

Dispensa del corso di  
“Reti di Telecomunicazioni”

**N. Blefari-Melazzi**  
**M. Listanti**  
**A. Roveri**

# **RETI PER APPLICAZIONI TELEMATICHE**

Versione 2.0 (Dicembre 1998)  
VERSIONE RIDOTTA

A.A. 1998-1999

# INDICE

<b>I</b>	<b>TELEMATICA .....</b>	<b>7</b>
I.1	INFRASTRUTTURE.....	9
I.1.1	WAN per dati di prima generazione.....	9
I.1.2	Reti integrate nei servizi.....	12
I.1.3	Rete a larga banda integrata nei servizi (B-ISDN).....	14
I.1.4	Reti in area locale e metropolitana.....	18
I.1.5	Internet.....	21
I.2	ARCHITETTURE PER APPLICAZIONI TELEMATICHE .....	26
I.2.1	Il modello OSI.....	28
I.2.2	Il modello Internet.....	31
I.3	LE APPLICAZIONI E I SERVIZI OSI.....	31
I.3.1	Applicazioni telematiche.....	31
I.3.2	I servizi telematici.....	34
I.4	LE APPLICAZIONI INTERNET.....	36
<b>II</b>	<b>STRATO FISICO .....</b>	<b>45</b>
II.1	CARATTERISTICHE DEL SERVIZIO.....	45
II.2	FUNZIONALITÀ DEL SERVIZIO.....	52
II.3	FUNZIONE MODEM .....	56
II.3.1	Trasmissione in banda base.....	60
II.3.2	Trasmissione in banda traslata.....	78
II.3.3	Condizione di sincronismo.....	86
II.4	PROTOCOLLI DI STRATO FISICO .....	88
II.4.1	Caratteristiche meccaniche.....	89
II.4.2	Caratteristiche elettriche.....	91
II.4.3	Caratteristiche funzionali.....	95
II.4.4	Caratteristiche procedurali.....	100
<b>III</b>	<b>STRATO DI COLLEGAMENTO.....</b>	<b>115</b>
III.1	CARATTERISTICHE DEL SERVIZIO.....	115
III.2	FUNZIONALITÀ DEL SERVIZIO.....	118
III.2.1	Funzionalità del DL-servizio con modalità orientata alla connessione. ...	118

III.2.2	<i>Funzionalità del DL-servizio con modalità senza connessione.</i>	127
III.3	PROTOCOLLI DI STRATO DI COLLEGAMENTO	130
III.3.1	<i>Delimitazione delle trame.</i>	134
III.3.2	<i>Rivelazione degli errori.</i>	138
III.3.3	<i>Recupero in caso di errori.</i>	142
III.3.4	<i>Controllo di flusso.</i>	152
III.3.5	<i>Gestione del collegamento.</i>	152
III.4	IL LAP-B	158
<b>IV</b>	<b>STRATO DI RETE</b>	<b>165</b>
IV.1	CARATTERISTICHE DEL SERVIZIO	165
IV.2	CONTROLLO DEL TRAFFICO	166
IV.3	INDIRIZZAMENTO	168
IV.4	INSTRADAMENTO	170
IV.4.1	<i>Generalità sugli algoritmi di instradamento.</i>	171
IV.4.2	<i>Procedure basate sulla ricerca del percorso più breve</i>	176
IV.4.3	<i>Instradamento ottimo</i>	179
IV.4.4	<i>Instradamento a deflessione</i>	181
IV.4.5	<i>Instradamento diffusivo</i>	182
IV.4.6	<i>Instradamento in reti interconnesse</i>	184
IV.5	FUNZIONALITÀ DEL SERVIZIO	187
IV.6	PROTOCOLLO X.25 DI LIVELLO 3	193
<b>V</b>	<b>STRATO DI TRASPORTO</b>	<b>217</b>
V.1	CARATTERISTICHE DEL SERVIZIO	218
V.2	FUNZIONALITÀ DEL SERVIZIO	219
V.2.1	<i>Instaurazione di connessioni con prefissata qualità di servizio</i>	220
V.2.2	<i>Trasferimento di dati e controllo di flusso</i>	222
V.2.3	<i>Abbattimento di una connessione</i>	223
V.2.4	<i>Parametri di qualità di servizio</i>	225
V.2.5	<i>Metodi di specifica del servizio di trasporto</i>	227
V.3	PROTOCOLLI DI STRATO DI TRASPORTO	229
V.3.1	<i>Protocolli OSI</i>	230
<b>VI</b>	<b>LE RETI IN AREA LOCALE E METROPOLITANA</b>	<b>245</b>

---

VI.1	TOPOLOGIE.....	245
VI.2	ARCHITETTURE DI UNA LAN.....	253
VI.3	TECNICHE DI TRASMISSIONE NELLE LAN .....	255
VI.4	STRATO MAC.....	257
VI.5	PROTOCOLLI AD ACCESSO CASUALE.....	258
VI.5.1	<i>Protocollo ALOHA</i> .....	259
VI.5.2	<i>Protocollo Slotted ALOHA</i> .....	262
VI.6	PROTOCOLLO CSMA/CD.....	264
VI.7	PROTOCOLLO A TESTIMONE SU BUS .....	271
VI.8	PROTOCOLLO A TESTIMONE SU ANELLO.....	275
VI.9	LO STRATO LLC NELLE LAN.....	280
VI.9.1	<i>Servizio di strato LLC</i> .....	281
VI.9.2	<i>Protocollo di strato LLC</i> .....	286
VI.10	LA MAN DQDB.....	292
<b>VII</b>	<b>INTERNET.....</b>	<b>301</b>
<b>VIII</b>	<b>LA RETE ISDN.....</b>	<b>303</b>
VIII.1	ARCHITETTURA DI UNA ISDN .....	303
VIII.1.1	<i>Interfacce di accesso</i> .....	303
VIII.1.2	<i>Funzioni di trasporto</i> .....	312
VIII.1.3	<i>Potenzialità di trasferimento</i> .....	315
VIII.1.4	<i>Il blocco dei protocolli ISDN</i> .....	319
VIII.2	LO STRATO 1 ISDN .....	325
VIII.3	PROTOCOLLO CSMA/CR.....	329
VIII.4	IL LAP-D .....	330
<b>IX</b>	<b>BIBLIOGRAFIA.....</b>	<b>343</b>
IX.1	LISTA DELLE FIGURE, TABELLE ED EQUAZIONI .....	344
IX.1.1	<i>Figure</i> .....	344
IX.1.2	<i>Tabelle</i> .....	348
IX.1.3	<i>Equazioni</i> .....	349





# I TELEMATICA

Il termine telematica è un neologismo di origine francese, nato verso la fine degli anni settanta per sottolineare, con un acronimo delle parole *telecomunicazioni* e *informatica*, quegli aspetti delle tecnologie dell'informazione in cui il trattamento e il trasporto assumono ruoli complementari e inscindibili.

Già dalla seconda metà degli anni sessanta era iniziata un'integrazione sempre più spinta di risorse di elaborazione nei sistemi di telecomunicazione. Attualmente, ormai in prossimità della fine del secolo, questa integrazione sta raggiungendo un completo grado di maturazione, come è dimostrato dal largo impiego di unità di elaborazione sia come *apparecchiature terminali* nella fornitura di servizi di telecomunicazioni, sia come *apparecchiature infrastrutturali* al servizio delle *reti di telecomunicazione*, e cioè di quelle infrastrutture che rendono possibile il trasporto dell'informazione a distanza.

Nel caso di elaboratori utilizzati come apparecchiature terminali, e cioè come origine e/o come destinazione di informazioni, l'obiettivo dell'impiego è consentire quanto è richiesto in una molteplicità di applicazioni finalizzate ai più svariati campi delle attività umane, e cioè, per es., a una automazione delle procedure di ufficio, a un tele-controllo di processi produttivi o a una tele-elaborazione su modelli matematici. In tutte queste applicazioni è richiesto il colloquio a distanza tra apparecchiature in grado di elaborare e/o memorizzare l'informazione ovvero il colloquio tra queste ultime e operatori umani.

Circa poi l'utilizzazione dei calcolatori come apparecchiature infrastrutturali, l'obiettivo è quello di sfruttare le potenzialità di risorse elaborative per:

- consentire il trasporto e l'utilizzazione dell'informazione, assicurando il rispetto di opportune regole procedurali nello svolgimento del colloquio e mettendo in atto tutti quei provvedimenti in grado di fronteggiare (possibilmente con successo) eventi di natura aleatoria (per es. guasti di apparecchiature o disturbi trasmissivi) che potrebbero compromettere lo scambio delle informazioni;
- controllare l'evoluzione della comunicazione, per ciò che riguarda in particolare la sua inizializzazione e, se richiesto, l'aggiunta di valore a servizi di base;
- fornire tutti gli ausili necessari per lo svolgimento delle operazioni di

gestione del servizio e dell'infrastruttura che ne consente la fornitura, avendo come obiettivo la migliore utilizzazione delle risorse disponibili in condizioni sia normali che anormali (per es. in condizioni di guasto).

In definitiva, il livello di integrazione tra trattamento e trasporto dell'informazione, e cioè tra le tecniche che sono l'oggetto dell'informatica e delle telecomunicazioni, è oggi così elevato da rendere eccessivamente totalizzante il termine telematica se il suo significato fosse quello inizialmente proposto. In suo luogo gli si attribuisce qui un significato più ristretto, secondo cui *la telematica è "un insieme disciplinare che tratta i principi e le soluzioni tecniche su come è possibile porre in corrispondenza applicazioni residenti in sistemi di elaborazione tra loro interconnessi con risorse di telecomunicazione"*.

In questo quadro rientrano:

- le tecniche impiegate per strutturare e per codificare le informazioni gestite da sistemi di elaborazione al fine di un colloquio tra uomo e macchina o tra macchina e macchina;
- l'organizzazione delle funzioni, per il trasferimento e per l'utilizzazione dell'informazione, svolte dalle apparecchiature terminali e infrastrutturali affinché abbia luogo la corrispondenza tra applicazioni (*architetture di comunicazione*);
- le infrastrutture per il trasporto a distanza dell'informazione la cui origine e/o la cui destinazione sono sistemi di elaborazione;
- i modi operativi secondo cui avviene il trasporto dell'informazione da un'origine a una destinazione;
- le modalità necessarie per assicurare l'accesso a risorse di trasferimento condivise quando questo avviene con un controllo affidato direttamente alle sorgenti di informazione, senza coordinamenti di tipo centralizzato;
- le tecniche utilizzate per controllare il traffico che interessa un'infrastruttura, per indirizzare l'informazione verso la destinazione desiderata e per assicurare un corretto instradamento nell'operazione di trasporto;
- le *applicazioni* e i *servizi*, che sono stati normalizzati nel rispetto di opportuni riferimenti architettureali e che per questo motivo godono dell'attributo "telematico".

