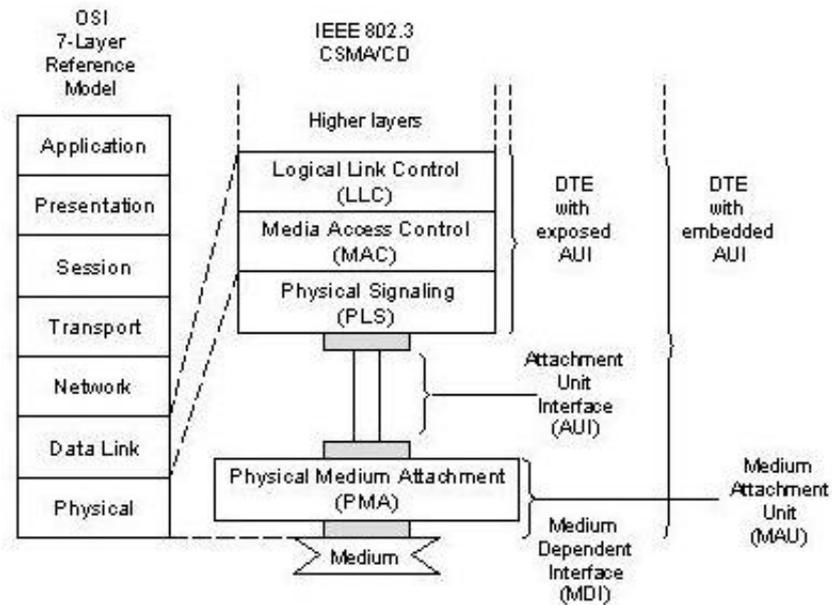


Fig. 1.2. - Relazione tra il modello di riferimento OSI e IEEE 802.3.

Come si può vedere il Livello Fisico in 802.3 è costituito dai seguenti sottosistemi:

- il *Medium Attachment Unit* (MAU);
- l'*Attachment Unit Interface* (AUI);
- il *Physical Signaling Sublayer* (PLS).

Il MAU, detto anche *transceiver*, fornisce l'interfaccia funzionale, elettrica e meccanica tra il DTE e il particolare mezzo fisico, realizzando la trasmissione/ricezione dei pacchetti tra l'interfaccia Ethernet ed il mezzo trasmissivo, la rilevazione delle collisioni e la segnalazione del corretto funzionamento del circuito di rilevazione di queste tramite l'invio, alla fine di ogni trasmissione, di un segnale detto SQE Test (*Signal Quality Error Test*). Il

sistema di trasmissione del segnale elettrico adoperato è quello definito *baseband*, che consiste nell'inviare i bit che compongono il pacchetto da spedire, direttamente sul mezzo fisico senza effettuare a priori alcuna modulazione del segnale elettrico (tecnica denominata *broadband*).

Il MAU è a sua volta composto da due parti: la PMA (*Physical Medium Attachment*) e la MDI (*Medium Dependent Interface*), che contiene l'hardware necessario per creare una connessione fisica ed elettrica diretta con il cavo di rete.

L'*Attachment Unit Interface* (AUI), detta anche cavo transceiver nello standard Ethernet V2.0, fornisce un percorso per i segnali tra l'interfaccia Ethernet ed il MAU. Tipicamente si tratta di un cavo schermato, di lunghezza massima di 50 m, con connettori a 15 poli.

Il PLS, sebbene parte del Physical Layer, è normalmente implementato localmente alla funzione MAC e svolge quattro funzioni principali: trasmissione dei dati in uscita, ricezione dei dati in ingresso, Carrier Sense (rilevamento della portante), Error Sense (rilevamento degli errori).

Il PLS fornisce l'interfaccia tra il MAC layer e l'AUI; esso, infatti, si occupa principalmente di codificare i pacchetti, ricevuti dal livello MAC, in stringhe seriali di bit da inviare sul mezzo fisico e di decodificare le stringhe seriali di bit, ricevuti dal MAU, in pacchetti. Lo schema di codifica/decodifica utilizzato è quello **Manchester**. Nessuna delle versioni 802.3 utilizza, infatti, una codifica diretta dei dati (ad esempio, zero volt per lo zero e cinque volt per l'uno), perché sarebbe difficile rilevare le collisioni e delimitare con facilità l'inizio e la fine di ogni singolo bit senza riferimento ad un clock esterno. Con una codifica diretta dei dati non sarebbe possibile, ad esempio, distinguere in ricezione una stringa di bit 0001000 da una 10000000 o 01000000, in quanto risulta impossibile distinguere un mittente inattivo da un bit 0.

Con la codifica Manchester ogni periodo di bit viene suddiviso in due intervalli uguali: un bit binario 1 viene spedito con voltaggio alto durante il

primo intervallo e basso nel secondo, mentre un bit binario 0 viene spedito esattamente nel modo opposto, prima voltaggio basso poi alto.

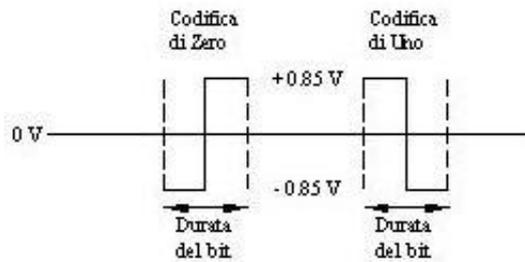


Fig. 1.3. - Codifica Manchester

I principali vantaggi di tale codifica sono:

- la facilità di sincronizzazione fra mittente e destinatario;
- il codice trasmissivo è *bilanciato*, cioè vi è uguale energia per lo zero e per l'uno, e quindi la trasmissione di dati, anche se genera diverse quantità di zeri e uni, non produce componenti in corrente continua, molto dannose perché ostacolano la trasmissione dei segnali;
- è facile rilevare le collisioni.

Si noti però che tale codifica richiede, a parità di velocità di trasmissione, una banda doppia rispetto alla codifica diretta, dal momento che ogni bit richiede la trasmissione di due valori distinti.

La funzione di Carrier Sense del PLS consente al MAC di determinare lo stato del mezzo trasmissivo: il PLS restituisce lo stato del segnale di Carrier Sense pari a ON quando c'è un'attività di trasmissione da parte del nodo, oppure viene rilevata un'attività di ricezione o collisione dalla rete.

La funzione di Error Sense consente al nodo di rilevare se il MAU e l'associata AUI sono operativi e collegati. Ciò viene fatto tramite l'analisi del segnale SQE Test inviato dal MAU alla fine di ogni trasmissione e che indica che la logica di rilevamento delle collisioni è stata testata ed è funzionante.

1.3. Formato della frame Ethernet

Il “cuore” del sistema Ethernet è la frame Ethernet, che viene utilizzata per trasferire i dati da una stazione all’altra della rete. Ogni frame Ethernet è costituita da una sequenza di bit raggruppati in un preciso formato, definito dallo standard IEEE 802.3 ed illustrato nella figura seguente:

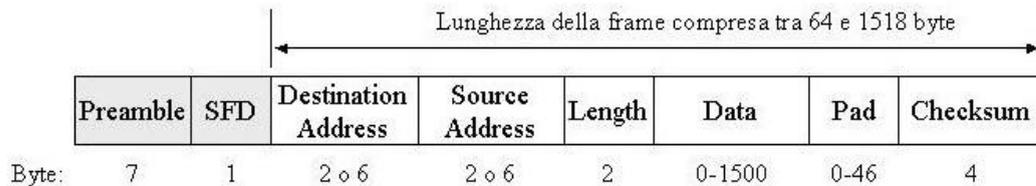


Fig. 1.4. - Formato della frame Ethernet.

Ogni frame contiene sempre i seguenti campi:

- il Preambolo;
- lo Start of Frame Delimiter o SFD;
- il Destination Address;
- il Source Address;
- il campo Length;
- il campo Dati;
- il Pad;
- il Frame Check Sequence o FCS.

Tutti i campi hanno una lunghezza fissa eccetto il campo Dati e il campo Pad.

Il **Preambolo** contiene una sequenza di 7 byte tutti uguali a 10101010, la cui codifica Manchester produce, a 10 Mbps, un'onda quadra a 10 Mhz per 5,6 μ s, che consente al ricevitore di sincronizzare il suo clock con quello del trasmettitore.

Lo **SFD** contiene un singolo byte costituito dalla sequenza 10101011, che immediatamente segue il preambolo e segnala al ricevente l’inizio del pacchetto stesso.

Il **Destination Address** ed il **Source Address** indicano rispettivamente l’indirizzo della stazione cui è destinato il pacchetto e quello della stazione che

lo ha spedito. Entrambi possono contenere 2 o 6 byte (ma tale dimensione deve essere la stessa per tutte le stazioni collegate ad una LAN!), anche se gli indirizzi usati sono sempre a 6 byte e sono unici a livello mondiale (sono cablati dentro l'interfaccia). È possibile specificare un singolo destinatario, un gruppo di destinatari (multicast) oppure un invio in broadcast a tutte le stazioni (indirizzo costituito da una sequenza di uni). Infatti, il bit di livello più alto del Destination Address è 0 per gli indirizzi ordinari e 1 per gli indirizzi di gruppo. Gli indirizzi di gruppo permettono a più stazioni di ascoltare i messaggi provenienti da un singolo indirizzo. Quando una frame viene inviata a un indirizzo di gruppo, tutte le stazioni del gruppo la ricevono. La spedizione a un gruppo di stazioni è detta **multicast**. Un indirizzo composto di tutti 1 è detto **indirizzo di broadcast**, ovvero il pacchetto sarà letto da tutte le stazioni connesse alla rete locale. Altra caratteristica interessante dell'indirizzamento è quella di utilizzare il bit 46 (vicino al bit di ordine più alto) per distinguere gli indirizzi locali da quelli globali. Gli indirizzi locali sono assegnati dall'amministratore di rete e non hanno significato al di fuori della rete locale. Invece, gli indirizzi globali sono assegnati dall'IEEE per assicurare che nessuna stazione nel mondo abbia un indirizzo uguale ad un'altra.

Il campo **Length** contiene due byte ed indica il numero di byte contenuti nel campo **Dati**, da un minimo di 0 fino ad un massimo di 1500. Occorre sottolineare che il campo Length assume un diverso significato nelle implementazioni commerciali di Ethernet, infatti in questo caso il suo posto è preso dal campo "tipo del protocollo", il quale contiene l'identificativo del protocollo che opera a livello *network* relativamente al modello di riferimento **OSI**. Nel caso del protocollo IP il suo identificativo è 0x1800. Malgrado questa differenza i due formati possono comunque coesistere, infatti poiché la lunghezza del campo dati non supera mai i 1500 byte, se questo campo contiene un valore superiore a 1500 vuol dire che indica il protocollo, nel caso contrario indica la lunghezza del campo dati.

Se il campo Dati è più corto di 46 byte, ai precedenti campi viene aggiunto un campo **Pad** (cuscinetto), calcolato in modo da garantire che la lunghezza totale del pacchetto, senza considerare il preambolo e lo SFD, non sia mai inferiore a 512 bit. Nei successivi paragrafi, si vedrà perché 802.3 impone tale lunghezza minima dei pacchetti.

Infine, l'ultimo campo della frame Ethernet è il **FCS**, che contiene un codice hash di 32 bit per i dati, utilizzato per effettuare un controllo ciclico di ridondanza (**CRC**) sull'intera frame, in modo che la stazione ricevente possa accorgersi se ha ricevuto una *frame* non corretta, ed in tal caso scartarla. Il CRC viene calcolato dalla stazione trasmittente sui campi Destination Address, Source Address, Length e Data e inserito in coda come ultimi 4 byte della frame. Lo stesso algoritmo CRC viene utilizzato dalla stazione ricevente per calcolare il valore CRC per la frame che essa sta ricevendo. Il valore calcolato viene poi confrontato con quello appeso dalla stazione trasmittente, fornendo così un meccanismo di *error-detection* (rilevazione d'errore) nel caso i dati vengono ricevuti corrotti per i rumori sul cavo.

Occorre notare che in 802.3 esiste una differenza tecnica tra i termini "pacchetto" e "frame". Un pacchetto è usato per descrivere l'intera sequenza seriale di bit trasmessa, così come viene vista sul mezzo fisico, dal primo bit del preambolo fino all'ultimo bit del campo FCS. Invece, una frame è definita come la parte della flusso (*stream*) seriale di bit compresa tra il Destination Address e il FCS inclusi. Quindi, dalle precedenti definizioni, dimensioni ammesse per le frame variano da un minimo di 64 byte ad un massimo di 1518.

Nel seguito, laddove non specificato, si utilizzeranno indifferentemente i due termini pacchetto e frame, riferendosi sempre alla seconda definizione data dall' 802.3.

1.4. Metodo di accesso CSMA/CD

Come abbiamo avuto modo di vedere in precedenza, il protocollo utilizzato in Ethernet per arbitrare l'uso del canale tra le stazioni collegate è il CSMA/CD, il cui acronimo significa Carrier Sense Multiple Access with Collision Detection.

Esso consiste in un protocollo totalmente distribuito, senza stazioni master, per permettere alle stazioni di condividere in maniera *fair* l'utilizzo del mezzo trasmissivo comune. Si tratta di un metodo a contesa, non deterministico, che cioè non garantisce un tempo di attesa limitato superiormente. Poiché mediante il collegamento a bus i trasmettitori delle stazioni si trovano ad essere “in parallelo”, è necessario evitare che più stazioni trasmettano contemporaneamente. Tuttavia, il protocollo non esclude che ciò possa comunque avvenire, e prevede un meccanismo di riconoscimento di tale evento da parte delle stazioni coinvolte in modo che possano ritentare la trasmissione in un tempo successivo.

Il protocollo opera in tre diverse fasi:

- *Carrier Sense* (ascolto della portante): ogni stazione che deve trasmettere ascolta il bus e decide di trasmettere solo se questo è libero (*listen before talking*).
- *Multiple Access*: nonostante il carrier sense è possibile che due stazioni, trovando il mezzo trasmissivo libero, decidano contemporaneamente di trasmettere; la probabilità di questo evento è aumentata dal fatto che il tempo di propagazione dei segnali sul cavo non è nullo, e quindi una stazione può credere che il mezzo sia ancora libero anche quando un'altra ha già iniziato la trasmissione.
- *Collision Detection* (Rilevazione delle collisioni): se si verifica la sovrapposizione di due trasmissioni si ha una “collisione”; per rilevarla, ogni stazione, mentre trasmette un pacchetto (o *frame*),

ascolta i segnali sul mezzo trasmissivo, confrontandoli con quelli da lei generati (*listen while talking*).

Occorre evidenziare che la collisione non è un errore trasmissivo, ma è banda impiegata per arbitrare il canale. La presenza di un numero limitato di collisioni su una rete locale non è quindi un sintomo di malfunzionamenti, ma è funzionale all'arbitraggio della rete stessa.

A seguito di un'avvenuta collisione si intraprendono le seguenti azioni:

- la stazione trasmittente sospende la trasmissione e trasmette una sequenza di *jamming* (interferenza trasmissiva), costituita da 32 bit per 802.3, in modo da consentire a tutte le stazioni di rilevare l'avvenuta trasmissione;
- le stazioni in ascolto, riconoscendo il frammento di collisione, costituito dalla parte di pacchetto trasmessa più la sequenza di jamming, scartano i bit ricevuti;
- la stazione trasmittente ritenta la trasmissione dopo un tempo pseudo-casuale per un numero di volte non superiore a 16.