Nel seguito di questo capitolo la trattazione è rivolta a tre dei temi ora elencati, e precisamente alle infrastrutture (par. I.1), alle architetture (par. I.2), nonché alle applicazioni e ai servizi con distinzione di quelli legati a un ambiente OSI (par. I.3) dagli altri sviluppati in ambito Internet (par. I.4).

## I.1 Infrastrutture

In questo paragrafo saranno presentate alcune infrastrutture tipiche per applicazioni telematiche. La trattazione sarà rivolta prevalentemente alle *reti in area geografica* (WAN-Wide Area Network), senza tuttavia trascurare le *reti in area locale* (LAN, Local Area Network) o *area metropolitana* (MAN-Metropolitan Area Network).

Nell'ambito delle reti in area geografica si distingueranno poi le reti dedicate (*WAN per dati*) da quelle integrate nei servizi (*ISDN- Integrated Services Digital Network*) sia a banda stretta (*N-ISDN -Narrowband ISDN*) che a banda larga (*B-ISDN -Broadband ISDN*), con la precisazione che oggi il paradigma ISDN è di attualità prevalentemente per gli aspetti riguardanti le potenzialità di accesso.

Infine nell'organizzazione della trattazione, è sembrato opportuno partire con un cenno alle WAN per dati di prima generazione, nel seguito indicate come *reti X.25*. Ciò infatti consentirà di meglio inquadrare le infrastrutture che sono ora di attualità. Tra queste un primo approccio, denominato nel seguito *rete FR* (Frame Relay) e diretto successore delle reti X.25, si è sviluppato con i criteri architettonici del paradigma ISDN e appare attualmente indirizzato in modo esclusivo alle comunicazioni di dati.

Il secondo riferimento riguarda invece Internet e cioè la rete che ha acquistato un ruolo dominante negli ultimi anni come infrastruttura telematica a livello mondiale: attualmente come WAN per dati e, in prospettiva a breve termine, anche con apertura alle comunicazioni multimediali.

### I.1.1 WAN per dati di prima generazione

Le reti per dati in area geografica (WAN per dati) sono di concezione e di realizzazione molto più recente (anni '70) rispetto a quelle telefoniche. Esse quindi si sono sviluppate in un ambiente di comunicazione ormai fortemente orientato all'impiego delle tecniche numeriche e sono perciò inquadrabili come reti numeriche integrate (IDN, Integrated Digital Network). Nel passato, le WAN per dati sono state realizzate con modi di trasferimento sia a circuito, che a pacchetto. Oggi le scelte tecniche sono esclusivamente verso le reti con orientamento al pacchetto.

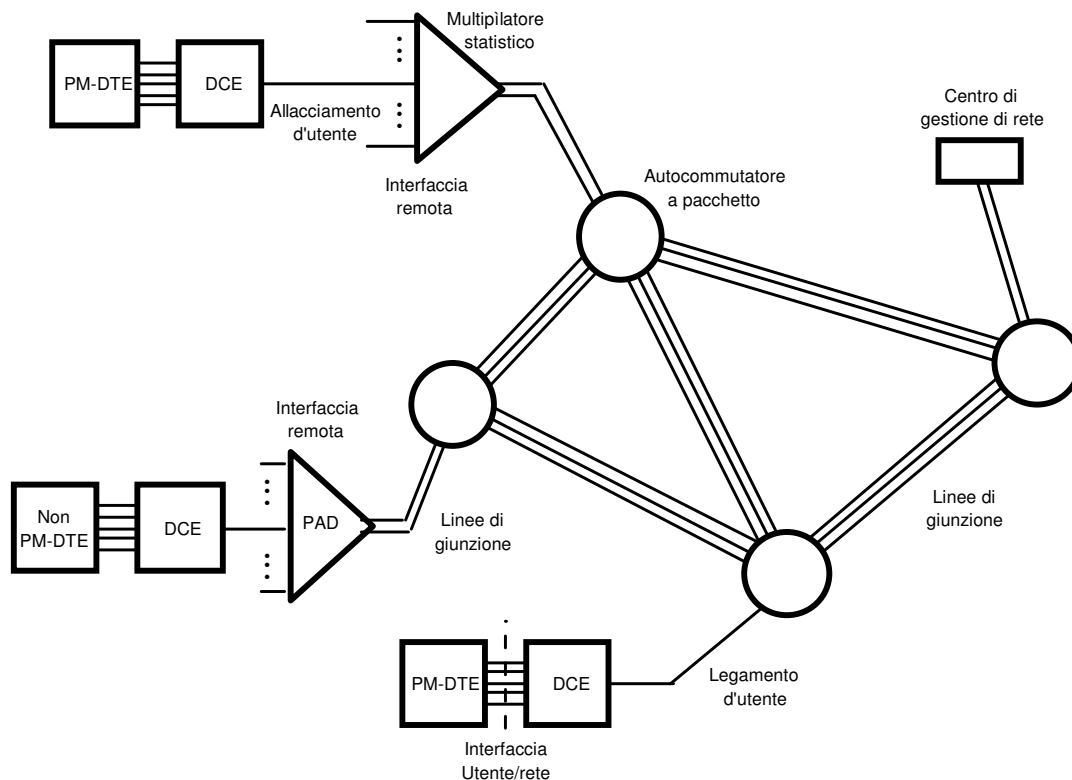
Le WAN a circuito per dati sincroni hanno seguito gli stessi principi applicati nel caso delle IDN telefoniche. Tecniche opportune sono state usate per sotto-multiplexare flussi di dati a ritmo binario inferiore a 64 kbit/s (sotto-ritmo) su un canale avente capacità di 64 kbit/s. Il modo di trasferimento a

circuito per flussi di dati a sotto-ritmo è stato realizzato secondo due modalità. In primo luogo, si è utilizzato un pre-commutatore operante a sotto-ritmo. Come seconda modalità si è impiegata una normale rete di connessione operante a 64 kbit/s facendo ricorso a modalità di riempimento/svuotamento degli intervalli temporali al fine di trattare correttamente i flussi a sotto-ritmo.

Sempre nel passato, reti per dati a circuito sono state realizzate anche per comunicazioni di dati asincroni (start-stop). Tali reti dedicate sono state progettate, almeno all'origine, per trattare il servizio telex o dati start-stop fino a 300 bit/s.

In Fig. I.1 è mostrato un modello di WAN per dati a pacchetto di prima generazione. In esso si distinguono:

- la sezione interna, costituita da nodi di commutazione, interconnessi da linee di giunzione numeriche operanti con una capacità non inferiore a 64 kbit/s;
- la sezione di accesso, comprendente le apparecchiature di concentrazione, di moltiplicazione e di interfacciamento dei DTE;
- il centro di gestione di rete, preposto a funzioni di supervisione per una efficiente utilizzazione delle risorse.



PM - DTE : DTE operante a pacchetto  
 Non PM - DTE : DTE non operante a pacchetto

Fig. I.1 - Rete per dati a pacchetto

Le tecnologie allora disponibili e la necessità di assicurare elevata capacità di traffico hanno portato per i nodi di commutazione a criteri di progetto di tipo modulare, secondo strutture distribuite in cui unità specializzate per funzioni (interfaccia di linea, gestione dei protocolli) erano collegate tra loro da strutture di interconnessione ad alta velocità.

Nella prima generazione delle WAN per dati a pacchetto, il funzionamento della sezione interna ha applicato i principi del modo di trasferimento a pacchetto con servizio di trasferimento *con connessione* (commutata o semi-permanente) o *senza connessione*. Il caso di connessione commutata è stato denominato servizio a *chiamata virtuale* (VC-Virtual Call), mentre quello di connessione semi-permanente è noto come servizio a *circuito virtuale permanente* (PVC- Permanent Virtual Circuit). Infine, quando la scelta di servizio di trasferimento si è orientata verso una soluzione senza connessione, si è parlato di servizio *datagramma* (DG-Datagram).

Anche nella sezione di accesso sono stati impiegati gli stessi principi di modo di trasferimento con un servizio che, in linea di principio, è potuto essere

diverso da quello adottato nella sezione interna. In pratica, se il servizio era DG nella sezione interna, poteva essere DG anche nella sezione di accesso, ma si sono avuti casi in cui si è preferita per l'accesso la soluzione VC/PVC. Se invece il servizio era VC/PVC nella sezione interna, era normalmente VC/PVC nella sezione di accesso.

Per ciò che riguarda i modi di accesso a una rete e pacchetto di prima generazione, in Fig. I.1 sono previsti tre casi. Due di questi riguardano DTE che sono in grado di pacchettizzare l'informazione e cioè *DTE operanti a pacchetto* (PM-DTE, Packet Mode DTE), mentre il terzo si riferisce a DTE che, come i terminali start-stop, non hanno questa capacità (*DTE non operanti a pacchetto*). Relativamente ai primi due casi, in Fig. I.1 vengono distinte due possibilità di accesso: una prima con linea d'utente individuale; l'altra con accesso intermedio a un multiplatore statistico, su cui convergono altri DTE. Per ciò che riguarda poi il terzo caso, dispositivi per l'*assemblamento/ disassemblamento dei pacchetti* (PAD - Packet Assembling/Disassembling) sono impiegati per interfacciare DTE non operanti a pacchetto; l'accesso avviene tipicamente tramite connessioni commutate sulla rete telefonica.

I primi due casi considerati in Fig. I.1 si riferiscono a DTE operanti a pacchetto che si interfacciano con il proprio DCE secondo la normativa della Racc. X.25. Apparecchi terminali dello stesso tipo possono anche accedere a una rete a pacchetto per il tramite di una rete a circuito (telefonica o per dati); in questo caso l'interfaccia DTE-DCE è definita nella Racc. X.32. Con riferimento poi al terzo caso, l'interfaccia DTE-DCE per terminali start-stop che accedono a una rete a pacchetto tramite una unità PAD è normalizzata nella Racc. X.28.

Infine, per consentire uno scambio di informazione tra due DTE, di cui uno operante a pacchetto e l'altro senza questa capacità per il tramite di un PAD ovvero tra due PAD, sono previste le procedure descritte nella Racc. X.29.

### *I.1.2 Reti integrate nei servizi*

Una ISDN (Integrated Service Digital Network) è una rete che, evolvendo dalla IDN telefonica, è in grado di offrire una completa connettività numerica da utente a utente per la fornitura di un'ampia gamma di servizi, vocali e non vocali, ai quali è possibile accedere attraverso un insieme ristretto di interfacce utente-rete normalizzate.

Da questa definizione emergono i tre aspetti chiave che sono alla base di una ISDN:

- la completa *connettività numerica da utente a utente*;
- la capacità di fornire una *molteplicità di servizi* di telecomunicazione;

- la normalizzazione di un *ristretto numero di interfacce utente-rete*.

Per ciò che riguarda i punti 1) e 2), questi non richiedono ulteriori commenti. D'altra parte, il punto 3) sottolinea che, per raggiungere una completa integrazione di trattamento dei vari servizi, il primo passo essenziale è la normalizzazione di un numero limitato di interfacce utente-rete. Ciò consente di superare, con evidenti vantaggi economici, la situazione offerta dalle reti dedicate, in cui l'accesso a singoli servizi (o a limitati insiemi di questi) richiede interfacce diverse da rete a rete.

Una ISDN è pienamente descritta dalle caratteristiche di servizio offerte alle sue interfacce di accesso, piuttosto che dalle soluzioni sistemistiche e tecnologiche adottate nella sezione interna. Ciò comporta una completa indipendenza tra l'architettura della sezione interna e i servizi offerti agli utenti; l'evoluzione di questi due aspetti può avvenire in modo completamente separato. Conseguentemente una ISDN può essere realizzata in una varietà di configurazioni, a seconda dell'ambiente di comunicazione dal quale emerge, in termini dello stato iniziale della rete e della sua evoluzione successiva.

Nella prima fase di sviluppo (già conclusa) la ISDN si è basata sulla IDN telefonica, introducendo in questa le potenzialità connesse alla *linea di utente numerica*. Attraverso tale rete è stato possibile accedere ad altre infrastrutture dedicate ad un servizio, quali quelle per dati a pacchetto.

Compatibilmente poi con quanto consentito dall'evoluzione tecnologica e dai vincoli economici, la ISDN ha incorporato progressivamente funzioni aggiuntive e prestazioni migliorate, incluse quelle fornite dalle strutture dedicate ad un servizio. Chiariamo questo punto.

In primo luogo, a conclusione della prima fase, i nuovi servizi introdotti nella ISDN sono stati basati su connessioni commutate o semi-permanenti con capacità di trasferimento uguale a 64 kbit/s; attualmente sono già disponibili o sono previste a breve termine anche connessioni commutate o semi-permanenti con capacità di trasferimento superiore ed, eventualmente, inferiore a 64 kbit/s.

Inoltre, le funzionalità di una ISDN non devono limitarsi al trasporto dell'informazione, ma devono comprendere anche trattamenti di questa con lo scopo di fornire *servizi a valore aggiunto*. Le risorse di elaborazione necessarie per questo scopo potranno essere inserite direttamente nella ISDN come risorse condivise e, quando richiesto da specifici servizi, anche negli apparecchi terminali, curandone la compatibilità con quanto offerto dalla rete.

In terzo luogo, essenziale per le funzioni di una ISDN è l'integrazione delle sue funzionalità per tutto ciò che concerne la gestione della rete, nei suoi vari aspetti coinvolti dai problemi di esercizio, di amministrazione e di

manutenzione. Per la soluzione di questi problemi un ruolo chiave hanno ancora le risorse di elaborazione messe a disposizione dalla ISDN.

### *1.1.3 Rete a larga banda integrata nei servizi (B-ISDN).*

Il primo esempio di integrazione nella fornitura di servizi è rappresentato dalla ISDN che, come descritto nel paragrafo precedente, è stata definita nei suoi aspetti base nel corso degli anni '70. La sua caratteristica fondamentale è quella di fornire una connettività numerica da estremo a estremo per il supporto di servizi con banda inferiore a 2 Mbit/s, tipicamente il servizio telefonico e i servizi dati a bassa velocità.

Occorre osservare, comunque, che il concetto di integrazione adottato nella ISDN è limitato alle modalità di accesso dell'utente alle funzionalità di trasporto; quest'ultime rimangono però differenziate per tipologia di traffico e di servizio ed è compito della centrale locale dirigere il traffico verso la rete di trasporto più adatta (circuito, pacchetto, segnalazione).

Nel corso degli anni '80 lo scenario dei servizi si è rapidamente modificato rispetto a quello considerato nella definizione della ISDN. Sono emerse nuove esigenze di trasferimento in rete geografica principalmente dovute a:

- il supporto di flussi dati ad elevata capacità (oltre i 10 Mbit/s) quali quelli derivanti dalla interconnessione di reti locali o dal collegamento tra elaboratori remoti per il supporto di applicazioni distribuite;
- il supporto di flussi numerici derivati dalla codifica di immagini sia fisse che in movimento con diverso grado di definizione;
- il supporto di servizi richiedenti il trasferimento di diverse tipologie di flussi informativi (es. audio, video e/o dati), normalmente indicati come servizi multimediali.

Queste nuove esigenze hanno dato luogo ad un processo di avanzamento del concetto di integrazione in cui fossero compresi anche il trasferimento di flussi numerici ad alta velocità e le relative funzionalità di accesso. L'approdo finale di questo processo è stato la definizione della Rete Numerica Integrata nei Servizi a Larga Banda (Broadband ISDN, B-ISDN).

L'ITU, l'ente di standardizzazione internazionale nel settore delle comunicazioni, nella raccomandazione I.121, definisce la B-ISDN come "la rete universale in grado di supportare differenti tipi di applicazioni e categorie di utenti". L'obiettivo più ambizioso nella definizione della B-ISDN è quello di estendere il concetto di integrazione non solo alle funzionalità di accesso d'utente, ma anche a quelle di trasporto individuando una modalità unica di trattamento in rete dei flussi informativi, ovvero un unico modo di

trasferimento.

La definizione del modo di trasferimento più adatto al supporto della molteplicità di flussi informativi previsti in un ambiente a larga banda ha richiesto vari anni di studio. La convergenza verso una soluzione si è avuta intorno alla fine degli anni '80 ed il risultato è stato la definizione del Modo di Trasferimento Asincrono (Asynchronous Transfer Mode – ATM) che l'ITU definisce come “il modo di trasferimento per la realizzazione della B-ISDN”.

E' noto che i requisiti che devono essere soddisfatti da un modo di trasferimento sono i seguenti:

- flessibilità nella assegnazione della banda;
- efficienza di utilizzazione delle risorse;
- minimizzazione del ritardo di trasferimento e della sua variabilità;

I modi di trasferimento tradizionali, circuito e pacchetto, non potevano essere direttamente utilizzati nella B-ISDN:

- il modo trasferimento a circuito, se, da un lato, garantisce un ritardo di trasferimento in rete estremamente basso e soprattutto costante, dall'altro, a causa dell'uso della moltiplicazione statica, pone notevoli problemi di flessibilità nell'assegnazione della banda e di efficienza di utilizzazione delle risorse trasmissive;
- il modo trasferimento a pacchetto, viceversa, basandosi sull'uso della moltiplicazione dinamica, è intrinsecamente orientato al raggiungimento di elevati valori di efficienza ed è naturalmente flessibile nell'assegnazione della banda; il suo punto debole riguarda il ritardo di trasferimento che è variabile in funzione dello stato di congestione in cui si trova la rete al momento del transito dei pacchetti appartenenti ad una comunicazione.

Alle precedenti considerazioni va aggiunto che i vincoli di integrità posti dai vari tipi di applicazioni possono essere estremamente diversi, imponendo o meno la necessità di prevedere meccanismi di controllo d'errore e quindi un trattamento differenziato dei vari tipi di servizi.

Da quanto detto, risulta evidente che l'adozione di un unico modo di trasferimento poteva emergere solo dalla ricerca di un compromesso tra i fattori ora descritti. L'ATM è il risultato di questo compromesso.

L'ATM è un modo di trasferimento orientato al pacchetto; i pacchetti, indicati con il nome di celle, sono di lunghezza fissa. Il formato di una cella è costituito da 5 byte di intestazione (header) e da 48 byte di campo informativo (payload). Le celle sono assegnate ad una sorgente su domanda, in dipendenza delle sue caratteristiche di attività; la gestione delle eventuali contese di utilizzazione è orientata al ritardo.

L'ATM è un modo di trasferimento orientato alla connessione (connection oriented); ogni qualvolta due sistemi terminali devono scambiarsi informazioni occorre preliminarmente instaurare una connessione, indicata con il nome di "Virtual Channel" (VC) che sarà abbattuta al termine del trasferimento informativo.

I vantaggi derivanti dal concetto di connessione sono: i) semplicità di commutazione, si evita che un nodo interno alla rete debba scegliere il cammino per ogni singola cella; ii) possibilità di garantire fissati livelli di qualità di servizio, all'atto dell'instaurazione della connessione la rete può preventivamente verificare che le risorse a disposizione siano sufficienti a garantire i valori di QoS dichiarati dagli utenti.

I problemi relativi alla variabilità del ritardo di transito sono stati risolti cercando di diminuire il più possibile i fenomeni di accodamento all'interno dei nodi; ciò è possibile operando in due direzioni:

- aumentando la velocità delle linee e quindi diminuendo il tempo di trasmissione delle unità informative;
- diminuendo il carico elaborativo dei nodi di commutazione legato al processamento di ogni singola cella in modo da aumentare il throughput del nodo stesso.

Il secondo obiettivo si è raggiunto:

- assegnando ai nodi interni della rete (nodi di transito) solo l'esecuzione dell'insieme delle funzioni che sono comuni a tutti di applicazione;
- facendo migrare verso la periferia della rete (terminali o nodi di accesso) tutte le altre funzioni che sono invece dipendenti dal tipo di applicazione.

Si osservi inoltre che, le funzioni appartenenti al secondo insieme sono normalmente caratterizzate anche dalla maggiore complessità elaborativa (es. controllo di sequenzialità, d'errore, equalizzazione dei ritardi).

L'architettura della B-ISDN risultante dall'applicazione dei principi ora esposti è mostrata in Fig. I.2. Questa risulta essere composta da due strati: i) lo strato interno (core) costituisce una rete di trasporto universale e raggruppa le funzionalità a banda stretta, a larga banda e di segnalazione comuni a tutte le applicazioni; ii) lo strato esterno (edge) a cui appartengono i terminali nativi ATM ed, eventualmente, le centrali di accesso verso segmenti di rete non ATM; tale strato comprende le funzionalità dipendenti dal tipo di applicazione, ovvero svolge funzioni specifiche dei particolari servizi.