Il modello CSMA/CD consisterà quindi in periodi di contesa e altri di trasmissione, con periodi di quiete (*idle*) che si avranno quando tutte le stazioni saranno inattive, come indicato nella figura seguente.

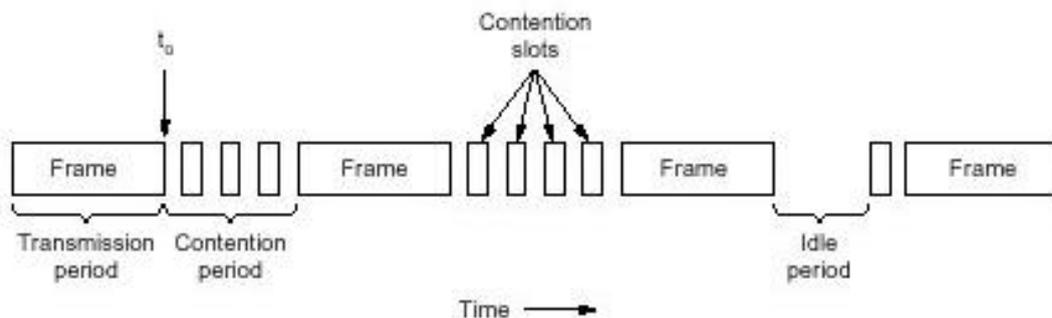


Fig. 1.5. - Stati del modello CSMA/CD: contesa, trasmissione, inattività.

1.4.1. L'algoritmo di regressione binaria esponenziale

La schedulazione della ritrasmissione in base ad un tempo di attesa pseudo-casuale evita che dopo una collisione le stesse stazioni che l'hanno generata ritrasmettano contemporaneamente; il tempo di attesa è determinato da un algoritmo di backoff detto *truncated binary exponential backoff* (BEB) o **algoritmo di regressione binaria esponenziale**.

Dopo una collisione, il tempo si considera discretizzato (slotted) con uno *slot time* pari a 51,2 μ s (corrispondenti al tempo di trasmissione di 512 bit, ossia 64 byte, pari alla lunghezza minima di una frame senza contare il preambolo ed lo Start Frame Delimiter). Il tempo di attesa prima della prossima ritrasmissione è un multiplo intero dello slot time, e viene scelto a caso in un intervallo i cui estremi dipendono da quante collisioni sono avvenute:

- dopo n collisioni, il numero r di slot time da lasciar passare è scelto a caso nell'intervallo $0 \leq r < 2^k$, con $k = \min(n, 10)$;
- dopo 16 collisioni la stazione rinuncia alla trasmissione del pacchetto e lo scarta (inviando un messaggio di errore al livello superiore).

La crescita esponenziale dell'intervallo di randomizzazione, che dopo 10 collisioni viene fissato a 1023 slot time, garantisce una buona adattabilità ad un numero variabile di stazioni, infatti:

- se il range fosse sempre piccolo, con molte stazioni si avrebbero praticamente sempre collisioni;
- se il range fosse sempre grande, per esempio sempre 1023, non ci sarebbero quasi mai collisioni ma il ritardo medio (metà range*slot time) causato da una collisione sarebbe molto elevato.

Facendo in modo che l'intervallo di randomizzazione cresca a mano a mano che avvengono collisioni consecutive, l'algoritmo assicura un basso ritardo quando poche stazioni collidono, ma assicura anche che le collisioni siano risolte in un intervallo di tempo ragionevole se molte stazioni collidono.

1.4.2. Round Trip Delay

Affinché le stazioni siano sempre in grado di rilevare le eventuali collisioni è necessario che siano rispettati alcuni vincoli tra i parametri di progetto della rete. In particolare, per garantire il corretto funzionamento del CSMA/CD è necessario che la stazione trasmittente si accorga di un'eventuale collisione entro una finestra temporale chiamata *collision window*, la quale assicura di rilevare una collisione prima di aver trasmesso completamente il pacchetto più corto (infatti, la collisione può essere rilevata **soltanto durante** la trasmissione, e non dopo). Inoltre, il frammento di collisione, costituito dalla somma della parte di pacchetto trasmessa più la sequenza di jamming che viene posta in coda, deve avere una lunghezza inferiore a 576 bit time (64 bit per la trasmissione della sequenza preambolo/SFD più i 512 bit della frame di lunghezza minima); ciò significa che l'ultimo bit di jamming deve essere trasmesso entro 575 bit time (57.5 μ s) dall'inizio della trasmissione.

Più precisamente, se la collisione viene rilevata durante il preambolo, la sequenza preambolo/SFD viene completata (tutti i 64 bit) prima di appendere la sequenza di jamming di 32 bit, il che fissa la dimensione minima consentita del frammento di collisione a 96 bit time (64 + 32). Se, invece, il preambolo viene completato prima della rilevazione della collisione, la stazione aggiungerà alla fine i 32 bit di jamming. Entrambe queste condizioni vengono in genere indicate come “in-window collision”.

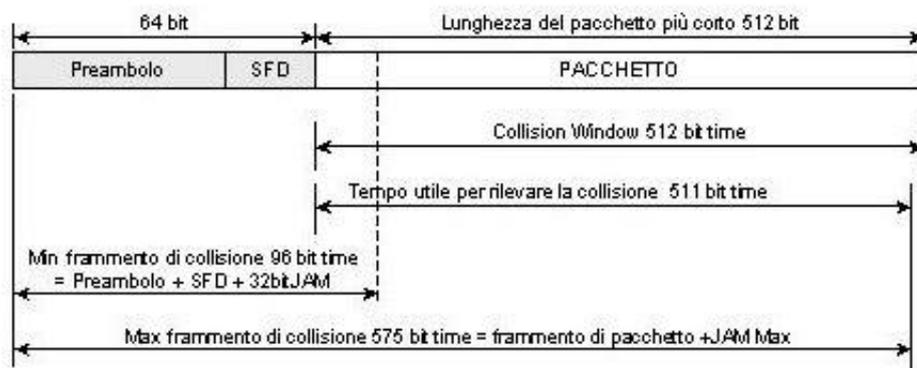


Fig. 1.6. - Collision window e minimo e massimo frammento di collisione.

La lunghezza massima del frammento di collisione pone però dei limiti più restrittivi sul tempo di ritardo della rete rispetto a quelli della collision window; infatti, se la collisione venisse rilevata dopo 511 bit time dalla trasmissione dello SFD e quindi la stazione interrompesse la trasmissione per trasmettere la sequenza di jamming, ne risulterebbe un frammento di collisione superiore a 576 bit time.

I frammenti di collisione con lunghezza superiore al massimo consentito non vengono visti come tali dalla stazione in ascolto, bensì come pacchetti contenenti errori quali *CRC error* o *alignment error*. Se la stazione trasmittente rileva la collisione entro un tempo superiore alla collision window, incrementa il contatore delle *late collision*.

Il tempo di ritardo massimo che può intercorrere da quando la stazione ha trasmesso il primo bit del preambolo a quando viene propagato l'ultimo bit di jamming nel segmento cui essa è collegata viene chiamato *round trip delay*. Esso non deve essere superiore al massimo frammento di collisione cioè 575 bit time (si veda la fig.1.6).

Nel calcolo del round trip delay si considera sempre una rete di estensione massima, e quindi contenente, come si vedrà in seguito, dei ripetitori; pertanto si prende in esame la sequenza di jamming del ripetitore che è di 96 bit.

Il round trip delay viene anche definito come il tempo necessario ad un bit per propagarsi da un estremo all'altro della rete e poi al contrario.

Per comprendere il motivo per cui la trasmissione di un pacchetto da parte di una stazione debba durare per un periodo di tempo superiore al round trip delay, è sufficiente analizzare il caso peggiore, cioè quello in cui due stazioni, A e B, sono alle estremità opposte di una rete di estensione massima, illustrato nella figura seguente.

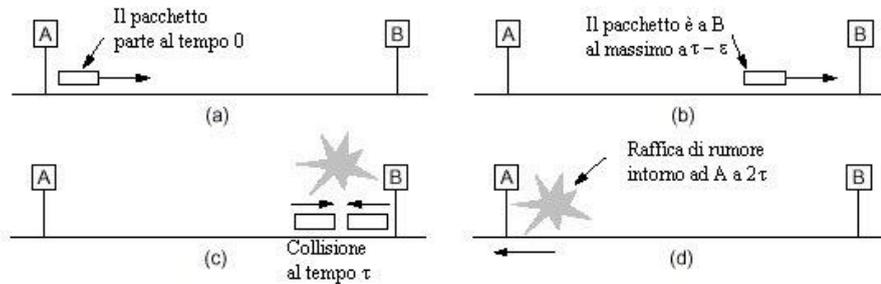


Fig. 1.7. - Rilevazione di una collisione

Si supponga che A debba trasmettere e trovi il mezzo libero. Sia τ il tempo di propagazione della frame fino all'estremità opposta. Si supponga che anche B debba trasmettere, e verifichi la disponibilità del mezzo un istante prima che il primo bit della trasmissione di A la raggiunga (cioè al tempo $\tau - \epsilon$). Non appena B inizia a trasmettere essa si accorge di ricevere più potenza di quanto sta spendendo rilevando così la collisione, per cui interrompe la trasmissione e invia la sequenza di jamming. Ma affinché anche A si accorga della collisione, la sua trasmissione deve durare almeno finché l'inizio della trasmissione di B non si è propagato fino a lei, vale a dire fino al tempo 2τ circa. Ciò impone un valore minimo alla lunghezza di ogni pacchetto trasmesso: se, infatti, A trasmettesse un pacchetto troppo piccolo e si avesse una collisione con B, la sua trasmissione terminerebbe prima dell'arrivo del primo bit trasmesso da B, per cui A concluderebbe, erroneamente, che la trasmissione del suo pacchetto fosse stata completata con successo.

Quindi una trasmissione deve durare almeno il tempo necessario ad un bit per propagarsi da un estremo all'altro (da A a B) e poi al contrario (da B ad A) e questo tempo è proprio il *round trip delay*.

1.4.3. Parametri del protocollo

Il round trip delay è uno dei parametri di progetto delle reti CSMA/CD, mentre gli altri parametri sono la dimensione massima della rete, la velocità di trasmissione (in bit/s), la lunghezza minima di ogni pacchetto e la distanza minima tra i pacchetti.

Nota la velocità di propagazione dei segnali sul cavo (circa 2/3 della velocità della luce nel vuoto) e decisa la velocità di trasmissione è possibile definire uno degli altri parametri e calcolare i rimanenti. In Ethernet la velocità di trasmissione è di 10 Mbps e la dimensione minima di ogni pacchetto è fissata in 512 bit più 64 bit di preambolo per la sincronizzazione e di *start frame delimiter*; la durata della trasmissione di un pacchetto è quindi di almeno 57.6 μ s, e questo è il massimo round trip delay ammissibile. La metà di tale tempo è il massimo tempo di propagazione di un segnale da un estremo all'altro della rete, che, alla velocità di propagazione di circa $2 \cdot 10^8$ m/s, corrisponderebbe ad una estensione massima di oltre 5 Km. In pratica, però, l'attenuazione introdotta dai cavi non consente la realizzazione di una rete di tale estensione senza ripetitori. Essi, insieme ai vari elementi attivi e passivi di collegamento, introducono dei ritardi nella propagazione dei segnali, i quali impongono, per non superare il massimo round trip delay, una riduzione dell'estensione totale dei cavi. È anche necessario introdurre un certo margine di sicurezza nei parametri temporali per considerare le tolleranze dei componenti.

1.4.4. Caratteristiche funzionali del CSMA/CD

Il protocollo CSMA/CD è responsabile delle seguenti operazioni:

- *Trasmissione dei pacchetti*: durante questa fase il MAC layer accetta un pacchetto dal livello superiore e fornisce una stringa seriale di bit al livello fisico per la loro trasmissione sul mezzo fisico.
- *Trasmissione in modalità differita di un pacchetto*, quando il canale è occupato (segnale Carrier Sense ON).
- *Generazione del campo FCS* per i pacchetti trasmessi.
- *Ricezione dei pacchetti*: durante questa fase il MAC riceve una stringa seriale di bit dal livello fisico e fornisce il pacchetto al livello superiore. Nel caso in cui il pacchetto non sia indirizzato alla stazione ricevente (singolo o multicast), né sia un pacchetto

broadcast, oppure la sua lunghezza sia inferiore al valore minimo ammesso di 64 byte, viene scartato.

- *Controllo del campo FCS in ricezione*: il MAC verifica che non ci siano errori nel pacchetto ricevuto confrontando il valore contenuto nel campo FCS del pacchetto ricevuto con quello calcolato localmente. In caso di errori scarta il pacchetto senza richiederne la ritrasmissione: il MAC gestisce infatti sempre un protocollo non connesso. Ciò è dettato dal fatto che, visto il bassissimo tasso d'errore delle LAN, si preferisce un protocollo datagram ad alte prestazioni. Occorre sottolineare altresì che nessun protocollo del sottolivello MAC garantisce la consegna affidabile: anche in assenza di collisioni, il ricevente potrebbe non aver ricopiato correttamente il pacchetto per diverse ragioni (ad esempio, per mancanza di spazio di buffer oppure per un'interruzione non raccolta).
- *Spaziatura dei pacchetti*: il MAC garantisce che tra due pacchetti consecutivi intercorra un lasso di tempo minimo pari al cosiddetto *Inter-Frame Gap* (IFG) o *Inter-Packet Gap* (IPG). Questo tempo (pari, nel caso in cui la velocità di trasmissione è 10 Mbps, a 9,6 μ s) serve a delimitare la fine di un pacchetto e a separarlo da quello successivo.
- *Rilevazione delle collisioni*: il MAC interrompe la trasmissione quando rileva una collisione.
- *Schedulazione delle ritrasmissioni*: il MAC schedula la ritrasmissione a seguito di un'avvenuta collisione dopo il periodo di tempo calcolato tramite l'algoritmo di backoff.
- *Jamming*: il MAC trasmette un messaggio di jamming a seguito della rilevazione di una collisione e dopo aver interrotto la trasmissione del pacchetto.

- *Generazione del preambolo*: in trasmissione il MAC prepone un preambolo al pacchetto che deve essere trasmesso.
- *Rimozione del preambolo* in ricezione.

Nella figura seguente vengono illustrati i diagrammi di flusso della trasmissione e della ricezione di una frame secondo il CSMA/CD.

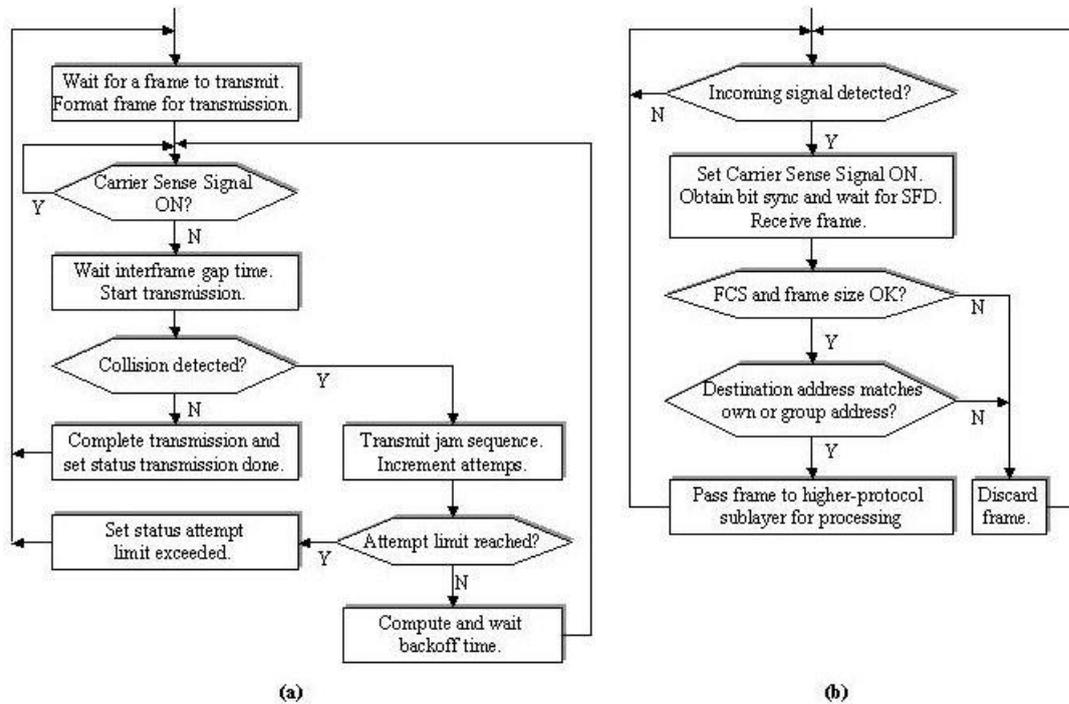


Fig. 1.8. - Protocollo CSMA/CD: (a) trasmissione di una frame; (b) ricezione di una frame.

1.4.5. Domini di collisione

In una singola rete CSMA/CD il mezzo trasmissivo è condiviso tra tutte le stazioni che se contendono l'utilizzo. Al crescere del numero di stazioni e/o del traffico aumenta la probabilità di collisioni e quindi diminuisce l'efficienza della rete.

È possibile suddividere la rete in più sottoreti in modo che la contesa del mezzo avvenga soltanto tra le stazioni appartenenti ad una singola sottorete. Si dice che ciascuna sottorete costituisce un singolo **collision domain** o dominio di collisione. Le stazioni separate da repeater fanno parte dello stesso collision domain, mentre fanno parte di collision domain diversi le stazioni separate da

apparecchiature di rete che lavorano a livelli OSI superiori al Fisico (bridge, switch, router o gateway) e che quindi sono in grado di decodificare indirizzi MAC e filtrare i pacchetti.

1.4.6. Prestazioni di 802.3

La natura non deterministica del CSMA/CD rende complessa la valutazione delle prestazioni. I valori che si trovano in letteratura sono abbastanza diversi in funzione del fatto che l'autore sia un sostenitore o un detrattore. Comunque sia, le prestazioni osservate sono molto buone, migliori di quelle stimabili in via teorica.

Peraltro, queste ultime sono fortemente influenzate dal modello di traffico che si assume. Di solito lo si assume poissoniano, ma in realtà è bursty e per di più *self similar*, ossia il suo andamento su un lungo periodo è simile a quello su un breve periodo, ricordando in questo le caratteristiche dei frattali.

La pratica ha mostrato che 802.3:

- può sopportare un carico medio del 30% (3 Mbps) con picchi del 60% (6 Mbps);
- sotto carico medio:
 - il 2-3% dei pacchetti ha una singola collisione;
 - qualche pacchetto su 10.000 ha più di una collisione.

È però indubbiamente vero che bisogna anche considerare il numero di stazioni attive sulla LAN e la direzione dei flussi di traffico. A parità di traffico totale, se vi sono poche stazioni molto attive le prestazioni sono migliori di quando vi sono molte stazioni mediamente meno attive. Nel caso limite di due sole stazioni le prestazioni possono raggiungere il 90%.

In una rete ben funzionante con un carico del 30% è altamente improbabile osservare un pacchetto che non può essere trasmesso perché supera il limite di 16 tentativi.

Esaminiamo ora brevemente le prestazioni di 802.3 in condizioni di traffico pesante e carico costante, vale a dire con k stazioni sempre pronte a trasmettere. Assumeremo una probabilità costante di ritrasmissione per ogni slot: se ogni stazione trasmette durante uno slot di contesa con probabilità p , la probabilità A che una stazione acquisisca il controllo del canale in quello slot sarà:

$$A = kp(1-p)^{k-1}$$

A è massimo quando $p=1/k$, con $A \rightarrow 1/e$ se $k \rightarrow \infty$. La probabilità che l'intervallo di contesa abbia esattamente j slot è, $A(1-A)^{j-1}$, quindi il numero di slot per contesa è dato da:

$$\sum_{j=0}^{\infty} jA(1-A)^{j-1} = \frac{1}{A}$$

Poiché ogni slot ha una durata di 2τ , l'intervallo di contesa medio, w , sarà $2\tau/A$. Assumendo p ottimale, il numero di slot di contesa non sarà mai maggiore di e , e w sarà al massimo $2\tau e \approx 5.4\tau$.

Se in media una frame impiega P secondi per la sua trasmissione, quando molte stazioni devono inviare frame,

$$\text{Efficienza del canale} = \frac{P}{P + \frac{2\tau}{A}}$$

Più lungo è il cavo, più lungo è l'intervallo di contesa. Non permettendo cavi più lunghi di 2.5 km e non più di quattro ripetitori, il tempo per percorrere tutto il giro può essere limitato a 51.2 μ s, che a 10 Mbps corrisponde a 512 bit, la lunghezza minima della frame.

Riformulando la precedente equazione in termini di lunghezza di frame, F , di larghezza di banda, B , di lunghezza del cavo L e di velocità di propagazione del segnale c , per il caso ottimale di e slot di contesa frame. Con $P = F / B$, la precedente equazione diventa:

$$\text{Efficienza del canale} = \frac{1}{1 + \frac{2BL e}{cF}}$$

Quando il secondo termine del denominatore è grande, l'efficienza della rete sarà bassa. Più precisamente, aumentando la larghezza di banda della rete o la distanza (il prodotto BL) si riduce l'efficienza per una data dimensione di frame. Sfortunatamente, molta ricerca sull'hardware di rete ha lo scopo di aumentare questo prodotto. Si vuole una ampia larghezza di banda su distanze lunghe, ciò suggerisce che forse 802.3 può non essere il modello migliore per queste applicazioni.

Nella seguente figura l'efficienza del canale viene calcolata rispetto al numero di stazioni pronte per $2\tau = 51.2 \mu\text{s}$ e una velocità di dati di 10 Mbps, usando la precedente equazione. Con un intervallo di tempo di 64 byte non sorprende che frame di 64 byte non siano efficienti. D'altra parte, con frame di 1024 byte e valore asintotico di e slot di 64 byte per intervallo di contesa, il periodo di contesa è di 174 byte e l'efficienza 0.85.

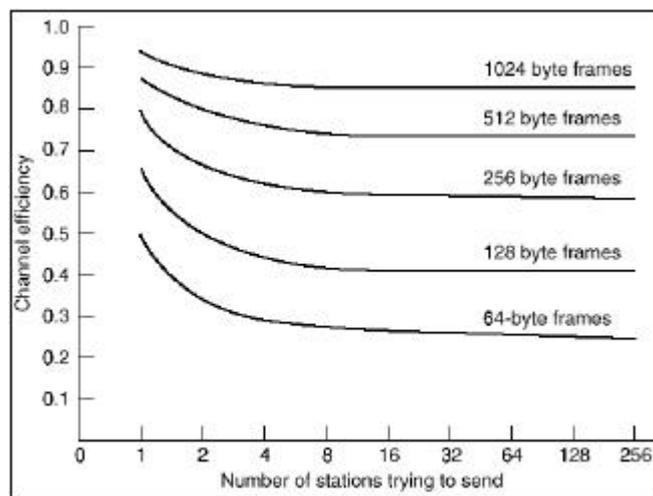


Fig. 1.9. - Efficienza di 802.3 a 10 Mbps con tempi di slot di 512 bit.

Per determinare il numero medio di stazioni pronte a trasmettere in condizioni di carico pesante, possiamo usare la seguente constatazione. Ogni frame occupa il canale per un intervallo di contesa e uno di trasmissione, per un totale di $P + w$ secondi. Il numero di frame per secondo è quindi $1 / (P + w)$. Se ogni stazione genera frame alla velocità media di λ frame/s, quando il

sistema è nello stato k , la velocità totale di trasmissione di tutte le stazioni attive è $k\lambda$ frame/s. Poiché all'equilibrio la velocità di input e di output devono essere identiche, possiamo eguagliare queste due espressioni e risolvere rispetto a k (si noti che w è una funzione di k).

1.5. Principali configurazioni 802.3

Dello standard 802.3 sono state definite diverse versioni, tra le quali citiamo:

- 10Base5
- 10Base2
- 10BaseT
- 10BaseF.

Esse operano tutte secondo lo stesso protocollo MAC alla velocità di 10Mbps e utilizzano la stessa sistema di trasmissione del segnalazione di banda di base, differenziandosi invece per il mezzo fisico adoperato per la trasmissione, per il numero massimo di stazioni collegabili, per la lunghezza massima per segmento e per l'elettronica dei *transceiver*.

1.5.1. 10Base5

Lo standard 10Base5 è il primo storicamente ad essere stato definito. La sua notazione indica che esso opera a 10 Mbps e usa segnalazione di banda di base e può sopportare segmenti lunghi fino a 500. Possono essere installate fino ad un massimo di 100 macchine su un segmento, alla distanza minima di 2,5 m l'una dall'altra.

All'interfaccia di rete viene collegata una estremità di un corto cavo (pochi metri), detto *transceiver drop cable* o AUI (*Attachment unitInterface*), all'altra estremità del quale è connesso il *transceiver* che si aggancia, con un dispositivo detto *vampire tap*, al cavo (che non viene interrotto).

Il segmento 10Base5 è costituito da un cavo coassiale da 50Ω di tipo RG213, chiamato anche “cavo thick” o “cavo giallo” o “Thick Ethernet”, le cui specifiche sono:

- impedenza $50 \pm 2 \Omega$;
- velocità di propagazione minima $0.77 c$, dove c è la velocità della luce;
- attenuazione massima del segmento (500 m) 8.5 dB misurata a 10 MHz e 6 dB misurata a 5 MHz.

Fino a cinque segmenti 10Base5 possono essere combinati assieme in una singola rete Ethernet, usando fino a quattro *repeater*, che elettricamente amplificano e condizionano il segnale da un segmento all’altro del cavo, consentendo così di estendere la lunghezza dell’intera rete fino a 2.5 Km senza la minima perdita di qualità del segnale.

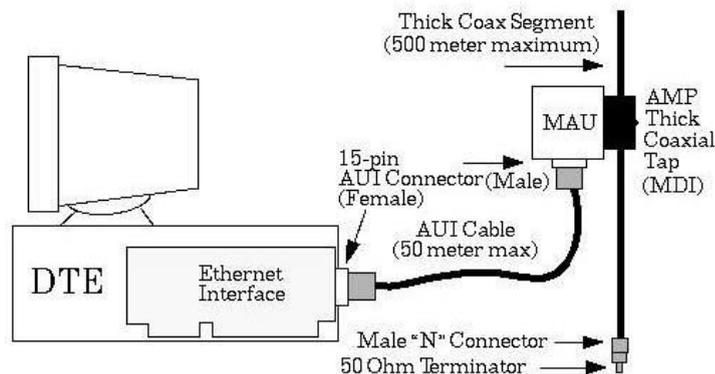


Fig. 1.10. - Cablaggio Thick Ethernet

1.5.2. 10Base2

Lo standard 10Base2 è una variante del 10Base5 in cui si utilizza un cavo coassiale più sottile (*thin* o Cheapernet) con una maggiore attenuazione che riduce la massima lunghezza del segmento a 185 m, ma con un minor costo e una maggiore flessibilità di utilizzo.

Le caratteristiche principali sono:

- velocità trasmissiva: 10 Mbps;
- lunghezza massima del segmento: 185 m;

- lunghezza massima del cavo AUI: 50 m;
- distanza minima tra due transceiver: 0.5 m;
- numero massimo di transceiver: 30;
- Il segmento 10Base2 è costituito da un cavo coassiale da 50 Ω di tipo RG58 A/U, le cui specifiche minime richieste sono:
 - impedenza $50 \pm 2 \Omega$;
 - velocità di propagazione minima 0.65 c, dove c è la velocità della luce;
 - attenuazione massima del segmento 8.5 dB a 10 MHz e 6 dB a 5 MHz.
- il transceiver è integrato nell'interfaccia di rete ed è collegato al cavo coassiale tramite un connettore a T di tipo BNC.

Fino a cinque segmenti 10Base2 possono essere combinati assieme in una singola rete Ethernet, usando fino a quattro *repeater*, a patto che la lunghezza totale dei segmenti non superi i 925 m.

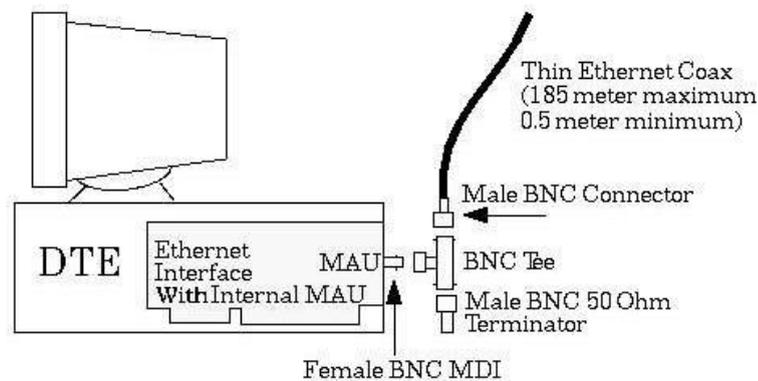


Fig. 1.11. - Cablaggio Thin Ethernet

1.5.3. 10BaseT

Lo standard 10BaseT ammette solo collegamenti punto-punto tra stazioni tramite UTP (Unshielded Twisted Pair) di categoria 3 o doppiamente schermato, consistente di due fili isolati delicatamente intrecciati. La particolarità di questo standard implica l'utilizzo di ripetitori multiporta (hub)

per poter connettere più di due stazioni in rete e la topologia è quindi di tipo stellare.

Il MAU 10BaseT, oltre alle solite funzioni, si occupa di verificare costantemente l'integrità del collegamento, inviando sulla linea, quando è libera un segnale detto TP_IDLE. L'elemento MDI è costituito dai circuiti driver e receiver per il doppino e da una presa RJ-45 (jack a 8 contatti con chiave centrale).

Pin Number	Signal
1	TD+
2	TD-
3	RD+
4	Unused
5	Unused
6	RD-
7	Unused
8	Unused

Fig. 1.12. - Assegnazione dei contatti su RJ-45.

Il segmento 10BaseT è costituito da un cavo avente come minimo due coppie intrecciate da 100 Ω , una per la trasmissione e l'altra per la ricezione, con le seguenti caratteristiche minime:

- impedenza $100 \pm 15 \Omega$ misurata tra 1 e 16 MHz;
- velocità di propagazione minima 0.585 c, dove c è la velocità della luce;
- attenuazione massima del segmento 11.5 dB nelle frequenze comprese tra 5 e 10 MHz;

Normalmente, la lunghezza massima del segmento è :

- 100 m, nel caso in cui si utilizzano UTP di categoria 3;
- 150 m, nel caso in cui si utilizzano doppini di categoria 5, i quali sono simili ai doppini di cat.3, ma con più intrecciamenti per centimetro e isolamento in Teflon, che porta a meno interferenze e una migliore qualità del segnale su distanze più lunghe.