Come sarà più approfonditamente descritto nel seguito, l'architettura funzionale mostrata in Fig. I.2 determina le caratteristiche dell'architettura protocollare della B-ISDN. E' il caso di anticipare che l'architettura

protocollo della B-ISDN si articola in quattro strati:

- 1) Strato Fisico (Physical layer – PH): questo strato comprende tutte le funzioni di gestione del mezzo trasmissivo e quelle riguardanti la trasmissione e la ricezione dei bit informativi; inoltre comprende le funzioni di adattamento delle celle alle caratteristiche specifiche dell'interfaccia trasmissiva. Le funzionalità dello strato ATM sono eseguite in qualsiasi elemento di rete (terminale o nodo).
- 2) Strato ATM (ATM Layer): comprende le funzioni di trattamento delle celle, in particolare, appartengono allo questo strato le funzioni di elaborazione delle informazioni di indirizzamento contenute nelle celle e le funzioni di commutazione di quest'ultime; si osservi che, concordemente a quanto illustrato in precedenza, le funzioni ora descritte costituiscono l'insieme delle funzioni comuni a tutte le applicazioni, qualsiasi altra funzione è specifica di una particolare applicazione e, come tale, non fa parte dello strato ATM; le funzionalità dello strato ATM sono eseguite in qualsiasi elemento di rete.

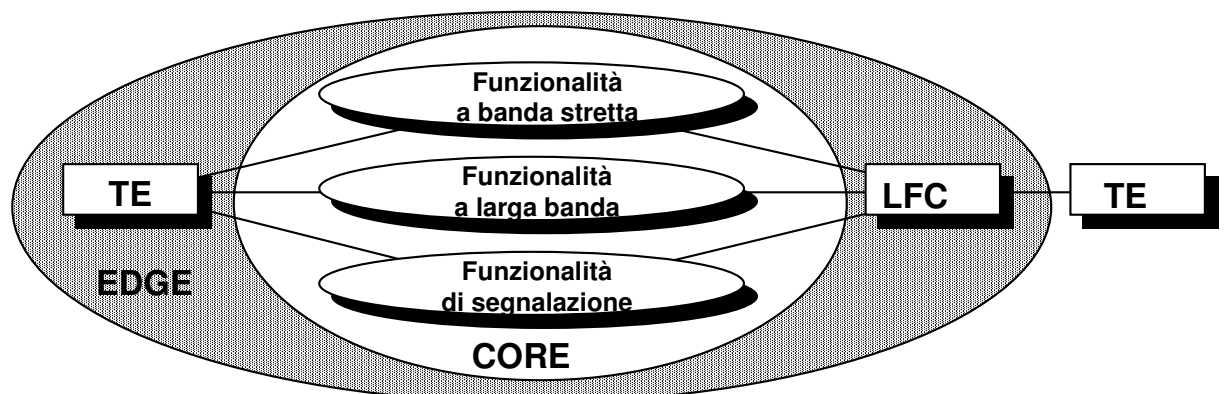


Fig. I.2 - Architettura funzionale della B-ISDN  
LFC. Local Function Capabilities, TE: Terminal Equipment

- 3) Strato di Adattamento (ATM Adaption Layer – AAL): comprende le funzioni necessarie all'adattamento del servizio offerto dallo strato ATM ai requisiti posti dalla specifica applicazione; è evidente come le funzioni di tale strato siano dipendenti dal tipo di applicazione e quindi queste sono eseguite solo dagli elementi compresi dello strato esterno della rete (terminali o nodi di accesso).
- 4) Strati superiori (higher layers): comprendono le funzioni esterne alla B-ISDN e dipendono dalle caratteristiche dei protocolli posti al di sopra dello strato di adattamento; in particolare possono riguardare sia funzionalità applicative, nel caso il terminale sia nativo ATM, oppure funzionalità di altri protocolli (IP, Frame Relay, LLC/MAC, ecc.) nel caso in cui la rete ATM sia connessa

ad altre reti funzionanti secondo paradigmi diversi da quello ATM.

#### 1.1.4 Reti in area locale e metropolitana

Negli ambienti di ufficio si è già da tempo affermata una capillare diffusione dei PC per coadiuvare l'esecuzione del lavoro individuale. A tale distribuzione delle risorse elaborative spesso però non corrisponde un'equivalente distribuzione di risorse di altro tipo, ugualmente necessarie all'esecuzione dei singoli compiti. Ad esempio, archivi di dati (data base) e periferiche costose (stampanti di qualità, dispositivi grafici, grosse memorie di massa, ecc.) rimangono centralizzate per ragioni di efficienza di gestione e/o di costo. Ciò rende assolutamente necessario un mezzo rapido di comunicazione che metta in grado il PC del singolo utente di reperire, da un lato, le informazioni necessarie al proprio lavoro e di utilizzare, dall'altro, i dispositivi che rimangono centralizzati.

Per soddisfare tali esigenze sono state sviluppate le reti in area locale (LAN) con il compito di interconnettere apparecchiature terminali che sono sorgenti o collettori di dati negli ambienti di ufficio. Queste apparecchiature terminali sono nel seguito indicate come “*stazioni*” in conformità alla terminologia utilizzata per le LAN. Come loro elementi distintivi, queste reti:

- hanno una *limitata estensione geografica*, con uno sviluppo lineare non superiore orientativamente ad una decina di chilometri e in generale racchiuso entro un edificio o un gruppo di edifici tra loro vicini;
- sono normalmente *infrastrutture private*, al servizio di una singola organizzazione, che normalmente ne cura anche la gestione;
- interconnettono le stazioni di loro competenza con un *unico* canale trasmissivo che è *condiviso* dalle sorgenti di informazione quando queste si trovano nello stato di trasferimento;
- utilizzano modi di trasferimento *orientati al pacchetto*;
- come conseguenza della loro limitata estensione geografica, consentono una *capacità di trasferimento relativamente elevata* (maggiore orientativamente di un Mbit/s) con un *basso tasso di errore* (minore, sempre orientativamente di  $10^{-6} \div 10^{-7}$ ).

Quest'ultima caratteristica consente rapidi scambi informativi tra macchine cooperanti e quindi non introduce limitazioni nella velocità elaborativa del sistema distribuito.

Da un punto di vista logico, il canale trasmissivo comune supporta un *unico mezzo di comunicazione* che costituisce la via di percorrenza dell'informazione per il trasporto da una stazione all'altra. Nel seguito questa

via sarà chiamata il *mezzo* della LAN.

Questo mezzo supporta un trasferimento di tipo *multiplato dinamicamente*: cioè ad una stazione che, in un certo istante, abbia una *unità informativa* (UI) da emettere, viene assegnata l'intera capacità di trasferimento del mezzo per una durata corrispondente allo svolgimento dell'operazione di trasporto.

Nelle LAN oggi disponibili sul mercato il mezzo è condiviso con una *multiplazione a divisione di tempo*. Non sono tuttavia da escludere, in prospettiva, soluzioni alternative che però, per brevità, non verranno ulteriormente considerate nel seguito di questa trattazione.

La multiplazione dinamica a divisione di tempo implica che tra le stazioni possano determinarsi situazioni di contesa: ciò si verifica quando la richiesta di accesso al mezzo presentata da due o più stazioni comporterebbe, se soddisfatta immediatamente, l'impegno della risorsa comune in intervalli di tempo sovrapposti. Per risolvere queste situazioni di contesa, considerazioni di economia di sistema hanno guidato verso l'adozione di un controllo distribuito, e cioè verso una cooperazione tra le stesse stazioni e senza la necessità di una unità di controllo centralizzato.

Questa soluzione deriva da quella adottata per l'interconnessione tra microprocessori in un ambiente multiprocessore. In tali sistemi è spesso presente un mezzo di comunicazione ad alta velocità (bus) che interconnette i vari processori, a cui questi accedono per lo scambio dei dati. È evidente tuttavia che il grado di integrazione delle comunicazioni raggiungibile in un ambiente multiprocessore non è pensabile in una LAN che, anzi, ha lo scopo di permettere la comunicazione anche tra dispositivi eterogenei.

La soluzione a controllo distribuito, se da un lato aumenta la complessità dell'interfaccia tra ogni stazione e il mezzo, dall'altro consente un elevato grado di flessibilità e un aumento dell'affidabilità di tutto il sistema per l'eliminazione di elementi centralizzati. Si osserva che tali requisiti sono di fondamentale importanza in un ambiente, quale quello di un ufficio, per sua natura estremamente variabile, sia per il numero che per la dislocazione fisica delle stazioni.

Infatti un architettura a controllo distribuito è intrinsecamente più affidabile rispetto ad una soluzione analoga a controllo centralizzato. In questo caso, un guasto nell'unità centrale di controllo determina, nella maggior parte dei casi, un periodo di fuori servizio dell'intera rete. Ciò costringe spesso a replicare tale unità per ridurre la probabilità di fuori servizio. Invece, nelle architetture a controllo distribuito il guasto di una o più stazioni non compromette il funzionamento della rete, che può continuare ad offrire il

proprio servizio alle stazioni ancora funzionanti.

La modalità di condivisione ora introdotta, basata su una moltiplicazione in cui l'accesso al mezzo è regolato tramite un controllo distribuito, è denominata *accesso multiplo*, mentre le procedure di cooperazione atte a risolvere le situazioni di contesa in questo ambiente di comunicazione prendono il nome di protocolli di *controllo di accesso al mezzo* (MAC, Medium Access Control).

Come già sottolineato all'inizio di questo paragrafo, le LAN sono nate e si sono sviluppate come reti dedicate a servizi di comunicazione di dati. Attualmente, anche in relazione all'aumento delle loro capacità di trasferimento, alcuni tipi di LAN sono utilizzate anche per comunicazioni multimediali e quindi per l'intero insieme delle applicazioni telematiche che sono oggi di interesse.

Da quanto esposto emergono le sostanziali differenze tra LAN e WAN. Quest'ultima è una rete che copre un'area geografica anche molto estesa, con le dimensioni, ad esempio, di un'intera nazione. Normalmente mette a disposizione dell'utenza connessioni con capacità di trasferimento limitata (dell'ordine del Mbit/s) ed è, nella maggioranza dei casi, di tipo pubblico. Infatti, gli utenti sono di tipo eterogeneo e l'organizzazione che gestisce la rete è diversa da quella di cui gli utenti fanno parte.

Una terza classe di reti, che si colloca in posizione intermedia tra LAN e WAN, è rappresentata dalle *reti in area metropolitana* (MAN, Metropolitan Area Network). Come lo stesso nome suggerisce, una MAN è una rete che è concepita per interconnettere stazioni distribuite su di un'area geografica corrispondente ad un quartiere cittadino o ad un'intera città. Inoltre una MAN usa una tecnologia simile a quella delle LAN (ad es. un canale trasmissivo condiviso, un mezzo multiaccesso, ecc.). È inoltre caratterizzata da un aumento delle capacità di trasferimento; si raggiungono infatti, mediante l'uso di mezzi trasmissivi in fibra ottica, capacità superiori a 100 Mbit/s.