Grazie alla topologia stellare, il 10BaseT presenta notevoli vantaggi in termini di fault-tolerance e di fault-isolation rispetto al 10Base2 e al 10Base5. Infatti, in una rete basata su cavo coassiale, una rottura del cavo, un connettore staccato o una spina rotta, portano inevitabilmente al crash e all'isolamento dell'intero bus, rendendo complicata l'individuazione del guasto. Invece, la topologia a stella, tipica del 10BaseT e del 10BaseF, di cui si dirà dopo, dedica un link ad ogni utente. Pertanto, se un singolo link dovesse guastarsi, solo l'utente su quel link se risentirebbe, continuando a garantire il funzionamento del resto della rete. Ciò è reso possibile dal link test e dall'auto-partizionamento delle sue porte realizzato da un hub 10BaseT, di cui si parlerà brevemente in seguito.

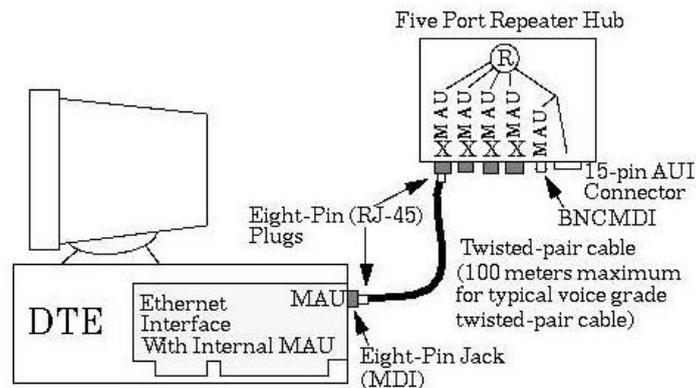


Fig. 1.13. - Cablaggio Ethernet tramite Hub

1.5.4. 10BaseF

Lo standard 10BaseF si occupa di regolamentare l'utilizzo della fibra ottica come mezzo trasmissivo per LAN 802.3. Questo tipo di collegamento è piuttosto costoso per i prezzi dei connettori e dei terminatori, ma offre livelli di rumore molto bassi ed è un metodo scelto in passato spesso quando si trattava di collegare edifici o hub molto lontani. Il 10BaseF si suddivide in 3 sotto-standard che sono:

- 1) 10BaseFP;
- 2) 10BaseFB;
- 3) 10BaseFL.

Il 10BaseFP (FP sta per Fiber Passive) permette l'utilizzo della topologia a stella passiva basata sul concetto di splitter ottico, ovvero un ripartitore di segnale luminoso; questa tecnica implica che gran parte del segnale luminoso vada perso nella stella e per questo motivo è necessario che i transceiver abbiano un'elevata dinamica. La lunghezza massima di fibra ottica che si può avere tra la stella ottica passiva ed il MAU è di 500 m.

Nel 10BaseFB (FB sta per Fiber Backbone) la fibra ottica è ammessa solo sulle dorsali e i transceiver sono integrati ai repeater. I transceiver ottici generano inoltre due tipi di segnali:

1. SIDL (synchronous idle): permette di mantenere la sincronizzazione tra il repeater ed il MAU. Consiste in un impulso di durata pari a due periodi di bit ad una frequenza di 2.5 MHz.
2. RF (remote fault): segnala un guasto del link. Consiste in un impulso di durata pari a tre periodi di bit ad una frequenza di 1.667 MHz.

Questi segnali sincroni vengono utilizzati dal ricevitore per sincronizzare il clock con il trasmettitore evitando di doversi risincronizzare su ogni pacchetto.

Nel 10BaseFL (FL sta per Fiber Link) la fibra ottica può interconnettere anche le stazioni con topologia punto-punto e stellare e la lunghezza massima del segmento è di 2000 m.

1.6. Repeater 802.3

1.6.1. Definizione di un repeater.

Un repeater (o ripetitore) è un dispositivo che consente di accrescere l'estensione di una topologia di rete fisica, al di là delle normali restrizioni imposte dall'utilizzo di un singolo cavo, in termini di distanza e di numero di nodi collegati. Il repeater opera a livello Fisico, come illustrato nella figura seguente, e il suo comportamento è identico per tutte le implementazioni a 10 Mbps ed indipendente dalla tecnologia di livello fisico.

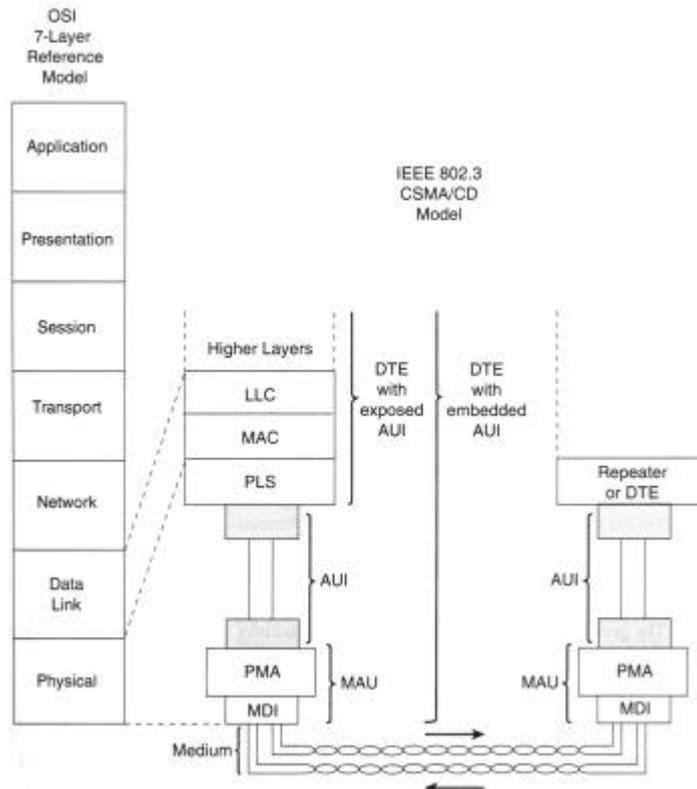


Fig. 1.14. - Architettura di un Repeater 802.3.

Nella sua forma più generica un repeater è un dispositivo ad n porte, nel quale i segnali ricevuti su una porta vengono amplificati e ritrasmessi a tutte le altre porte. Se un repeater scopre che due o più porte sono nello stato di ricezione, si viene a creare una collisione e il repeater invierà la sequenza di jam su tutte le porte, comprese quelle di ricezione. In questo modo, un nodo trasmittente collegato alla porta di un repeater si accorgerà che due o più dispositivi (ovvero esso stesso ed il repeater) stanno trasmettendo contemporaneamente, rileverà la collisione eseguirà il backoff.

Dal momento che il repeater ha almeno due porte, esso è in grado di collegare mezzi fisici differenti. Per esempio, esso può consentire ad un segmento di rete basato su cavo coassiale di essere collegato ad uno o più segmenti di rete basati su UTP. Il risultato è una singola rete CSMA/CD omogenea.

1.6.2. Funzioni principali di un repeater

Di seguito riportiamo alcune delle funzioni principali di un ripetitore, così come sono state definite dallo standard 802.3.

Signal Restoration

Lo scopo di un repeater è ripristinare il segnale ricevuto su un segmento e ritrasmetterlo sugli altri segmenti rimuovendo gli effetti di distorsione di ampiezza (causata dall'attenuazione del segnale) e di distorsione temporale (causata dal *jitter*, definito come lo scostamento del bit time dal valore nominale) che il segnale subisce quando si propaga attraverso ciascun segmento di rete.

A causa della rigenerazione realizzata nel repeater, ciascuna porta è isolata dall'altra e per il proprio funzionamento non fa affidamento su quello delle porte vicine, o di dispositivi connessi ad essa; inoltre le performance di ogni singola porta possono essere monitorate.

Preamble Restoration

Quando un pacchetto viene ricevuto ad una porta del repeater (o ad una porta qualsiasi, incluso una generica stazione Ethernet), alcuni bit del preambolo vengono persi, a causa del campionamento, temporizzazioni asincrone ed effetti di *clock recovery*. Nel caso di una generica stazione ciò non è importante dal momento che, essendo il preambolo ampio, il restante preambolo viene comunque rimosso dal MAC. Invece, in un repeater i bit di preambolo persi devono essere rigenerati, dal momento che ci possono essere diversi repeater interconnessi e una successiva perdita può lasciare un preambolo insufficiente perché una stazione remota sia in grado di decodificare con successo il pacchetto ricevuto.

Il ripetitore rigenera il preambolo proprio per evitare ciò, e a tal scopo esso deve includere una memoria FIFO, dal momento che esso deve immagazzinare la frame entrante, che è stata accorciata dalla perdita di

preambolo, finché non viene ripristinata la sua lunghezza originaria. Il numero di bit di cui è consentito ripristinare il preambolo è attentamente specificato, così da impedire l'eccessiva crescita del preambolo stesso.

Collision Detection

Quando un repeater sta ripetendo un pacchetto ricevuto su una porta a tutte le altre porte (quelle non disabilitate o partizionate), ed una di queste trasmettendo rileva una collisione, questa viene definita *transmit collision*. In questo caso, il repeater trasmette la sequenza di jam di almeno 96 bit su tutte le porte, in modo da assicurare che l'inizio della collisione sia segnalato a tutti i dispositivi nella rete, inclusa la originaria porta di ricezione.

Se un repeater rileva una collisione dalla porta dalla quale esso sta al momento ricevendo, essa viene definita *receive collision*. In questo caso, il repeater trasmette la sequenza di jam di almeno 96 bit a tutte le porte eccetto quella originaria di ricezione. Non c'è, infatti, bisogno di propagare l'inizio della collisione alla porta che ha già rilevato una collisione, bensì la collisione deve essere propagata al resto della rete.

Se una porta, inizialmente attiva in ricezione, diventa inattiva prima che siano stati ricevuti 96 bit, ciò viene pure classificato dal repeater come una *receive collision*. Frammenti di pacchetto più piccoli di 96 bit possono essere solo il risultato di una collisione, per cui tali frammenti vengono estesi dal repeater fino a 96 bit in modo da assicurare che essi siano esaminati da tutte le stazioni della rete. che i scarteranno come messaggi illegalmente breve (definiti come *runt packet*, cioè pacchetti più corti del normale).

MAU Jabber Lockup Protection (MJLP)

La funzione di MAU Jabber Lockup Protection permette al repeater 802.3 di usare lo stesso MAU di una stazione: ciò consente l'uso di uno dei MAU a 10 Mbps, precedentemente descritti, rendendo il repeater indipendente dal mezzo fisico.

Ciascun MAU ha un *jabber timer* che impedisce al DTE di trasmettere per un tempo eccessivo (20–150 ms), costringendo il MAU ad interrompere il processo di trasmissione dei dati e ad indicare una collisione. Comunque, quando il DTE è connesso ad un repeater, tali effetti sono non desiderabili.

Per impedire al MAU di entrare nello stato di jabber, il repeater implementa un *MJLP timer* di durata inferiore, che assicura che i valori dei jabber timer dei MAU non siano mai raggiunti. Il MJLP timer ha un valore di 4–7.5 ms. Se un repeater rileva un'attività di ricezione per un tempo superiore al MJLP timer, esso interrompe la trasmissione e la riabilita dopo un tempo che va da 9.6 a 11.6 ms.

Port Partitioning

Un repeater può facoltativamente isolare, per un determinato periodo di tempo, una porta che sperimenti una eccessiva durata della collisione oppure un numero eccessivo (oltre 30) di collisioni consecutive. In questo modo, tramite l'*auto-partizionamento* delle porte, il repeater può essere utilizzato per isolare automaticamente segmenti difettosi dal resto della rete.

Quando una porta è nello stato “partizionato”, il ripetitore continua a trasmettere alla porta, mentre i segnali ricevuti da essa vengono monitorati ma non ritrasmessi dal repeater al resto della rete.

Il repeater può riabilitare la porta se esso trasmette ad essa nel normale corso della ritrasmissione e non rileva alcuna collisione per un certo periodo di tempo, oppure se rileva da essa una corretta ricezione senza collisione.

Ritardo e Inter-Packet Gap Shrinkage

Sebbene i repeater sono dispositivi molto utili che consentono di espandere la rete e di interconnettere differenti mezzi fisici, essi introducono alcuni altri effetti che devono essere tenuti in conto nella costruzione di grosse reti. Due effetti principali devono essere considerati. Il primo consiste nel fatto che un ripetitore introduce un ritardo nella propagazione del segnale da una porta

all'altra, il quale deve essere conteggiato nel calcolo del round trip delay della rete. Il secondo effetto è definito "Inter-Packet Gap Shrinkage", ossia riduzione dell'Inter-Packet Gap, e può portare due pacchetti ad incollarsi e a diventare indistinguibili. La riduzione dell'Inter-Packet Gap può avvenire perché i pacchetti possono venire ritardati in modo diverso da un repeater. Infatti, il repeater non impiega sempre lo stesso numero di bit di preambolo per sincronizzare il suo ricevitore sul clock del trasmettitore. Quindi il numero di bit "persi" nel preambolo da un ripetitore, e che devono da questi essere rigenerati introducendo un ritardo di trasmissione, varia da pacchetto a pacchetto.

Quando si hanno due pacchetti consecutivi e il primo subisce un ritardo maggiore del secondo l'IPG si riduce. Ciò viene illustrato nell'esempio di fig.1.11, dove due pacchetti vengono emessi da una stazione con un minimo IPG ($9.6 \mu\text{s}$ o 96 bit time), ma il ritardo introdotto nella loro ritrasmissione da parte del repeater è, rispettivamente, di 15 bit time per il primo pacchetto e di 10 bit time per il secondo.

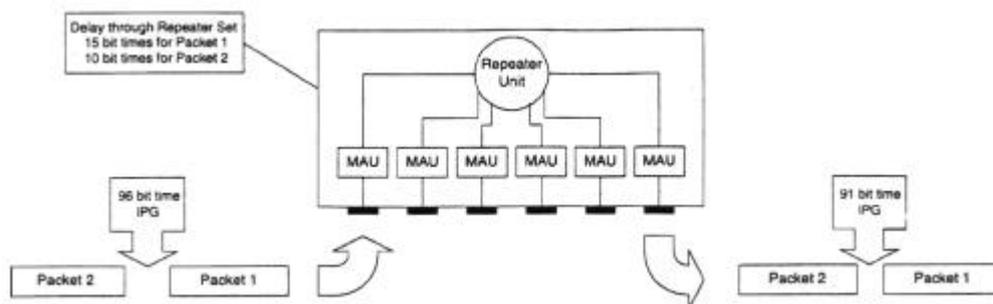


Fig. 1.15. - Esempio di Inter-Packet Shrinkage.

Ciò comporta che in uscita al ripetitore i due pacchetti saranno ora distanziati tra di loro di soli $9.1 \mu\text{s}$ anziché $9.6 \mu\text{s}$.

Se l'IPG tra due pacchetti diventa troppo piccolo un DTE o il ricevitore di un repeater potranno non essere più in grado di risincronizzare i loro clock con quello del trasmettitore, non riuscendo più così a decodificare correttamente il pacchetto in ingresso.

La conseguenza principale dei problemi dovuti al ritardo introdotti dai ripetitori e dell'IPG Shrinkage è che il numero di ripetitori ammessi all'interno di un dominio di collisione deve essere limitato.

1.7. Bridge

I bridge sono dispositivi che operano al livello MAC, quindi sopra al livello fisico al quale lavorano i ripetitori, ed il loro funzionamento è definito dallo standard IEEE 802.1D. Un bridge può collegare assieme tecnologie MAC identiche (come per esempio Ethernet ad Ethernet) oppure dissimili (come per esempio Ethernet a Token Ring). Esso utilizza le informazioni sul Destination Address e sul Source Address contenute nelle frame per tentare di inoltrare queste ultime in una maniera “più intelligente” piuttosto che semplicemente duplicare su tutte le altre porte ogni frame ricevuta (ciò viene definito nella terminologia dei bridge “flooding”).

Il più semplice tipo di bridge è un dispositivo a due porte. Nel caso di un bridge “Ethernet-Ethernet”, il dispositivo non solo consente l'estensione della topologia di rete fisica, ma anche isola le sottoreti su ogni porta in separati domini di collisione. Ciò significa che una collisione che avviene su un lato del bridge non viene trasferita all'altro. Solo le frame MAC valide che vengono generate su di una sottorete e che sono indirizzate ad un nodo sull'altra sottorete vengono propagate attraverso il bridge, in modo così da isolare il traffico locale all'interno della sola sottorete.

Tale ritrasmissione avviene con una modalità di “store & forward”, cioè la frame è ricevuta dal bridge per intera e poi eventualmente ritrasmessa. Questo permette di superare i limiti sulle distanze massime e sul numero massimo di sistemi collegabili.

1.7.1. Caratteristiche generali dei bridge

I bridge hanno le seguenti caratteristiche generali:

- operano, come già detto, al livello 2 del modello di riferimento OSI, sottolivello MAC, e per questo sono molto spesso detti MAC-Bridge;
- hanno algoritmi di instradamento molto semplici: ogni bridge calcola autonomamente le sue tabelle di instradamento senza interagire con gli altri bridge, tramite un algoritmo di routing isolato;
- si utilizzano in genere per interconnessioni locali, anche se sono stati usati nel passato, in modo un po' problematico, anche per interconnessioni geografiche.

I bridge possono essere realizzati secondo due filosofie diverse che differiscono per il luogo ove vengono memorizzate le *tabelle di instradamento* o *tabelle di filtraggio* (o *filtering database*):

- *transparent bridge*: sono i bridge conformi allo standard 802.1D, di derivazione Ethernet. Hanno le tabelle di instradamento a bordo e sono trasparenti, nel senso che i sistemi interconnessi alle LAN ignorano la loro presenza;
- *source routing bridge*: sono i bridge di derivazione Token Ring. Non hanno tabelle di instradamento a bordo, ma esse sono invece mantenute dai sistemi connessi alle LAN che in fase di trasmissione del pacchetto devono specificare esplicitamente il cammino che il pacchetto dovrà fare per giungere a destinazione, indicando tutti i bridge da attraversare (che quindi vengono indirizzati esplicitamente)¹.

Come già visto, i bridge per instradare i pacchetti sulla rete hanno bisogno di costruirsi delle tabelle di instradamento. Se la topologia della LAN suddivisa in più sottoreti è ad albero, la costruzione di tali tabelle può avvenire con un algoritmo molto semplice, in modo automatico, tramite un processo di apprendimento (*learning process*).

¹ Nel seguito verranno descritti soltanto i transparent bridge.

Poiché, tuttavia, spesso è preferibile avere topologie magliate per ragioni di affidabilità, occorre integrare il learning process con un algoritmo di *spanning tree* per riportare dinamicamente una topologia magliata ad una topologia ad albero, escludendo dall'operatività opportune porte di opportuni bridge.

Tale problema non esiste nei source routing bridge, in quanto il pacchetto, quando viene generato, contiene la specifica completa del cammino che dovrà seguire.

1.7.2. Frammentazione

I bridge che operano tra LAN eterogenee hanno il problema aggiuntivo della diversa lunghezza massima del campo dati del pacchetto MAC. La dimensione massima del campo dati del pacchetto varia, infatti, a seconda degli standard, ad esempio è di 1500 byte per 802.3, 17749 byte per 802.5 (Token Ring) e 4478 byte per FDDI.

Poiché è impossibile violare tali dimensioni massime, quando un bridge deve ritrasmettere un pacchetto di dimensione superiore a quella massima ammessa ha due possibilità: scartare il pacchetto o frammentarlo.

La frammentazione è un tipico compito del livello 3 (network) e non può essere realizzata da un bridge in modo generalizzato per tutti i protocolli.

1.7.3. Architettura fisica di un bridge

I bridge sono costituiti da una o più CPU, una memoria e due o più interfacce che interconnettono LAN.

La ROM contiene il software che realizza tutte le funzionalità del bridge in conformità con lo standard IEEE 802.1D. La memoria RAM contiene le tabelle di instradamento, i buffer per i dati ed un'area di memoria utilizzata dal software per le strutture dati interne.

L'interfaccia è costituita per una parte da dispositivi elettronici conformi ai diversi standard per le LAN (ad esempio, 802.3, 802.5, FDDI) e per la restante

parte da dispositivi per la connessione ai diversi mezzi fisici trasmissivi (ad esempio: UTP, STP, fibra ottica, cavo coassiale).

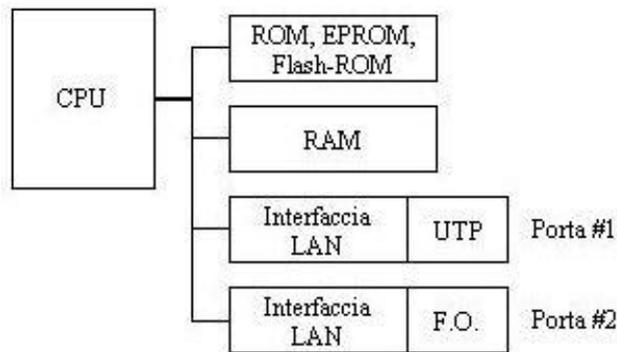


Fig. 1.16. Architettura fisica di un bridge.

1.7.4. Architettura logica di un bridge

Oltre a tale organizzazione fisica si può considerare un bridge costituito, da un punto di vista logico, dai seguenti tre elementi:

- le *porte*, che possono essere due o più;
- l'entità di ritrasmissioni e filtraggio chiamata *MAC relay entity*;
- le entità di livello superiore chiamate *higher layer entities*;

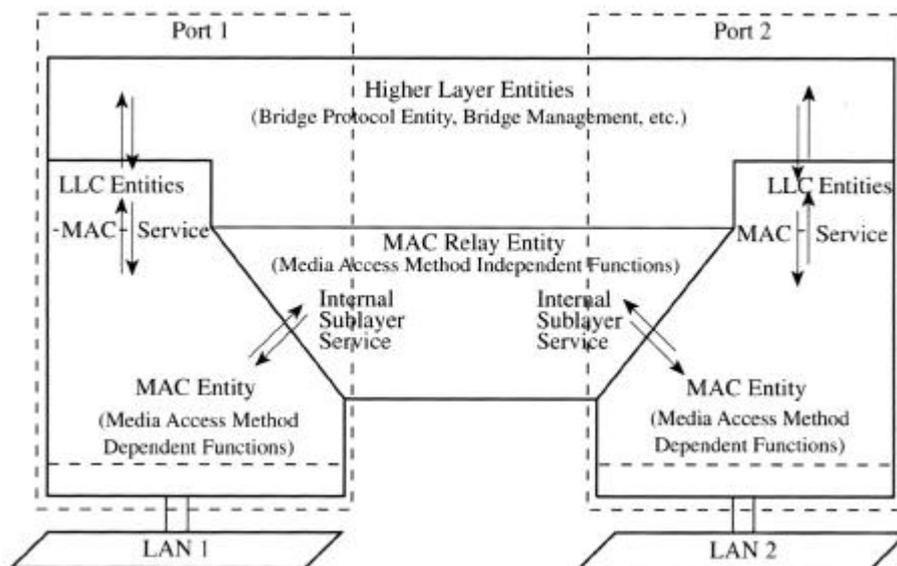


Fig. 1.17. Architettura logica di un bridge.

Ogni porta riceve/trasmette i pacchetti dalla/alla LAN a cui è connessa usando il servizio fornito dall'entità MAC associata a tale porta. L'entità MAC di ogni porta tratta tutte le funzioni di metodo di accesso al mezzo nel modo specificato dai relativi standard IEEE 802.

L'entità di ritrasmissioni (MAC relay entity) si occupa di ritrasmettere i pacchetti tra due porte, filtrare i pacchetti ed apprendere le informazioni di filtraggio.

Le entità di livello superiore (higher layer entities), che fanno uso delle procedure di Logical Link Control (LLC), fornite separatamente per ogni porta, sono principalmente due:

- *bridge protocol entity* che si occupa del calcolo e della configurazione della topologia della rete suddivisa (algoritmo di Spanning Tree)
- *bridge management entity* che si occupa di governare e controllare le funzioni del bridge.

1.7.5. Tabella di instradamento

La tabella di instradamento di un bridge, detta anche tabella di filtraggio o filtering database, è costituita da un insieme di righe (*entry*), ciascuna contenente l'indirizzo MAC a 48 bit di una stazione, l'identificativo della porta del bridge cui essa è connessa e lo stato della porta.

Alcune entry sono statiche, ossia esplicitamente configurate tramite operazioni di management, mentre altre sono dinamiche, ossia vengono registrate automaticamente dal processo di apprendimento.

Il filtering database fornisce le informazioni al processo di *forwarding* (inoltro), per decidere se inoltrare un pacchetto avente un certo indirizzo di destinazione su una data porta.

Una entry dinamica non viene creata se esiste già una entry statica relativa allo stesso indirizzo MAC. All'atto della creazione di una entry statica, una eventuale entry dinamica relativa allo stesso indirizzo MAC viene rimossa.

Le entry dinamiche sono soggette ad un meccanismo di timeout: se l'entry non viene aggiornato per un tempo superiore al parametro *ageing time* (valore di default cinque minuti), la entry viene automaticamente rimossa; al contrario, le entry statiche non sono soggette a timeout.

L'informazione di stato associata ad ogni porta del bridge governa la sua partecipazione alla suddivisione della LAN. A livello di management si dichiara la porta attiva (*enabled*) o non attiva (*disabled*). Le porte attive possono trovarsi in vari sottostati in funzione delle decisioni prese dall'algoritmo di Spanning Tree. Una porta attiva si dice in uno stato di *forwarding* se partecipa alla ritrasmissioni di pacchetti, di *learning* se si limita ad apprendere, di *listening* se si limita ad ascoltare il traffico e di *blocking* se è stata definita come porta di backup dall'algoritmo di Spanning Tree.

1.7.6. Funzioni chiave: Filtering, Learning e Forwarding

Le funzioni principali di un bridge sono:

- ricevere, filtrare e ritrasmettere i pacchetti,
- mantenere le informazioni richieste per prendere le decisioni di filtraggio;
- governare e controllare (*management*) quanto sopra citato.

I pacchetti trasmessi da una sottorete S1 verso una sottorete vengono confinati dai bridge nelle LAN che formano il percorso tra S1 e S2. Questo tipo di filtraggio è il più comune e serve a ridurre il traffico globale.

Qualora esistano più percorsi per raggiungere S2, i bridge operano un ulteriore filtraggio per prevenire la duplicazione di pacchetti.

Altri tipi di filtraggio possono essere effettuati intervenendo direttamente sul bridge con sistemi di management e definendo filtri esclusivi e filtri inclusivi. I primi limitano l'apprendimento del learning process, imponendo classi di pacchetti che devono essere sempre filtrati, cioè non ritrasmessi; i secondi, invece, includono classi di pacchetti che non devono essere filtrati. I pacchetti che superano il filtraggio vengono ritrasmessi sulla LAN di destinazione con il

protocollo di livello MAC proprio di tale rete e mantenendo la sequenza con cui i pacchetti sono stati ricevuti.

Il processo di *learning* (o apprendimento) osserva il Source Address (MAC) dei pacchetti ricevuti su ogni porta e crea o aggiorna le entry dinamiche del filtering database, condizionatamente allo stato delle porte. Il Source Address indica al learning process che la stazione con quell'indirizzo è raggiungibile attraverso la porta che ha ricevuto il pacchetto. Tale metodologia di apprendimento è anche detta di *routing isolato – backwarding learning*, in quanto un indirizzo di sorgente attuale crea o aggiorna una entry dinamica della tabella di instradamento relativamente ad una destinazione che potrà essere utilizzata in futuro.

Quando un pacchetto viene ricevuto su una porta del bridge, esso viene affidato al processo di inoltramento (*forwarding*) che deve deciderne un eventuale accodamento per la trasmissione su altre porte. Condizione necessaria è che sia la porta di ricezione sia le porte di destinazione si trovino in stato di forwarding.

Il processo di inoltramento consulta la tabella di instradamento per determinare su quale porta eventualmente accodare il pacchetto in funzione del suo indirizzo di destinazione. Se viene trovata una entry contenente un indirizzo uguale a quello di destinazione del pacchetto, questo viene inoltrato sulla porta corrispondente. In caso contrario, il pacchetto viene inoltrato su tutte le porte disponibili (eccetto quella di ricezione): si dice, in questo caso, che viene realizzato il “flooding” del pacchetto. Finché una specifica associazione Source Address-porta non viene rilevata e appresa da un pacchetto ricevuto, il filtering database non sarà in grado di “risolvere” ogni pacchetto indirizzato a quell'indirizzo MAC, per cui ciò comporterà il flooding a tutte le porte disponibili finché l'indirizzo non sarà appreso dal filtering database.

Lo standard 802.1D specifica il funzionamento dei bridge “store & forward”. In questo tipo di bridge, la frame deve essere ricevuta per intero prima che i processi di filtering e forwarding possano essere iniziati. Ciò consente alla

frame di essere scartata alla porta di ricezione se essa fallisce il controllo CRC, impedendo il non necessario filtraggio e/o inoltro di frame errate.