Un esempio di MAN è costituito dalle reti via cavo per la distribuzione del segnale televisivo (Cable Television - CATV). Tali reti interconnettono una centrale locale, posta a livello di quartiere o di intera città, con le singole postazioni d'utente. Data la particolarità del servizio supportato, la trasmissione è analogica ed è di tipo essenzialmente diffusivo e unidirezionale, mentre la topologia è ad albero.

Una MAN può essere di tipo sia privato che pubblico. Nel primo caso, la sua funzionalità è del tutto simile ad una LAN, in cui si ha la possibilità di estendere la connettività a sedi della stessa organizzazione geograficamente

distanti. In questo caso, spesso la MAN viene indicata con il termine *LAN ad alta velocità* (High Speed LAN - HSLAN).

Se la MAN è invece di tipo pubblico, essa è gestita da un'unica organizzazione, ma è utilizzata da utenti di tipo diverso. In questo caso la MAN si configura come una rete di interconnessione per un'utenza localizzata di tipo particolare, avente esigenze di comunicazione sofisticate, quali, ad esempio, il trasferimento di dati ad alta velocità od l'interconnessione di LAN. Essa inoltre può funzionare come rete di raccolta di traffico da inoltrare verso una WAN per trasferimenti a lunga distanza.

#### *I.1.5 Internet*

Internet è una infrastruttura per comunicazioni di dati (*rete di calcolatori*), che si è sviluppata seguendo le linee-guida del modello omonimo. Attualmente sta conoscendo un successo a livello mondiale che non ha precedenti per la sua estensione geografica, per la numerosità dei suoi utenti e per la vastità di interessi coinvolti. A titolo puramente orientativo e con riferimento alla fine del 1997, si parla di oltre 20 milioni di calcolatori connessi e si stimano in circa 70 milioni gli utenti serviti (tenendo conto che uno stesso calcolatore può essere usato da più utenti).

Il modello architetturale di Internet sarà discusso con più dettaglio nel seguito (Cap. VII). Qui è importante sottolineare che Internet non è una nuova rete, progettata e costruita ex novo o necessariamente alternativa alle soluzioni discusse in precedenza, ma l'unione di diverse reti, spesso pre-esistenti ed eterogenee tra loro. Le reti componenti Internet sono però omogenee al loro interno e possono operare in accordo ad uno qualsiasi dei paradigmi già introdotti.

La struttura fisica di Internet è costituita quindi da un certo numero di reti componenti e da alcuni dispositivi che le interconnettono; tali dispositivi di interconnessione, mediante l'uso di opportuni protocolli di comunicazione, consentono a calcolatori connessi a reti diverse di scambiarsi informazioni.

Elemento caratterizzante di Internet è il modo in cui questa rete è nata e si è sviluppata. I primi studi sulla interconnessione di calcolatori furono effettuati alla fine degli anni sessanta. Si voleva allora sperimentare una tecnica che permettesse la condivisione di linee di comunicazione da parte di utenti attestati su sistemi diversi, sfruttando la tecnica della commutazione di pacchetto. In quegli anni il Ministero della Difesa degli Stati Uniti (*Department of Defense, DoD*) incaricò l'organizzazione *ARPA (Advanced Research Project Agency)* di sviluppare e sperimentare una rete di calcolatori.

Dovendo avere applicazioni militari, uno dei requisiti fondamentali che tale rete doveva soddisfare era la resistenza ad attacchi che ne modificassero la configurazione (ad. es. per interruzione di linee di giunzione o di centri di commutazione). Sue caratteristiche dovevano quindi essere una struttura reticolare, magliata, non gerarchica e con un'elevata capacità di interconnessione e di interlavoro fra una grande varietà di nodi di rete. Per lo stesso motivo si scelse di adottare una modalità di trasferimento senza connessione e senza garanzie di qualità del servizio, rimandando queste ultime ai livelli superiori dell'architettura protocollare. La rete nata da questo progetto, e denominata ARPANET, è stata il primo esempio di rete geografica a commutazione di pacchetto. Nel 1983 tale rete fu separata in due parti: una civile (ARPANET) ed una militare (MILNET). Nel 1985, sempre negli Stati Uniti, la National Science Foundation finanziò lo sviluppo di una rete di trasporto a lunga distanza (NSFnet) e di reti regionali, che consentirono di interconnettere LAN di diverse università e di enti di ricerca alla rete ARPANET.

La rete Internet si è sviluppata a partire da questo nucleo iniziale fino ad estendersi in tutto il mondo e comprendendo non più solo organizzazioni pubbliche o di ricerca e al servizio della comunità scientifica ma anche organizzazioni commerciali ed utenti privati.

Le entità di calcolo all'interno di questa struttura sono chiamati host. Questi possono essere super-computer paralleli, mini-computer, workstation, semplici personal computer o calcolatori portatili. I ritmi binari di trasferimento delle informazioni possono essere molto variabili, sia nel tempo che in funzione delle coppie origine-destinazione considerate (da pochi bit/s fino a centinaia di Mbit/s).

A causa delle particolari esigenze sopra menzionate, e poste dall'organizzazione ARPA, si determinò una divergenza tecnica iniziale tra le scelte effettuate per ARPANET e quelle adottate dagli organismi di standardizzazione ISO e ITU-T (ex CCITT). La divergenza iniziale sulle scelte tecniche è perdurata nel tempo attraversando gli anni settanta ed ottanta, durante i quali ARPANET cresceva notevolmente negli Stati Uniti, espandendosi dall'ambito iniziale della difesa a quello dell'istruzione e della ricerca e conquistando sul campo un numero di utenti sempre maggiore. Nel frattempo ISO e ITU-T lavoravano alla definizione di un modello teorico di interconnessione tra sistemi aperti (l'Open Systems Interconnection, OSI) ed alla specifica di una serie di protocolli di comunicazione conformi a tale modello.

Fino alla fine degli anni ottanta era opinione diffusa che OSI ed i suoi protocolli sarebbero stati universalmente adottati.

Tale previsione si è però dimostrata non corretta: la realtà Internet si è infatti talmente diffusa, sia negli Stati Uniti che nel resto del mondo, che ha fatto impedire ad OSI di svilupparsi significativamente. La rete Internet ha raggiunto gli obiettivi che OSI si prefiggeva, costituendo una base di interconnessione aperta ed indipendente dai costruttori di sistemi di telecomunicazione.

Internet è oggi una infrastruttura di comunicazione le cui risorse sono dislocate spazialmente in tutto il mondo, ma che sono viste dal singolo utente in modo trasparente, senza cioè che sia necessario sapere dove sono fisicamente, ed a cui è molto facile accedere con strumenti progettati e realizzati per i suoi utenti.

Fra le principali caratteristiche ed i fattori di successo che hanno portato allo sviluppo esplosivo di Internet, si possono citare i seguenti:

- la modalità di trasferimento, nello strato di rete, è senza connessione e Internet non fornisce alcuna garanzia sulla qualità di servizio; Internet si impegna a fare del suo meglio, ma non è possibile, in generale, essere sicuri o contrattare che il trasferimento avvenga con determinate specifiche (integrità informativa, ritardo di trasferimento, grado di trasparenza temporale, etc.). Il compito di rendere la qualità di servizio adeguata alle esigenze degli utenti è demandato ai livelli applicativi residenti nei sistemi terminali. Il paradigma Internet offre quindi meno funzionalità del modello OSI (almeno negli strati bassi dell'architettura protocollare), ma la sua realizzazione è più semplice;
- la modalità di funzionamento di Internet è diversa rispetto a quella tipica di altre reti; infatti essa non richiede che tutti i sistemi che la compongono siano basati su di una identica e comune architettura protocollare;
- in molte reti geografiche le apparecchiature di rete deputate al trasporto, alla commutazione ed all'elaborazione delle informazioni sono di proprietà di società pubbliche o private che installano e gestiscono i sistemi di TLC mettendoli a disposizione dell'utenza; nel caso della rete Internet ciò è vero solo in parte, in quanto alcune apparecchiature di utente e molte sotto-reti private contribuiscono in modo significativo al trasporto delle informazioni;
- Internet offre i suoi servizi a costi accessibili; la tariffazione è prevalentemente di tipo "flat rate", cioè non dipende dalla quantità dell'informazione scambiata o dalla distanza che separa gli utenti;
- il software che realizza i protocolli di comunicazione di Internet è stato fornito gratuitamente e risulta di semplice implementazione; inoltre,

inizialmente ed ancora oggi, tali protocolli sono stati distribuiti insieme al sistema operativo Unix. La grande diffusione di Unix, soprattutto nella comunità scientifica, ha trainato con se anche i protocolli Internet;

- altrettanto gratuito è spesso il software applicativo necessario per sfruttare appieno le capacità di comunicazione dell'infrastruttura; gli applicativi più utili e più facili da usare si diffondono in tal modo con grande rapidità, come è avvenuto in passato per applicazioni quali *Gopher* (applicativo per la ricerca di informazione) e, più recentemente, per il *WWW (World Wide Web)*, cioè per una applicazione multimediale di banca di dati distribuita.
- il processo di standardizzazione è snello, rapido e prevede delle verifiche implementative: si richiede che esistano almeno due diverse implementazioni interoperanti di un protocollo ed una solida esperienza in campo prima di dichiararlo standard;
- le idee di lavoro e le specifiche dei protocolli sono state sempre ampiamente e liberamente diffuse (anche e soprattutto per mezzo della stessa Internet); la disponibilità di informazioni generali e dettagliate sui protocolli, e delle relative implementazioni, anche nei primissimi stadi di sviluppo, ha svolto un ruolo essenziale nella loro diffusione. Un simile impegno verso la documentazione pubblica e gratuita a questo livello di dettaglio è inusuale. I benefici di questo modo di procedere hanno avuto significative conseguenze non solo sulla realizzazione di Internet ma anche, più in generale, sullo sviluppo delle telecomunicazioni.