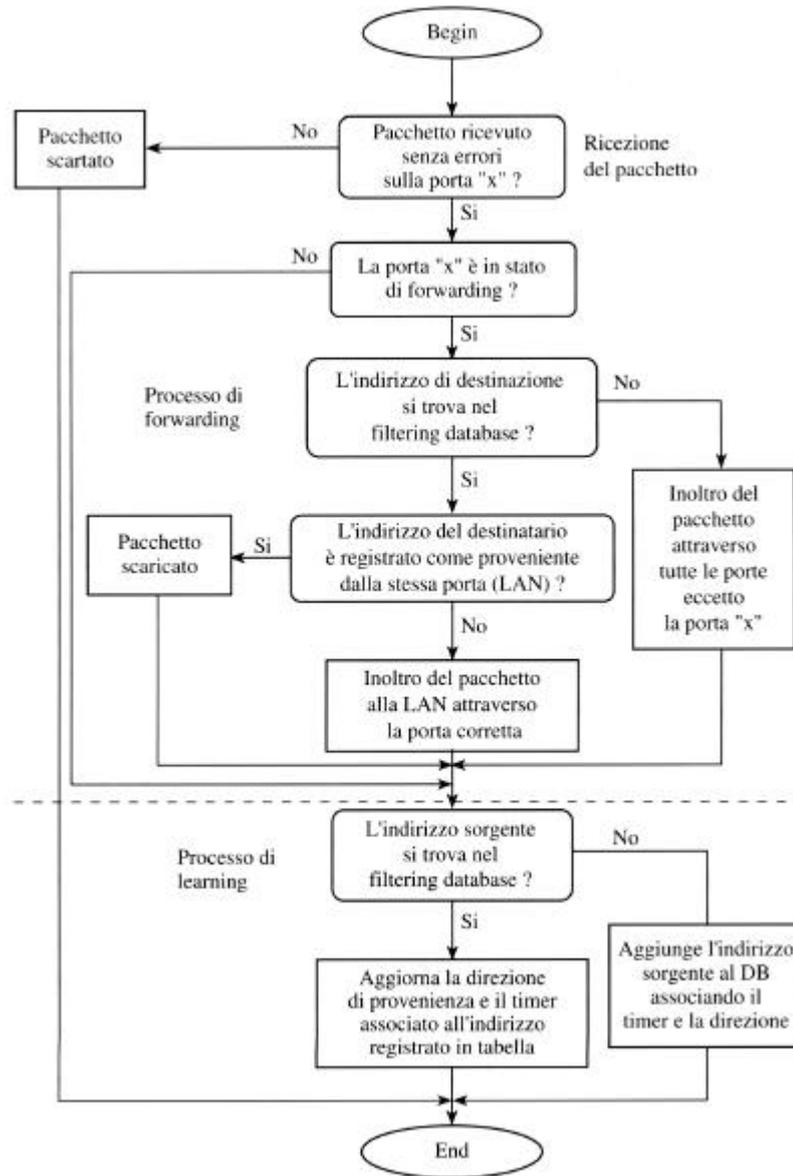


Fig. 1.18. - Diagramma di flusso delle principali operazioni di un bridge

1.7.7. Algoritmo di Spanning Tree

L'algoritmo di Spanning Tree è un protocollo *packet-based* che consente ai bridge di identificare dinamicamente la topologia della rete e configurarla per evitare che nella rete si creino dei percorsi chiusi (loop). Impedire la

formazione di loop, in modo che tra due nodi qualunque esista un cammino unico, è molto importante per diversi motivi, tra i quali due chiave sono:

- Un bridge farà il flooding di una frame su molte porte se essa contiene un indirizzo di destinazione sconosciuto o di tipo broadcast. Tale flooding può essere replicato su di un altro bridge, che può a sua volta replicarlo. Se questi bridge hanno più di un'interconnessione (per cui si crea un loop), tale situazione può perpetuarsi e ripetersi sino a consumare alla fine tutta la larghezza di banda disponibile di tutte le LAN.
- Nelle frame a livello MAC non c'è un contatore (*hop count*) che consenta di scartare una frame che continua a girare su di un loop.

Sostanzialmente, l'algoritmo di Spanning Tree è un algoritmo distribuito che opera nei seguenti tre passi:

- *Elezione del root bridge*: poiché si vuole identificare un albero, il primo passo consiste nell'identificare la radice dell'albero, cioè il *root bridge*, che, per definizione, è il bridge avente il minore indirizzo MAC a 48 bit.
- *Selezione della root port*: per ogni bridge si identifica la porta più "conveniente" per interconnettere il bridge verso il root bridge, la quale viene detta, appunto, *root port*.
- *Selezione del designated bridge*: per ogni LAN si sceglie quale bridge è designato a interconnettere la LAN con il root bridge. Questo passo è particolarmente importante quando esistono più cammini tra le LAN e il root bridge. Ogni LAN ha quindi un solo designated bridge che è il bridge "più vicino" al root bridge e che si incaricherà di trasmettere verso di esso i pacchetti. La porta del designated bridge che interconnette la LAN è detta *designated port*.
Il root bridge è l'unico bridge avente tutte designated port.

Al termine di questi tre passi si può procedere alla messa in stato di blocking delle porte che non sono né root né designated.

Ogni porta di un bridge ha un *path cost* che può essere configurato da management e che indica il costo di attraversamento di quella porta. Per ogni bridge e per ogni porta del bridge si definisce inoltre il *root path cost* come il costo totale di percorso per raggiungere il root bridge.

Il root path cost delle porte serve a due scopi:

- all'interno di un bridge, per scegliere quale sia la root port (quella che ha root path cost minore);
- tra le non root port dei bridge che si collegano su una LAN, per scegliere la designated port (quella che ha root path cost minore).

Nella figura seguente viene illustrato un esempio di applicazione dell'algoritmo di Spanning Tree ad una rete locale estesa con topologia a maglia. Nella fig. 1.19.a il bridge 1 è stato eletto come root bridge ed è quindi il designated bridge per la LAN 1 e la LAN 2, il bridge 2 è il designated bridge per la LAN 3 e la LAN 4 e il bridge 4 è il designated bridge per la LAN 5. Le porte del bridge 3 e del bridge 5 collegate alla LAN 5 sono state messe in stato di blocking. La fig. 1.19.b mostra lo Spanning Tree di questa configurazione attiva della rete locale estesa.

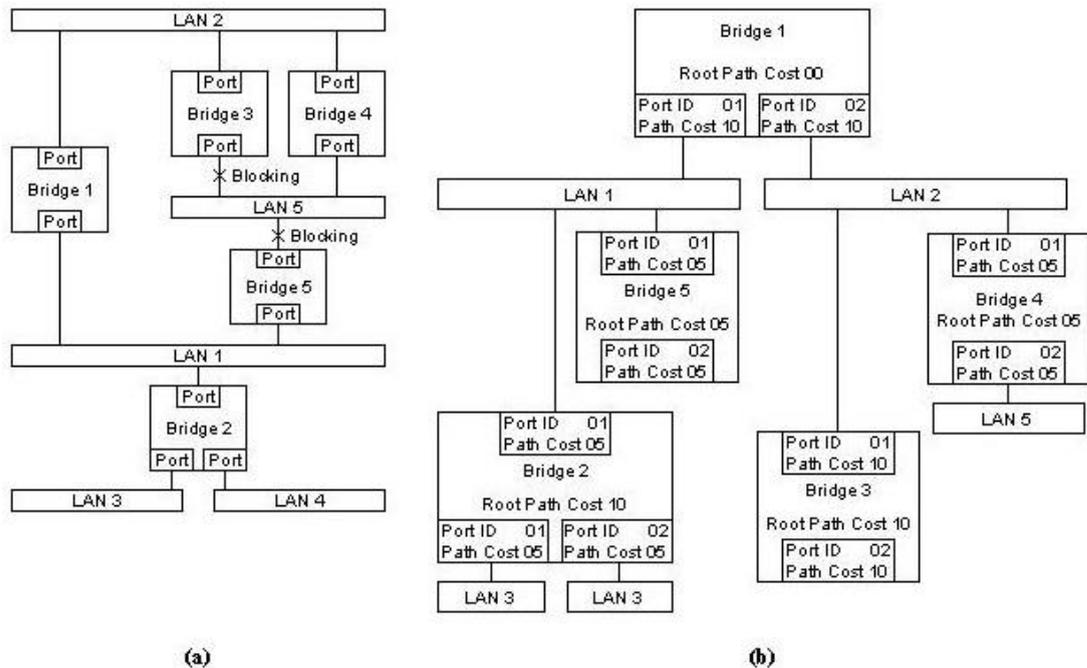


Fig. 1.19. - Esempio di SpanningTree: (a) Topologia attiva ad albero su di una LAN magliata; (b) spanning tree della configurazione attiva.

L'algoritmo di Spanning Tree viene eseguito periodicamente al fine di rilevare eventuali cambiamenti nella topologia della rete o l'aggiunta/rimozione di altri bridge trasparenti.

1.7.8. Prestazioni di un Bridge 802.3

Le prestazioni di un bridge sono importanti in quanto determinano le prestazioni della rete locale che esso suddivide. I parametri più importanti sono:

- numero massimo di pacchetti al secondo che un bridge può filtrare (cioè ricevere e passare);
- numero massimo di pacchetti al secondo che un bridge può ritrasmettere;
- *tempo medio di latenza*, cioè tempo di attraversamento del bridge da parte di un pacchetto.

Per minimizzare la possibilità di perdita di pacchetti è preferibile che un bridge sia *full-speed*, cioè che i primi due parametri siano uguali al massimo teorico. Questo è tanto più difficile da realizzare quanto più i pacchetti sono corti (massimizza il numero di decisioni sul filtraggio che un bridge deve prendere nell'unità di tempo), per cui la proprietà di full-speed deve essere verificata con pacchetti tutti di lunghezza minima.

Nel caso di 802.3, un bridge è full-speed quando è in grado di inoltrare 14880 PPS (Packet Per Second) da 64 byte (pacchetti più corti). Tale numero si può calcolare partendo dalla lunghezza del pacchetto, aggiungendo preambolo e delimitatori e considerando la necessità di rispettare l'Inter-Packet Gap tra i pacchetti, fissato, come già detto, da IEEE 802.3 in 9.6 μ s.

Il tempo di latenza di un bridge esprime il tempo che intercorre da quando il pacchetto incomincia ad entrare nella porta ricevente a quando esso

incomincia ad uscire dalla porta trasmittente. Il tempo di latenza non è fisso, ma varia a seconda della dimensione del pacchetto ricevuto: valori medi sono compresi tra 80 μ s e 1.2 ms.

La tabella seguente illustra le prestazioni di un bridge 802.3.

Dimensione pacchetto	Pacchetti al secondo	
	carico 50%	carico 100%
1518	403	812
1024	603	1206
512	1192	2385
256	2332	4664
128	4464	8928
64	8223	14880

Tab. 1.1. - Prestazioni di un bridge 802.3-802.3.

Capitolo 2

Fast Ethernet

Quasi sin dalla pubblicazione dell'IEEE 802.3, si sentì il bisogno di una Ethernet a velocità superiori ai 10 Mbps, a causa dell'aumento di richiesta di banda di rete da parte delle varie applicazioni e del numero del numero di stazioni collegate ad una LAN. Di conseguenza, ben presto l'attenzione si concentrò su come incrementare le performance complessive di Ethernet, prima con la versione a 100 Mbps, nota come Fast Ethernet, e più recentemente con lo sviluppo dello standard Gigabit Ethernet.

Nel 1992, infatti, l'IEEE riconvocò il comitato 802.3, con lo scopo di creare una LAN più veloce. Vennero presentate due proposte:

- una, da parte di HP e AT&T, prevedeva di ridisegnare completamente l'802.3, basandosi su un nuovo metodo di accesso detto Demand Priorità, per introdurre molte nuove caratteristiche come quelle per il traffico real-time e voce digitalizzata;
- l'altra, appoggiata tra gli altri da 3Com, Intel e Grand Junction Networks, proponeva di lasciare inalterata 802.3 e di renderla semplicemente più veloce.

Tra le due proposte, il comitato decise di scegliere la seconda, mentre i sostenitori della prima proposta formarono un nuovo comitato e standardizzarono la loro LAN (802.12 o 100VG AnyLAN).

Il lavoro del comitato 802.3 fu completato velocemente e il risultato, **802.3u** o, più comunemente, Fast Ethernet, venne ufficialmente approvato dall'IEEE nel giugno del 1995.

L'idea alla base di Fast Ethernet è quella di mantenere inalterati il protocollo di accesso al mezzo CSMA/CD, il formato della frame, le interfacce e le regole procedurali (in modo da non richiedere cambiamenti software ai livelli superiori) e il cablaggio esistente, riducendo, però, il *bit time* da 100 ns a 10 ns.

Dal momento che la dimensione minima del pacchetto non è stata alterata, si è dovuto ridurre di un fattore 10 il round trip delay e quindi il diametro della rete. Dal momento che già 10BaseT presenta un diametro della rete di 200m, la topologia adottata da Fast Ethernet è stata quella a stella, per cui lo standard è anche noto come **100BaseT**.

Lo standard 100BaseT è formato da cinque specifiche componenti. Queste definiscono il Media Access Control (MAC) layer, il Media Independent Interface (MII) layer e i tre physical layer (100BaseTX, 100BaseT4 e 100BaseFX), come illustrato nella figura seguente.

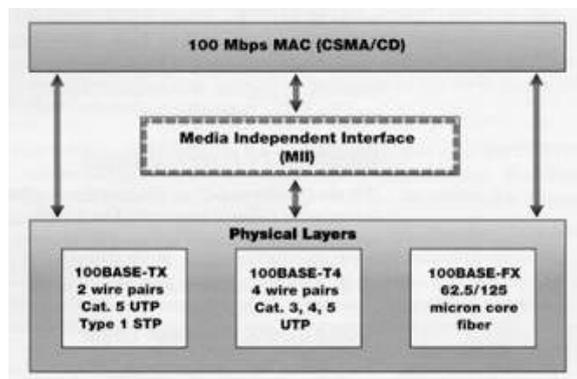


Fig. 2.1. - Architettura del protocollo di 100BaseT

Daremo ora, nei paragrafi seguenti, qualche cenno alle somiglianze e alle differenze tra Ethernet e Fast Ethernet.

2.1. Architettura di 100BaseT

Lo standard 100BaseT definisce un nuovo sottostrato, posto inferiormente al MAC, detto *Reconciliation Sublayer (RS)*, il quale fornisce la funzione di traduzione dei segnali a livello MII in primitive di servizio PLS (*Physical*

Layer Signaling). Quest'ultimo, lo ricordiamo, è un sottostrato del Physical Layer del modello OSI ed è responsabile della codifica/decodifica dei dati in fase di trasmissione e di ricezione, collocato, in 10BaseT tra la AUI ed il MAC (figura 1.2)

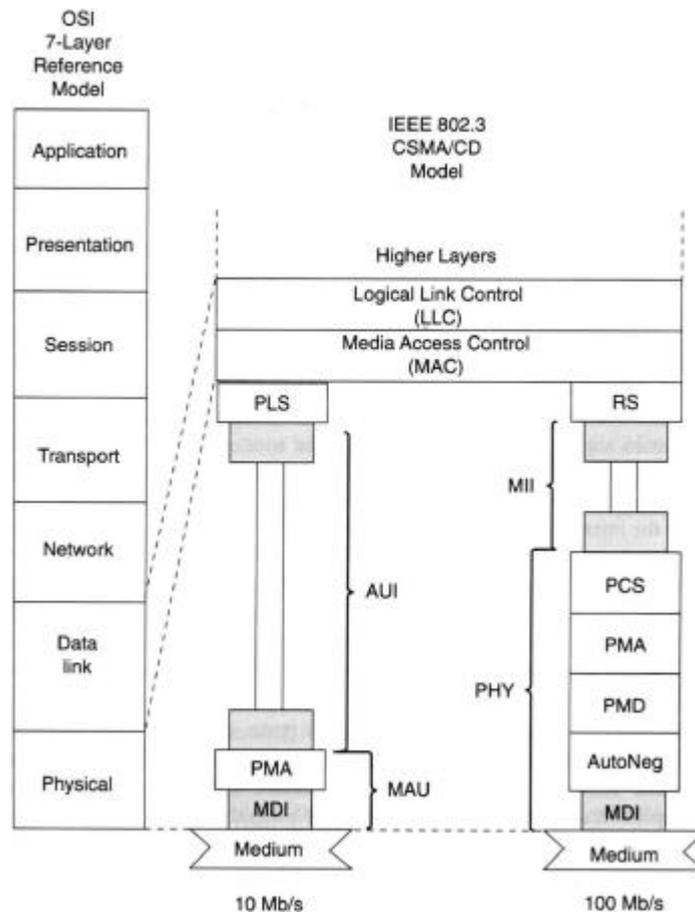


Fig. 2.2. - 100BaseT: relazione con il modello di riferimento ISO/OSI.

Il MAC layer è basato sullo stesso protocollo CSMA/CD di Ethernet, con la sola differenza che è 10 volte più veloce. I dati possono muoversi facilmente da Ethernet a Fast Ethernet senza richiedere traduzioni di protocollo, il che consente agli addetti ai lavori di non dovere imparare una tecnologia completamente nuova.

La Medium Independent Interface (MII) fornisce un'interconnessione semplice ed economica tra il MAC e i diversi sottostrati fisici (PHY), per cui è logicamente equivalente all'Attachment Unit Interface (AUI) di Ethernet.

I dati sono trasferiti attraverso la MII alla velocità di un *nibble* (4 bit) per ogni ciclo di clock, per cui i clock di trasmissione e di ricezione operano ad un quarto del data rate di 100 Mbps di 100BaseT. Dal momento che l'interfaccia supporta entrambe le velocità 10 Mbps e 100 Mbps, i clock operano a 2.5 o 25 MHz, rispettivamente.

Va sottolineato che a 100 Mbps non è possibile utilizzare codifiche come la Manchester, dal momento che l'elevato clock rate risultante violerebbe i limiti di 30 MHz fissati per l'uso su cavi UTP. Invece, vengono adoperati schemi di codifica dei bit che utilizzano uno o più gruppi di quattro bit per codificare ciascun simbolo, in modo da assicurare che ogni simbolo codificato abbia sufficienti transizioni al suo interno da abilitare il ricevitore a mantenere la sincronizzazione del clock.

Il sottolivello *Physical Layer Device* (PHY) è la porzione del livello fisico tra l'MDI (*Medium Dependent Interface*) e la MII che comprende i sottostrati *Physical Coding Sublayer* (PCS), *Physical Medium Attachment* (PMA) e, se presente, *Physical Medium Dependent*(PMD).

Il PCS è usato in 100BaseT per accoppiare la MII con il PMA; esso, infatti, contiene le funzioni per codificare i bit di dati in *code groups* (insiemi di sei simboli ternari per la variante T4; insiemi di cinque bit di codice per le varianti TX e FX) che possono essere trasmessi sul mezzo fisico. Sono definite due strutture PCS per 100BaseT: una per 100BaseX che usa la codifica 4B5B per generare un flusso full duplex a 125 Mbps, e una per 100BaseT4 che codifica i nibble di dati ricevuti dalla MII in *code groups* di tipo 6T usando la codifica 8B6T.

Il sottolivello PMA è la porzione del livello fisico che contiene le funzioni per la trasmissione, ricezione, clock recovery e skew alignment.

Il *Physical Medium Dependent* (PMD) è la porzione del livello fisico responsabile dell'interfaccia con il mezzo trasmissivo.

La *Medium Dependent Interface* (MDI) è l'interfaccia meccanica ed elettrica tra il mezzo trasmissivo e il PMA.

Consideriamo ora un po' più in dettaglio i tre physical layer specificati dallo standard 802.3u, vale a dire: 100BaseT4, 100BaseTX e 100BaseFX².

2.2. 100BaseT4

Il physical layer 100BaseT4 definisce la specifica per il 100BaseT per la trasmissione su doppino a quattro coppie di categoria 3, 4 o 5. Il collegamento tra stazione e hub usa, delle quattro coppie, due in modalità half duplex, cioè alternativamente in trasmissione o in ricezione, una sempre in trasmissione ed una sempre in ricezione (vedi figura seguente). La trasmissione avviene quindi su tre coppie contemporaneamente, mentre la quarta coppia, in ricezione, serve per permettere all'interfaccia fisica di rilevare la presenza di collisioni senza dovere introdurre complicazioni aggiuntive al protocollo MAC rispetto allo standard IEEE

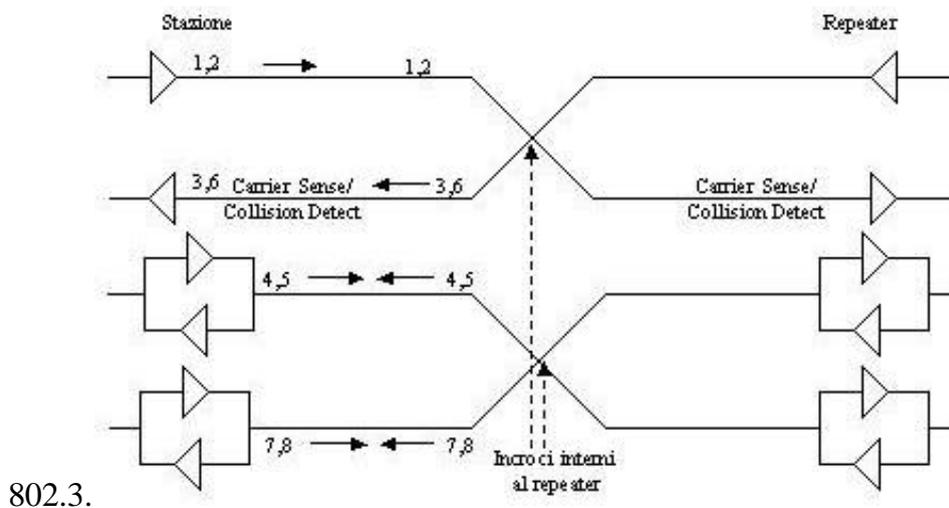


Fig. 2.3. - 100BaseT4: uso delle coppie.

Per trasmettere i pacchetti su tre coppie si utilizza una codifica di tipo **8B6T**, la quale suddivide un flusso binario a 100Mbps in tre flussi da 25 Mbaud (simboli, in questo caso, ternari al secondo). Infatti, trasmettere 100 Mbps divisi su tre canali significa trasmettere $100/3= 33.3$ Mbps su ogni canale, e trasformare ogni ottetto in sei simboli ternari significa associare ad ogni

² Il duo TX/FX viene anche chiamato 100BaseX

simbolo un'informazione pari a 8/6 di bit; quindi su ogni canale è necessario trasmettere

$$\frac{100 \times 6 / 8}{3} = 25 \text{ MHz (o Mbaud)}$$

che è ben al di sotto dei limiti fissati per l'UTP.

In questo schema trasmissivo partendo da 8 bit, che consentono 256 possibili combinazioni diverse di valori, si codificano 6 simboli ternari, che consentono di rappresentare 729 valori, e questo permette l'introduzione di informazioni aggiuntive per la sincronizzazione del flusso informativo e per il controllo degli errori.

2.3. 100BaseTX

Il physical layer 100BaseTX definisce la specifica per il 100BaseT per la trasmissione su due doppini di categoria 5 UTP o STP (Shielded Twisted Pair, doppino schermato). Con una coppia per la trasmissione e una per la ricezione, lo schema di cablaggio è identico a quello usato per il 10BaseT, così come identico è il connettore UTP usato, un RJ-45, cablato nella stessa maniera. Di conseguenza, 100BaseTX è un sistema full duplex: le stazioni possono trasmettere a 100 Mbps e ricevere a 100 Mbps nello stesso momento.

Dal momento che i fili usati possono sostenere un clock rate di 125 MHz e oltre, al posto della codifica binaria viene usata la tecnica chiamata **4B5B** a 125 MHz, già definita per FDDI, la quale utilizza cinque bit per spedire ogni possibile sequenza di quattro bit. Nella tabella 2.1 è riportata la codifica dei principali simboli MAC in 4B5B.

<i>Valore</i>	<i>Simbolo</i>	<i>Assegnazione</i>
00000	Q	stato di linea Quiet
11111	I	stato di linea Idle
00100	H	stato di linea Halt
11000	J	prima parte dello start delimiter
10001	K	seconda parte dello start delimiter
11110	0	quartetto di valore 0

01001	1	quartetto di valore 1
10100	2	quartetto di valore 2
10101	3	quartetto di valore 3
01010	4	quartetto di valore 4
01011	5	quartetto di valore 5
01110	6	quartetto di valore 6
01111	7	quartetto di valore 7
10010	8	quartetto di valore 8
10011	9	quartetto di valore 9
10110	A	quartetto di valore A
10111	B	quartetto di valore B
11010	C	quartetto di valore C
11011	D	quartetto di valore D
11100	E	quartetto di valore E
11101	F	quartetto di valore F
01101	T	simbolo di terminazione
00111	R	zero logico (reset)
11001	S	uno logico (set)

Tab. 2.1. Codifica dei simboli MAC in 4B5B.

Come si può osservare, la codifica 4B5B comprende tutti i possibili quartetti in ingresso, più alcuni simboli aggiuntivi per la gestione di un protocollo a livello fisico. L'overhead introdotto è pari ad un bit ogni quattro, cioè il 25%. Una trasmissione a 100 Mbps al livello MAC, quindi, invierà sul mezzo fisico 125 Mbps. I vantaggi della 4B5B sono essenzialmente che con essa:

- viene fornita un po' di ridondanza;
- vengono fornite abbastanza transizioni per permettere la sincronizzazione dei clock;
- vengono creati standard unici per i delimitatori di frame;
- viene mantenuta la compatibilità con FDDI a livello fisico.

2.4. 100BaseFX

Il physical layer 100BaseFX definisce la specifica per il 100BaseT su due cavi di fibra ottica multimodale, uno per direzione: è quindi anch'esso full duplex con 100 Mbps per direzione ed utilizza la stessa tecnica di codifica 4B5B di 100BaseTX.

Il cavo in fibra ottica, i convertitori ottico/elettrici e i connettori sono sostanzialmente più costosi e più difficili da installare rispetto agli analoghi mezzi fisici di rame. Di conseguenza, la fibra è in genere utilizzata solo quando necessario – tipicamente in applicazioni su lunghe distanze e ad ampia banda o per connettere LAN distanti geograficamente.

Il funzionamento full duplex consente alle distanze tra gli estremi dei link di essere estese considerevolmente fino ad un massimo di 2 km nel caso di fibra multimodale. L'uso della fibra a singolo modo è supportato da alcuni venditori per aumentare sostanzialmente la distanza anche al di là dei 2 km, ancora utilizzando il full duplex. Va comunque ricordato che, dal momento che i ripetitori non possono operare in maniera full duplex, una rete contenente ripetitori non potrebbe in ogni caso supportare tali distanze.

2.5. Diametro della rete

Il diametro della rete, ovvero la distanza massima tra due *end station* (due stazioni o una stazione e uno switch, bridge o router) sulla stessa LAN o dominio di collisione, è la principale differenza tra la tradizionale Ethernet e Fast Ethernet. La distanza fisica è limitata dal massimo round trip delay consentito dalla tecnologia e dal tipo di mezzo fisico adottato. A causa della maggiore velocità e del tipo di cablaggio, il massimo diametro di una Fast Ethernet basata su twisted-pair (100BaseTX o 100BaseT4) è di 205 m, laddove per una Ethernet basata su 10BaseT il diametro è di 500 m. Infatti, mentre nel caso di una rete basata su 10BaseT tra due stazioni poste ai suoi estremi ci possono essere fino a cinque segmenti e quattro ripetitori, nel caso di Fast Ethernet il numero di ripetitori consentiti è ridotto a due o a uno, a seconda della classe del ripetitore.

2.6. Ripetitori a 100Mbps

Al contrario dei ripetitori 10BaseT, che sono considerati essere tutti identici dal punto di vista funzionale, i ripetitori Fast Ethernet sono invece divisi in due classi distinte: Classe I e Classe II.

- Un ripetitore di Classe I presenta un maggiore ritardo temporale interno di uno di Classe II, e opera traducendo in forma digitale i segnali di linea in ingresso ad una sua porta e poi ritraducendoli in segnali di linea, quando deve inviarli sulle altre porte. Ciò consente di utilizzare un tale ripetitore per collegare sullo stesso dominio di collisione mezzi fisici differenti, come, per esempio, segmenti 100BaseTX a segmenti 100BaseFX. Questo ritardo maggiore è anche necessario per ripetitori 100BaseT4 e per modelli di tipo *stackable*³.
- Un ripetitore di Classe II ha un più piccolo ritardo interno, poiché esso tipicamente collega mezzi fisici identici sullo stesso dominio di collisione (per esempio, 100BaseTX a 100BaseTX).

In base allo standard, in una rete di dimensioni massime possono esistere solo un ripetitore di Classe I oppure fino a due ripetitori di Classe II.

2.7. Funzionamento full duplex

Un link punto-punto operante in modalità full duplex consente, come già visto, la simultanea trasmissione e ricezione dei dati, essendo costituito da due canali classici (half duplex) in parallelo, ciascuno monodirezionale. In tal modo, il link è non soggetto a collisione (poiché in ogni direzione c'è una sola stazione che può trasmettere e quindi per definizione non può collidere con nessun'altra) e quindi i limiti di distanza non sono più dettati dal livello MAC

³ Gli hub stackable sono hub "impilati" l'uno sull'altro su un unico stack, apparendo, in ogni caso, al resto della rete come un unico ripetitore logico. Tutto ciò consente di raggruppare più utenti.

ma solo dal livello fisico. Il full duplex è stato standardizzato dal comitato 802.3x, parallelamente allo sviluppo di 100BaseT, ed è utilizzabile sia in associazione con lo standard 10BaseT che al 100BaseT (e, vedremo pure, con Gigabit Ethernet). Le distanze massime ammesse sono tipicamente di 100 m su cavo UTP, 2 Km su fibra ottica multimodale e 50 Km su fibra ottica monomodale.

Un MAC operante in modalità full duplex richiede le seguenti variazioni rispetto ad un MAC CSMA/CD:

- La trasmissione di una frame non viene rinviata durante la fase di ricezione.
- Il conteggio dell'Inter-Packet Gap per la trasmissione di un pacchetto viene iniziato alla fine della trasmissione del pacchetto precedente, laddove nella modalità half duplex esso inizia nel momento in cui la stazione che desidera trasmettere trova il canale libero (vale a dire quando rileva il segnale di Carrier Sense a OFF).
- L'indicazione di una collisione nella modalità full duplex viene ignorata.

2.8. Auto-Negoziazione

Le possibili modalità di funzionamento di una scheda 100BaseT previste nello standard sono: 100BaseT4, 100BaseX full o half duplex, 10BaseT full o half duplex. La modalità full duplex è interessante per il collegamento tra switch in quanto consente di realizzare collegamenti dedicati a 20 o 200 Mbps.

La specifica 100BaseT descrive un processo di negoziazione che consente ai dispositivi posti ai due capi di un link di rete di scambiarsi automaticamente delle informazioni circa le loro caratteristiche ("duplex mode" e velocità trasmissiva) e di realizzare la configurazione necessaria perché entrambi operino al massimo livello comune. Per esempio, l'auto-negoziazione consente ad un adattatore 10/100 (cioè con possibilità di funzionamento a 10 e

a 100 Mbps) di funzionare in modalità 10BaseT se connesso ad un hub o switch 10BaseT, o in modalità 100BaseT se connesso ad un hub o switch 100BaseT.

L'Auto-Negoziazione è effettuata fuori banda usando una sequenza modificata dei segnali di test del collegamento detti *link integrity test pulses* usati in 10BaseT. L'informazione è trasmessa all'interno di un burst di link integrity test pulse vicini, chiamato un *Fast Link Pulse* (FLP), che viene generato all'accensione. I dati estratti dall'FLP informano la stazione ricevente delle capacità del trasmettitore all'altro capo del segmento. Se vengono rilevati degli FLP, l'algoritmo di auto-negoziazione determina il modo di funzionamento con le massime performance comuni e aggiorna entrambe le estremità del link.

Capitolo 3

Gigabit Ethernet

3.1. Lo standard 802.3z

Gigabit Ethernet rappresenta la naturale evoluzione dello standard Ethernet 10BaseT e di quelli Fast Ethernet, 100BaseTX e 100BaseFX, portando la velocità di trasmissione a 1Gbps.