La principale caratteristica della struttura attuale di Internet (una unione di diverse reti) è motivata da due considerazioni:

- nel passato è stato difficile far sì che un'unica tipologia di rete potesse rispondere in modo esauriente alle esigenze di comunicazione di ogni possibile utente. Ad esempio, le reti in area locale consentivano elevate velocità di trasferimento ma erano limitate geograficamente; il viceversa accadeva per le reti in area geografica. Inoltre la fisiologica esistenza di diversi produttori ha portato spesso a sviluppare sistemi di telecomunicazione non compatibili tra loro. Da ciò è derivato lo sviluppo di diversi paradigmi di comunicazione e la attuale co-esistenza di diverse reti di telecomunicazioni. Oggi è tecnicamente possibile realizzare un'unica infrastruttura di comunicazione che soddisfi i requisiti di tutti gli utenti in modo efficiente (è l'obiettivo della rete B-ISDN, basata sul modo di trasferimento ATM), ma non risulta economicamente vantaggioso sostituire tutti i sistemi già operativi;

- gli utenti desiderano una connettività universale. L'importanza e l'utilità di una rete di telecomunicazioni sono legate anche al numero dei suoi utenti. Ad esempio, la rete telefonica risulta molto utile proprio perché i suoi utenti sono numerosi (e quindi tramite essa è possibile comunicare con un grande numero di persone) e perché la rete appare ad essi come un solo sistema di comunicazione. La rete telefonica non avrebbe avuto lo sviluppo che conosciamo se fosse stata costituita da diverse sotto-reti con diversi standard e protocolli e non comunicanti tra loro. In tal caso infatti da un dato apparecchio telefonico non sarebbe stato possibile raggiungere tutti gli altri utenti ma solo un loro sotto-insieme.

In altre parole si desidererebbe una singola rete a cui tutti possano connettersi e tramite la quale raggiungere chiunque, ma è apparso sinora difficile realizzare una soluzione unitaria e quindi si è preferito rendere possibile la comunicazione tra diverse tipologie di reti.

Tali considerazioni hanno portato al concetto di *inter-rete*, e cioè di una infrastruttura *fisicamente* costituita da:

- un certo numero di reti componenti (nel seguito indicate come sotto-reti), cui sono collegati dei calcolatori (denominati nel seguito host);
- un certo numero di dispositivi necessari per interconnettere le sotto-reti componenti (denominati sistemi di interconnessione).

Da un punto di vista *logico*, l'elemento distintivo di una inter-rete è rappresentato dall'insieme di procedure di inter-lavoro (o protocolli di comunicazione) necessarie per interconnettere le sotto-reti e per permettere a qualunque host A connesso alla inter-rete di colloquiare con qualunque altro host B, anch'esso connesso alla inter-rete. La comunicazione tra A e B deve essere possibile indipendentemente dalle sotto-reti a cui A e B sono direttamente connessi e dal numero e dalla tipologia delle altre sotto-reti eventualmente coinvolte in tale comunicazione.

Internet è l'esempio più significativo di inter-rete oggi operativa su base mondiale. Sono da sottolineare le sue specificità. In primo luogo si vuole nascondere la struttura di rete ai suoi utenti. Questi ultimi e i programmi applicativi di loro interesse non devono preoccuparsi dei dettagli implementativi della rete fisica. Non si vuole imporre una pre-determinata *topologia* di rete. L'aggiunta di una nuova sotto-rete a Internet, e quindi alle sotto-reti pre-esistenti in Internet, non si effettua connettendo la nuova sotto-rete ad un nodo centrale o con l'aggiunta di connessioni fisiche dirette tra la nuova sotto-rete e *tutte* le altre pre-esistenti. E' sufficiente connettere la nuova sotto-rete ad *una* qualsiasi altra sotto-rete già connessa ad Internet. Le

informazioni sono quindi trasferite da una sotto-rete ad un'altra utilizzando, in generale, altre sotto-reti intermedie e attraversando diversi sistemi di interconnessione. Non vengono invece coinvolte sotto-reti intermedie qualora una sotto-rete che origina delle informazioni sia *direttamente* connessa alla sotto-rete cui queste informazioni sono dirette, ovvero quando è possibile trasferire informazione dall'una all'altra attraversando un solo sistema di interconnessione (cfr. par. **Errore. L'origine riferimento non è stata trovata.** e par. **Errore. L'origine riferimento non è stata trovata.**).

Una seconda specificità di Internet è che essa assicura l'indipendenza dell'interfaccia utente-rete dalle specificità della sotto-rete di accesso. L'insieme di operazioni necessarie per trasferire dati è indipendente dalla tecnologia della sotto-rete di accesso e dal sistema di destinazione. Un programma applicativo di comunicazione non deve tenere in conto la topologia e la struttura di sotto-rete.

I protocolli Internet si vanno ad aggiungere a quelli già esistenti all'interno delle sotto-reti componenti; questi ultimi non devono essere modificati. In tal modo alcune funzioni possono risultare duplicate e quindi si possono verificare delle inefficienze, ma questo modo di procedere determina notevoli semplificazioni operative.

Prima dell'avvento di Internet le reti per dati erano in gran parte diverse e non inter-comunicanti. Le sotto-reti di Internet sono rimaste diverse, nel senso che sono basate su diverse architetture e protocolli (e di per se autonome), ma i protocolli e i sistemi di interconnessione di Internet hanno fatto sì che potessero colloquiare tra loro.

Essendo Internet costituita da diverse tipologie di rete, ognuna, in generale, di proprietà di enti diversi, la sua modalità di gestione è necessariamente di tipo distribuito.

## **I.2 Architetture per applicazioni telematiche**

Tra le architetture di comunicazione che sono state finora sviluppate e che oggi (fine degli anni novanta) sono di attualità per applicazioni telematiche occorre menzionare il *modello OSI* e quello *Internet*, ambedue basati sui principi delle architetture a strati.

Il termine OSI, acronimo che fa riferimento alla "interconnessione di sistemi aperti" (*Open System Interconnection*), è correntemente utilizzato per definire la comunicazione tra sistemi di elaborazione che adottano un insieme di norme (le *norme OSI*), tali da consentire la cooperazione indipendentemente

dalla natura dei sistemi e quindi con lo scopo di rendere possibile l'interconnessione e il colloquio tra sistemi eterogenei (cioè tra sistemi che hanno origine da costruttori diversi), oltre che tra sistemi omogenei.

Il modello OSI è il risultato dell'attività di normalizzazione svolta dall'*ISO (International Standard Organization)* e fatta propria dall'*ITU-T (International Telecommunication Union - Telecommunication Standardization Sector)*. Lo sviluppo del modello OSI è iniziato alla fine degli anni settanta ed è tuttora in corso per la grande varietà degli aspetti ad esso connessi. Le linee-guida seguite dall'*ISO* e dall'*ITU-T* nella definizione del modello OSI sono state:

- disporre di norme la cui traduzione in prodotti fosse condizionata da vincoli tecnologici solo in modo debole;
- prevedere una estendibilità del modello e delle relative implicazioni normative verso evoluzioni future;
- comprendere una larga varietà di opzioni come compromesso tra esigenze spesso contrastanti, in quanto legate alla difesa di interessi industriali tra loro in competizione.

Il risultato è stato la produzione di norme, che sono ineccepibili dal punto di vista della generalità e della correttezza formale, ma che si sono dimostrate talvolta di difficile realizzazione pratica.

Il modello Internet ha invece avuto come origine l'organizzazione *ARPA (Advanced Research Project Agency)*, una emanazione del Dipartimento della Difesa negli Stati Uniti d'America. Attualmente lo sviluppo, il coordinamento e la gestione del mondo Internet fanno capo all'*IAB (Internet Architecture Board)*. Scopo principale dell'*IAB* è stato ed è tuttora l'adozione di soluzioni tecniche disponibili, che assolvano il loro compito e che consentano realizzazioni basate su posizioni consolidate. Ciò ha consentito di ottenere velocemente prodotti in linea con la normativa Internet e di adottare in modo tempestivo gli aggiornamenti necessari per rispondere a nuove esigenze.

La differenza tra i due modelli è dovuta sostanzialmente allo scopo diverso che le organizzazioni preposte si sono prefisse. Da un lato, con la modellistica OSI, gli organismi di normalizzazione formali (*ISO* e *ITU-T*) hanno puntato a definire un quadro normativo che fosse senza limitazioni circa la sua diffusione virtuale e che comprendesse ogni possibile accorgimento anche in vista degli aggiornamenti prevedibili. Dall'altro lato l'*IAB* ha cercato di rispondere in modo più rapido alle esigenze, anche a breve termine, dell'utente finale, fornendo quindi risposte concrete alle domande esistenti.

È il caso di sottolineare che, in un ambiente di comunicazione di dati, l'architettura del modello OSI e di quello Internet non sono le uniche che

possono essere derivate dai criteri architetturali esposti, ad es., in [1,9]. A criteri analoghi sono infatti ispirate altre architetture sviluppate dai singoli costruttori (*architetture proprietarie*), quali per es. la *System Network Architecture* (SNA) della IBM e la *Digital Network Architecture* (DNA) della DEC.

### 1.2.1 Il modello OSI

L'identificazione degli strati che compongono l'architettura OSI è stata effettuata in base al soddisfacimento delle seguenti regole generali:

- funzioni simili sono state raggruppate in uno stesso strato;
- funzioni diverse, sia per logica che per tecnologia impiegata nella loro realizzazione, sono state collocate in strati diversi;
- l'interfaccia tra due strati è stata stabilita in modo da minimizzare la complessità ed il numero di interazioni tra gli strati stessi.

L'applicazione di questi principi ha portato alla definizione di sette strati funzionali: lo strato di *applicazione* (strato 7); lo strato di *presentazione* (strato 6); lo strato di *sessione* (strato 5); lo strato di *trasporto* (strato 4); lo strato di *rete* (strato 3); lo strato di *collegamento* (strato 2); lo strato *fisico* (strato 1). L'insieme degli strati ora citati e la loro struttura gerarchica sono mostrati in Fig. I.3.

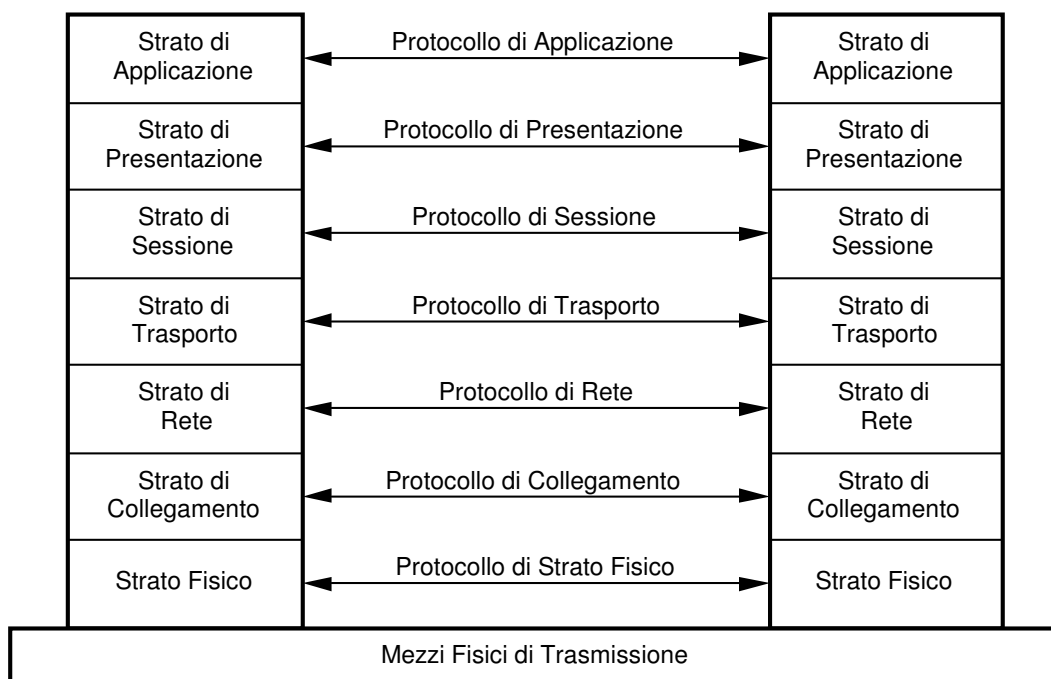


Fig. I.3 - Architettura del modello OSI.