Già nel 1995, infatti, l'idea di fare andare Ethernet a velocità dell'ordine del gigabit iniziò ad acquistare consensi, spingendo alla formazione di un gruppo di studio all'interno del comitato 802.3. Lo standard Gigabit Ethernet è stato, così, inizialmente approvato nel 1998 per la fibra ottica e il rame con il nome di IEEE **802.3z** e successivamente nel 1999 per l'UTP di categoria 5 con il nome di IEEE **802.3ab**, con i seguenti obiettivi:

- consentire il funzionamento half e full duplex a 1000 Mbps;
- mantenere lo stesso formato della frame Ethernet;
- utilizzare lo stesso protocollo MAC CSMA/CD;
- mantenere la compatibilità con le tecnologie 10BaseT e 100BaseT.

Nei successivi paragrafi, verrà illustrata la tecnologia Gigabit Ethernet, con particolare attenzione alle caratteristiche principali dei livelli fisico e data link e dell'interfaccia definita tra di essi.

3.2. GMII (Gigabit Media Independent Interface)

La figura seguente mostra la relazione tra il modello di riferimento OSI e il modello a strati definito per Gigabit Ethernet.

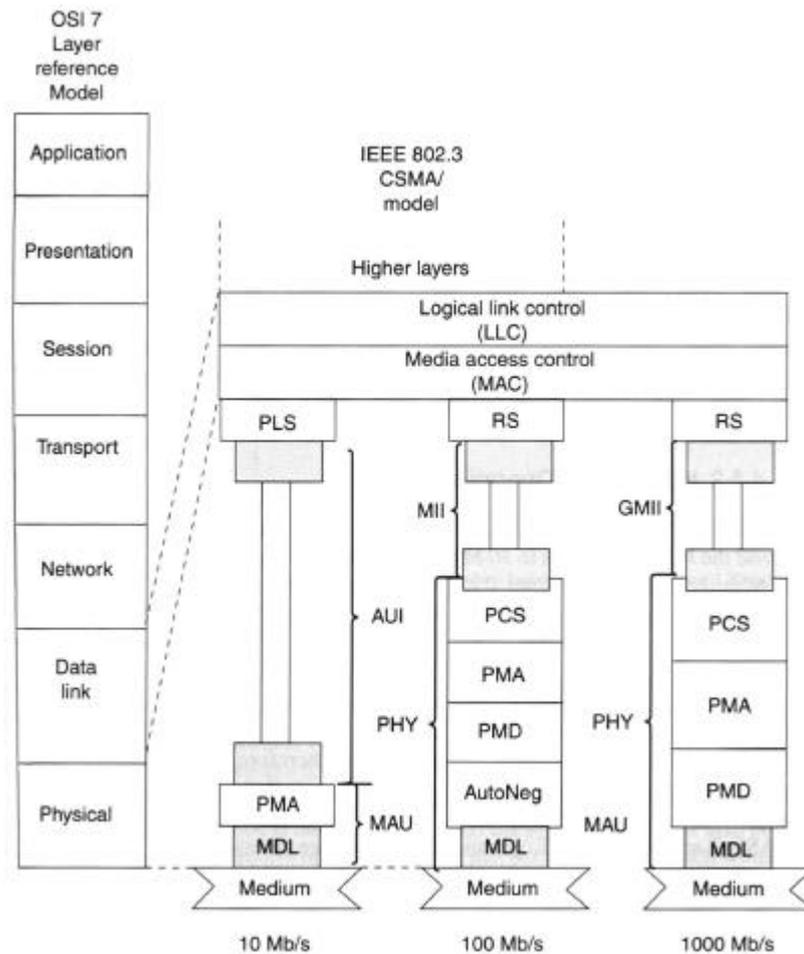


Fig. 3.1. - Gigabit Ethernet: relazione con il modello di riferimento ISO/OSI

La GMII è l'interfaccia tra il MAC layer ed il physical layer e presenta le seguenti caratteristiche:

- consente a ciascun mezzo fisico di essere utilizzato con il MAC layer;
- è un'estensione del MII usato in Fast Ethernet;
- supporta le velocità di trasmissione dai 10 ai 1000 Mbps;
- supporta sia il full duplex che l'half duplex.

La GMII fornisce due segnali di stato del mezzo: uno indica la presenza della portante (Carrier Sense) e l'altro indica l'assenza di collisioni (Collision

Detection). Il Reconciliation Sublayer (RS) mappa questi segnali nelle primitive di Physical Signalling (PLS) comprese dal sottolivello MAC esistente. Con la GMII è possibile collegare diversi mezzi fisici, come l'UTP e l'STP, la fibra ottica a modo singolo o multimodale, utilizzando lo stesso controller MAC.

3.3. Physical Layer

Il Physical Layer in Gigabit Ethernet è diviso in tre sottolivelli:

- il PCS (Physical Coding Sublayer);
- il PMA (Physical Medium Attachment);
- il PMD (Physical Medium Dependent).

Il **PCS** è il sottolivello del Physical Layer che fornisce un'interfaccia uniforme al Reconciliation Sublayer per tutti i mezzi fisici. Esso utilizza la codifica 8B/10B del Fiber Channel, nella quale gruppi di 8 bit sono rappresentati da gruppi di 10 bit (*code groups*), alcuni dei quali rappresentano appunto dati di 8 bit, mentre altri rappresentano simboli di controllo (come i simboli di estensione usati nel Carrier Extension). Il PCS genera i segnali di Carrier Sense e Collision Detection e inoltre gestisce il processo di auto-negoiazione tramite il quale il NIC (Network Interface Controller) comunica con la rete per determinare la sua velocità (10, 100 o 1000 Mbps) e la sua modalità di funzionamento (half duplex o full duplex).

Il **PMA** consente al PCS di supportare i diversi mezzi fisici "serial bit-oriented": esso infatti serializza i code groups e deserializza i bit ricevuti dal mezzo fisico in code groups.

Il **PMD** "mappa" il mezzo fisico sul PCS, definendo le codifiche dei segnali utilizzate dai diversi mezzi. Esso contiene il MDI (Medium Dependent Interface), che è l'effettiva interfaccia a livello fisico e definisce l'effettiva modalità di connessione, come i connettori, per i diversi tipi di mezzo fisico.

A livello fisico Gigabit Ethernet supporta quattro differenti mezzi di trasmissione, definiti in 802.3z (1000BaseX) e in 802.3ab (1000BaseT) e che andiamo rapidamente a illustrare.

3.3.1. 1000BaseX

Lo standard 1000BaseX è basato sul physical layer del Fiber Channel (canale a fibra). Il Fiber Channel è una tecnologia di interconnessione per collegare workstation, supercomputer, dispositivi di memorizzazione e periferiche, caratterizzata da un'architettura a 4 livelli. I due livelli più bassi FC-0 (che ha a che fare con il livello fisico) e FC-1 (che gestisce la codifica/decodifica dei bit), insieme allo schema di codifica 8B/10B, vengono utilizzati in Gigabit Ethernet (il che ha ridotto i tempi di sviluppo dello standard Gigabit Ethernet, essendo Fiber Channel una tecnologia già ben collaudata).

Tre tipi di mezzi fisici sono inclusi nello standard 1000BaseX:

- 1000BaseSX: viene utilizzata luce laser, con lunghezza d'onda pari a 850nm, su un cavo di fibra ottica multimodale;
- 1000BaseLX: viene utilizzata luce laser, con lunghezza d'onda pari a 1300nm, su un cavo di fibra ottica a modo singolo oppure multimodale;
- 1000BaseCX: viene utilizzato un cavo in rame STP (Shielded Twisted Pair) da 150 Ω e consente una distanza massima pari a 25 m.

Le distanze massime imposte dallo standard 802.3z per la fibra, a seconda del tipo e del diametro del nucleo, sono riassunte nella seguente tabella:

Standard	Tipo di fibra	Diametro	Distanza
1000BaseSX	multimodale	62.5 μm	275 m
	multimodale	50 μm	550 m
1000BaseLX	multimodale	62.5 μm	550 m
	multimodale	50 μm	550 m
	modo singolo	9 μm	5 Km

Tab. 3.1. - Distanze massime imposte da 802.3z per la fibra

3.3.2. 1000BaseT

1000BaseT è lo standard per Gigabit Ethernet per il rame e utilizza per la trasmissione quattro doppini UTP di categoria 5. Esso non utilizza lo schema di codifica 8B/10B di Fiber Channel e consente una distanza massima di 100 m.

3.4. MAC Layer

Nella tabella 3.1 sono mostrati i parametri di Gigabit Ethernet: essendo il valore del bit time a 1000 Mbps pari a 1 ns, l'Inter-Packet Gap per il Gigabit è un decimo di quello di Fast Ethernet.

Tutti gli altri parametri sono identici a quelli di Fast Ethernet, tranne lo slot time che, come vedremo, il cui valore è stato accresciuto a 4096 bit.

<i>Parametro</i>	<i>Valore</i>
slot time	4096 bit time
IPG	0.096 μs
Attempt Limit	16
Backoff Limit	10
jamsize	32 bit
min frame size	64 byte
max frame size	1518 byte

Tab. 3.2. - Principali parametri di Gigabit Ethernet.

Quando opera in maniera half duplex, Gigabit Ethernet adopera, come già detto, lo stesso protocollo CSMA/CD di Ethernet per risolvere la contesa sul mezzo condiviso. Il problema che però viene a presentarsi è dovuto al legame tra velocità di trasmissione, lunghezza minima della frame e lunghezza massima del cavo. Si è già visto come in Fast Ethernet, allo scopo di rispettare i vincoli di round trip delay, la lunghezza massima del cavo sia stata ridotta a 100 m, lasciando invariati la lunghezza minima della frame e lo slot time (512 bit time).

In Gigabit Ethernet vengono mantenute le stesse dimensioni minima e massima della frame di Ethernet, ma, essendo la velocità di trasmissione dieci volte maggiore di quella di Fast Ethernet, per mantenere lo stesso slot time la lunghezza massima del cavo dovrebbe essere ridotta a circa 10 m, il che sarebbe molto limitante.

Per risolvere questo problema ed assicurare comunque la compatibilità con Ethernet, lo slot size (vale a dire il numero di byte che possono essere trasmessi in uno slot time) viene portato da 64 a 512 byte, non alterando la dimensione minima della frame. Viene, quindi, esteso il cosiddetto “carrier event”, per cui se la frame è più corta di 512 byte essa viene “imbottita” con dei simboli speciali di estensione (trasmessi dopo il FCS), che, non contenendo dati, non contribuiscono al carico della rete. Tale processo prende il nome di *Carrier Extension*.

3.4.1. Carrier Extension

Il Carrier Extension consente, quindi, di mantenere le dimensioni minima e massima della frame anche per lunghezze dei cavi significative, permettendo così l'interoperabilità di Gigabit Ethernet con le esistenti reti 802.3.

Nelle frame “imbottite” i simboli di estensione sono inclusi sempre all'interno della *collision window*, per cui ai fini della collisione viene considerata l'intera frame imbottita (o *padded*). Comunque, il Frame Check

Sequence viene calcolato solo sulla frame originaria (senza simboli di estensione). I simboli di estensione vengono, infatti, rimossi prima che il FCS sia controllato dal ricevitore, per cui il LLC (Logical Link Control) layer non si rende conto del carrier extension. Nella figura seguente viene illustrato il formato della frame quando viene utilizzato il Carrier Extension.

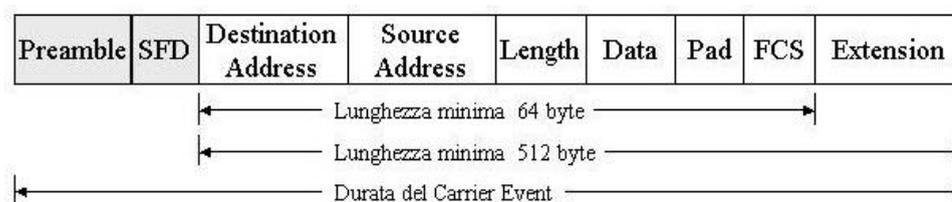


Fig. 3.2. - Formato della Frame Ethernet con il Carrier Extension

Sebbene il Carrier Extension risolva in maniera semplice il problema di mantenere lo stesso diametro di rete di 100BaseT, senza aumentare la dimensione minima della frame, esso aumenta la durata del tempo di trasmissione per le frame corte. In generale, infatti, per una frame da 64 byte, il tempo di trasmissione viene aumentato di otto volte (dato che vengono aggiunti 448 byte padding) sebbene i bit vengano trasmessi dieci volte più velocemente! Pertanto, in una rete consistente di sole frame di 64 byte, l'effettivo throughput ottenuto è solo il 25%. Comunque sia, il traffico di rete raramente consiste solo di frame corte, per cui l'overhead per frame introdotto è generalmente più piccolo. Una tecnica chiamata *Packet Bursting*, descritta nel successivo paragrafo, è stata sviluppata per risolvere il problema di migliorare le prestazioni in una Gigabit Ethernet condivisa, specie se il traffico di rete consiste di frame corte.

3.4.2. Packet Bursting

Il *Packet Bursting* è un'estensione del Carrier Extension, e più precisamente è un "Carrier Extension più un burst di pacchetti". Quando un DTE ha un certo numero di pacchetti da trasmettere, il primo pacchetto viene, se necessario, imbottito utilizzando il carrier extension. Se la trasmissione è stata

completata con successo (vale a dire senza collisione), gli eventuali successivi pacchetti vengono poi trasmessi uno dopo l'altro, distanziati del minimo Inter-Packet Gap, fino alla scadenza di un cosiddetto *burst timer* di 1500 byte. Occorre notare che i successivi pacchetti all'interno del burst non vengono imbottiti, poiché lo slot time è già passato e il DTE ha il controllo del mezzo. Una volta che il burst termina, il DTE contenderà un'altra volta per il mezzo.

In tal modo, il Packet Bursting aumenta sostanzialmente il throughput, consentendo a server, switch e altri dispositivi di inviare burst di piccoli pacchetti in modo da sfruttare pienamente tutta la banda disponibile.

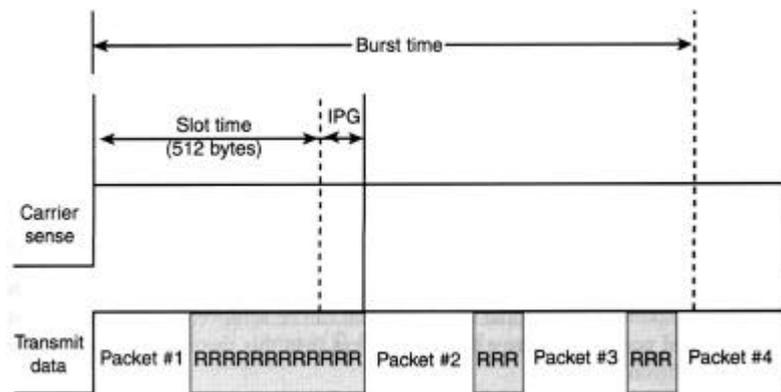


Fig. 3.3. - Packet Bursting.

Va comunque sottolineato il fatto che quando i dispositivi di rete operano in modalità full duplex, non è necessario estendere lo slot time, per cui essi non devono effettuare né il carrier extension né il packet bursting. I dispositivi full duplex, infatti, continueranno ad utilizzare il solito Inter Frame Gap (IFG) di 96 bit e la dimensione minima di 64 byte del pacchetto, così come previsto da Ethernet.