Con l'eccezione dello strato fisico, che s'interfaccia direttamente con il mezzo trasmissivo, i servizi degli strati del modello OSI sono stati definiti nelle

versioni sia con connessione che senza connessione. Qui di seguito vengono esposte in modo sintetico le finalità dei sette strati.

Lo *strato di applicazione* ha lo scopo di fornire ai processi applicativi residenti in un sistema i mezzi per accedere all'ambiente OSI, facendo sì che quest'ultimo costituisca una macchina virtuale in grado di associare un processo applicativo residente in un sistema con qualsiasi altro processo applicativo residente in un qualsiasi altro sistema remoto. Poiché gli utenti dei servizi offerti dallo strato di applicazione possono essere processi applicativi di qualsiasi natura, risulta arduo individuare un insieme di funzioni sufficientemente completo e generale che si adatti a tutte le possibili applicazioni. Per la descrizione delle funzioni interne allo strato di applicazione si segue quindi la strada di individuare, dapprima, elementi di servizio comuni a tutte le applicazioni e, in seguito, moduli disgiunti comprendenti elementi di servizio orientati a determinate classi di applicazioni.

Lo *strato di presentazione* ha due compiti fondamentali:

- rendere direttamente disponibili alle entità di applicazione i servizi dello strato di sessione relativi alla strutturazione, alla sincronizzazione e alla gestione del loro dialogo;
- risolvere i problemi di compatibilità tra due qualsiasi entità di applicazione per quanto riguarda la *rappresentazione dei dati* da trasferire e il modo di riferirsi a strutture di dati del sistema remoto.

In altre parole lo strato di presentazione ha lo scopo di sollevare le entità di applicazione da qualsiasi compito relativo alla *trasformazione della sintassi dei dati*. In generale, all'interno dello strato di presentazione esistono tre tipi di sintassi: due *sintassi locali*, utilizzate dalle singole entità di applicazione, e una *sintassi di trasferimento* utilizzata per lo scambio informativo vero e proprio. La sintassi di trasferimento è negoziata dalle entità di applicazione, mentre la conversione tra sintassi locali e sintassi di trasferimento è effettuata dallo strato di presentazione.

Lo scopo dello *strato di sessione* è quello di assicurare alle entità di presentazione una risorsa logica per organizzarne il colloquio. Ciò ha lo scopo di strutturare e di sincronizzare lo scambio dati in modo da poterlo sospendere, riprendere e terminare ordinatamente. Lo strato di sessione presuppone che siano risolti tutti i problemi relativi al trasporto dell'informazione tra i sistemi. Il suo compito consiste invece nel mascherare, per quanto possibile, eventuali interruzioni del servizio di trasporto e nel mantenere una continuità logica nell'evoluzione del colloquio tra entità di presentazione.

Lo *strato di trasporto* fornisce alle entità di sessione una risorsa virtuale

per il trasferimento trasparente delle unità di dati. Deve quindi, da un lato, colmare eventuali deficienze e fluttuazioni della qualità di servizio offerto dallo strato di rete e dall'altro, ottimizzare l'uso di questo servizio di rete in modo da poter garantire, al costo minimo, le prestazioni richieste. Procedendo dal basso verso l'alto nell'esame dell'architettura OSI, lo strato di trasporto è il primo che abbia esclusivamente un *significato da estremo a estremo*. In altre parole, mentre le funzionalità dei primi tre strati dell'architettura possono essere presenti sia nei sistemi terminali che nei nodi della rete di comunicazione, le procedure relative allo strato di trasporto sono eseguite solo dai sistemi terminali. Infatti, tutti i problemi relativi all'accesso dei sistemi terminali alla rete di comunicazione, all'instradamento e all'utilizzazione di eventuali reti di transito si considerano completamente risolti dallo strato di rete.

Lo *strato di rete* rende invisibile allo strato di trasporto il modo in cui sono utilizzate le risorse di rete, in particolare se sono utilizzate diverse sottoreti in cascata. L'unico elemento che è visibile all'interfaccia con lo strato di trasporto è la qualità di servizio complessiva offerta dallo strato di rete. Se il trasferimento interessa due sottoreti che offrono differenti qualità di servizio, la cooperazione tra le sottoreti può avvenire secondo due diverse modalità:

- le reti cooperano senza l'esecuzione di nessuna funzione addizionale; in questo caso la qualità di servizio risultante sarà sicuramente non superiore a quella corrispondente alla rete con minore qualità di servizio;
- sulla rete con qualità di servizio inferiore sono eseguite funzioni che hanno l'obiettivo di innalzare tale qualità di servizio allo stesso livello della seconda rete; in questo caso la qualità complessiva può essere uguale a quella della sottorete a più alta qualità.

La scelta tra queste due alternative dipende dalla differenza di prestazioni delle due sottoreti e da fattori di costo.

La funzione fondamentale eseguita dallo *strato di collegamento* è quella di *rivelazione e di recupero degli errori trasmissivi* verificatisi durante il trasferimento fisico. Tale funzione dovrebbe garantire allo strato di rete la fornitura di un servizio di trasferimento dell'informazione esente da errori.

Lo *strato fisico* ha il compito principale di effettuare il trasferimento fisico delle cifre binarie scambiate tra le entità di collegamento. Le funzioni eseguite al suo interno e i servizi forniti allo strato superiore hanno lo scopo di assicurare l'indipendenza della comunicazione dalle caratteristiche del particolare mezzo trasmissivo utilizzato.

### *I.2.2 Il modello Internet*

L'*architettura* del modello Internet può essere ricollegata ai concetti introdotti nel paragrafo I.1.5.

Partendo dal livello gerarchicamente più basso, comprende quattro strati.

Un primo strato, detto di *accesso alla rete* (Network Access Layer), consente la utilizzazione di risorse infrastrutturali (sotto-reti), tra loro anche non omogenee per tecnica realizzativa. Lo strato di accesso alla rete include le funzioni che nel modello OSI sono comprese negli strati fisico, di collegamento e di rete, quest'ultimo almeno per ciò che riguarda gli aspetti connessi al funzionamento di ogni singola rete componente (*sottostrato di rete basso*).

Un secondo strato, che è denominato *IP (Internet Protocol)* dal nome del protocollo che gli è proprio, consente l'interfunzionamento delle varie reti componenti con funzionalità che nel modello OSI sono collocabili in un *sottostrato di rete alto*. Il servizio di strato corrispondente è senza connessione.

Un terzo strato corrisponde allo strato di trasporto del modello OSI e viene denominato in base al protocollo che gli è proprio. Un primo tipo di protocollo di questo strato è il *TCP (Transmission Control Protocol)*, nell'ambito del quale il servizio di strato è con connessione. Un protocollo alternativo è l'*UDP (User Datagram Protocol)*, che a differenza del precedente è senza connessione;

Il quarto e ultimo strato del modello Internet è quello applicativo e corrisponde ai tre ultimi strati del modello OSI. Viene denominato strato dei Servizi Applicativi (Application Services)

## **I.3 Le applicazioni e i servizi OSI**

Vengono qui di seguito descritte alcune tra le applicazioni che sono state definite secondo il modello architetturale OSI.

### *I.3.1 Applicazioni telematiche*

*Posta elettronica* - Sviluppato nell'ambito del modello OSI, il *trattamento di messaggi* (MHS, Message Handling System) definisce un insieme di funzioni per l'invio di messaggi su reti di telecomunicazione pubbliche.

L'applicazione MHS è un tipico esempio di applicazione distribuita: essa implica l'esistenza di una rete di nodi intelligenti, ognuno dei quali cioè include tutti gli strati dell'architettura OSI. Ogni nodo è in grado di effettuare, per i messaggi ricevuti, una memorizzazione temporanea (*immagazzinamento*) e un inoltro verso un altro nodo più vicino al destinatario (*rilancio*). Una rete siffatta include: a) nodi di accesso, chiamati *agenti di utilizzazione (UA, User Agent)*,

che consentono la redazione, l'invio, la ricezione e la lettura di messaggi; b) nodi di transito, chiamati *agenti di trasferimento* (*MTA, Message Transfer Agent*), che sono in grado di interconnettersi con agenti di utilizzazione e con altri agenti di trasferimento.

L'applicazione MHS ha lo scopo di soddisfare tipiche esigenze di automazione d'ufficio, quali le possibilità di: 1) indirizzare i destinatari "per nome" invece che per numero, attraverso una funzione di *annuario*, realizzabile per mezzo di una banca di dati distribuita a estensione internazionale ; 2) specificare una lista di destinatari per uno stesso messaggio (*lista di distribuzione*); 3) assicurare l'identità di chi emette il messaggio (*autenticazione*); 4) essere informati circa l'avvenuta ricezione da parte del destinatario (*avviso di recapito*); 5) scambio di messaggi tra utenti diretti dell'MHS e utenti Teletex, Telefax e Videotex.

L'MHS può essere usato come base per la fornitura di servizi ulteriori rispetto alla messaggistica interpersonale. Come applicazione principale si può citare la sua possibile utilizzazione per *applicazioni EDI (Electronic Data Interchange)* e cioè per transazioni di tipo commerciale quali fatture e ordini d'acquisto.

L'applicazione che compete con MHS e che è stata sviluppata secondo il modello Internet è *SMTP (Simple Mail Transfer Protocol)*.

*Trasferimento di archivi.* - Questa applicazione (*FTAM, File Transfer, Access and Management*) definisce in ambiente OSI un insieme di funzioni necessarie a garantire il corretto scambio di informazione, memorizzata sotto forma di un archivio (file), tra processi applicativi residenti in sistemi remoti. Essa consente: 1) l'accesso ad archivi, per scrivere o per leggere informazioni; 2) il trasferimento completo o parziale di archivi; 3) la gestione di archivi per la loro creazione, la loro cancellazione e la modifica delle loro caratteristiche.

Poiché in genere ogni sistema di elaborazione usa modalità specifiche per descrivere e per organizzare i dati su memoria di massa, il fondamento dell'applicazione FTAM è un modello comune per la descrizione e per la rappresentazione di questi dati, e cioè una *memoria di massa virtuale* contenente *archivi virtuali*, ciascuno identificato da una serie di attributi quali nome, lista di parole chiave, storia, organizzazione, metodo di accesso, tipo dei dati contenuti.

Compito principale dell'applicazione FTAM è assicurare trasparentemente la corrispondenza tra la memoria di massa reale del sistema di elaborazione ove risiede l'archivio e la sua rappresentazione virtuale che è l'unica resa nota a un

processo applicativo remoto.

Il corrispondente di FTAM in ambiente Internet è FTP (*File Transfer Protocol*).

*Terminale virtuale.* - Questa applicazione (VT, *Virtual Terminal*) ha per obiettivo consentire in un ambiente OSI la gestibilità completa di qualsiasi tipo di terminale intelligente da parte di qualsiasi tipo di elaboratore. Essa si basa sulla definizione di vari modelli astratti di terminali, rappresentabili mediante opportune strutture di dati dalle quali si può derivare lo stato e il contenuto del terminale fisico.

Le principali proposte di terminali virtuali prevedono: 1) *modo a scorrimento (scroll mode)*, in cui la corrispondente struttura di dati è un vettore di elementi (*linee*) senza alcun attributo; 2) *modo a pagina (page mode)*, in cui la struttura di dati è una matrice bidimensionale ( $n$  righe per  $m$  caratteri) i cui elementi hanno l'attributo "protetto" o "non protetto"; 3) *modo a maschera (form mode)*, in cui la struttura di dati è un insieme di campi ognuno con attributi quali "tipo del contenuto", "coordinate del carattere iniziale del campo", "lunghezza del campo", "luminosità", "sottolineatura", "sola lettura".

L'applicazione corrispondente a VT in ambito Internet è "*telnet*".

*Annuario.* - È un'applicazione (*Directory*) di banca di dati distribuita relativa a persone, organizzazioni, sistemi e applicazioni in ambiente OSI. L'annuario contiene quelle informazioni che consentono una comunicazione tra persone e/o processi applicativi, e cioè numeri telefonici o di altre reti, indirizzi di posta elettronica, indirizzi residenziali.

Tra i servizi che l'annuario offre si possono citare la ricerca selettiva per nomi o per categorie e la risoluzione di nomi facilmente memorizzabili in indirizzi di vario tipo. Il controllo dell'accesso alla banca di dati secondo funzioni di sicurezza consente la creazione di diversi profili di "privilegi", che consentono, per es., l'accesso ai dati in scrittura solo agli utenti autorizzati.

Si noti che l'applicazione ora descritta non ha competitori in ambiente Internet.

*Trasferimento di documenti.* - Questa applicazione (*DTAM, Document Transfer and Manipulation*) è definita come la piattaforma comune per il trasferimento da estremo a estremo di applicazioni OSI e di servizi telematici. È strutturata secondo l'ottica di fornire funzionalità di trasferimento, di manipolazione e di gestione del documento elettronico in linea con i concetti

espressi dall'*ODA (Open Document Architecture)*. Ciò consente la definizione completa, nei contenuti e nei formati, di un qualsiasi documento.

Secondo DTAM, un documento è una collezione di informazioni di varia natura (testo, tabelle, diagrammi, grafici, disegni, voce e suoni) strutturate in modo da essere utilizzabili da un utente umano e rappresentabili in un sistema di elaborazione. L'*ODA* definisce due differenti strutture di documento: una *logica* e una *fisica*.

La struttura logica è orientata agli oggetti logici, e cioè agli elementi specifici nell'organizzazione concettuale del documento, e, in quanto tali, dipendenti dalla particolare natura del documento. Data la stretta dipendenza degli oggetti logici dal tipo dell'applicazione telematica, l'*ODA* non predefinisce alcun insieme di oggetti logici.

La struttura fisica è orientata agli oggetti di formato, e cioè alle possibili forme di organizzazione dell'informazione su un supporto fisico di uscita leggibile dall'operatore umano (video, carta, ecc.).

Se di un documento è disponibile solo la struttura fisica, esso viene presentato all'utente nella sua forma originale o nella sua forma finale senza possibilità di apportarvi modifiche. Se di un documento è disponibile solo la struttura logica, esso viene visualizzato in una forma eventualmente approssimata ai fini della percezione visiva ma tale da consentire l'accesso e la modifica di ogni singolo oggetto logico. Se di un documento sono disponibili entrambe le strutture, esso viene presentato all'utente esattamente nella sua forma originale o finale ma con la possibilità di accesso e di modifica di ogni singolo oggetto logico.

### *I.3.2 I servizi telematici*

Nella terminologia degli organi di normalizzazione si chiama *servizio telematico* una applicazione OSI offerta da un fornitore operante sulla rete pubblica e accessibile tramite un terminale con caratteristiche funzionali e di interfaccia appositamente normalizzate. Sono stati definiti e sono in esercizio da vari anni alcuni servizi telematici quali il *Teletex*, il *Videotex* e il *Telefax*. Per questi vengono date alcune sommarie notizie.

*Teletex*. - È un servizio per il trasferimento di testi da terminale a terminale, originariamente concepito come miglioramento del servizio *Telex*. Rispetto a quest'ultimo, presenta prestazioni migliorate in termini di ritmo di trasferimento (almeno 2400 bit/s), in termini di reti di trasporto utilizzabili (anche reti a pacchetto e non solo a circuito) e in termini di funzionalità del

terminale (per es., migliorate possibilità di composizione del documento da trasferire e di riproduzione di documenti ricevuti, un più vasto repertorio di caratteri).

*Videotex.* –E' un servizio di accesso a banche di dati di interesse generale (per es. orari dei mezzi pubblici di trasporto, notizie di cronaca, spettacoli, ecc.) per utenti residenziali o affari collegati a una rete di telecomunicazione pubblica. Le informazioni reperite sono visualizzate in forma sia alfa-numerica che grafica, utilizzando un repertorio di attributi di visualizzazione.

Le informazioni, visualizzate per pagine, vengono acquisite con una ricerca ad albero. L'utente parte da una prima pagina che gli offre un menu di possibilità. Ogni scelta successiva comporta la visualizzazione di una nuova pagina, la quale espande l'indice dei contenuti; finché, così procedendo, l'utente arriva all'informazione desiderata. Questa organizzazione ad albero comporta un tempo di accesso piuttosto lungo, ma dà la possibilità all'utente di accedere a un volume di informazioni teoricamente illimitato.

Al termine *Videotex*, che è la dizione ufficiale in seno agli organi di normalizzazione internazionali, corrispondono nei vari paesi nomi commerciali specifici: per es. *Prestel* in Gran Bretagna, *Minitel* in Francia, *Videotel* in Italia, *Bildschirmtext* in Germania Federale.

Il servizio offre anche le seguenti prestazioni: immissione di informazioni (per es. compilazione di questionari), transazioni interattive di vario tipo (per es. prenotazioni di spettacoli, di alberghi, di posti su mezzi di trasporto, trasferimento elettronico di fondi).

*Telefax.* - È un servizio per la riproduzione a distanza, normalmente in bianco e nero, di immagini fisse, che possono contenere dattiloscritti, manoscritti, disegni, grafici e fotografie (*documenti*). In questo caso la rappresentazione del documento è effettuata con tecnica a reticolo (*raster*), che simula il funzionamento dell'occhio nell'acquisizione di un'immagine.

Il trasferimento di una immagine è effettuato, per il tramite di una rete di telecomunicazione, emettendo un flusso informativo che rappresenta il risultato di una scansione ottica dell'originale effettuata nell'apparato emittente. Il contenuto informativo del documento è tradotto dal sistema di scansione in un segnale elettrico analogico o numerico. Questo, dopo opportune elaborazioni (codifica e modulazione) viene trasmesso al terminale ricevente, dove subisce le operazioni complementari e viene riconvertito in immagine da un sistema di registrazione. Estremamente importanti nel processo di registrazione sono la

*sincronizzazione* del dispositivo di registrazione e di quello di scansione, nonché il *mantenimento della relazione di fase* tra i due dispositivi per tutta la durata della trasmissione.

L'ITU-T ha definito tre gruppi di apparati *Telefax* da usare con la rete telefonica pubblica (*gruppi 1, 2 e 3*) e un gruppo da impiegare sulle reti per dati (*gruppo 4*). Mentre i gruppi 1 e 2 hanno oggi ormai solo un interesse storico, nei gruppi 3 e 4 il segnale emesso è in forma numerica. Inoltre, mentre il gruppo 3 non prevede alcun meccanismo di controllo degli errori di trasmissione e pertanto la copia ricevuta può presentare difformità dall'originale, il gruppo 4, che è quello raccomandato a lungo termine, utilizza procedure di controllo del trasferimento che garantiscono una riproduzione fedele del documento.

Sono in fase di commercializzazione anche terminali capaci di operare nel cosiddetto *modo misto* (Teletex-Telefax gruppo 4) in cui l'informazione organizzata a reticolo è combinata con quella organizzata a caratteri. Il modo misto, è talvolta considerato un quarto servizio telematico, in aggiunta ai tre qui descritti.

È da rilevare infine che il gruppo 4 è funzionalmente strutturato in tre classi: 1) *classe 1*: solo invio di informazioni del tipo a reticolo; 2) *classe 2*: può emettere immagini del tipo a reticolo e ricevere informazioni con modo misto; 3) *classe 3*: emette e riceve informazioni con modo misto.

#### **I.4 Le applicazioni Internet**

Nell'ambiente Internet, i protocolli di strato superiore al 4 sono denominati Servizi Applicativi e le loro funzioni non sono divise in termini di strato di sessione, presentazione ed applicazione.

I Servizi Applicativi in Internet operano prevalentemente in accordo al paradigma *client/server*; hanno cioè una componente client ed una server. Tale paradigma è importante anche perché è alla base di molti sistemi di comunicazione ed è fondamentale per comprendere le modalità di funzionamento di algoritmi distribuiti.

Il termine server si applica a qualunque processo che offre un servizio e che può essere raggiunto attraverso la inter-rete. I server accettano le richieste, le elaborano, effettuano il servizio richiesto e restituiscono il risultato al richiedente. Un processo applicativo è invece un client quando invia una richiesta ad un server ed aspetta la relativa risposta. Ad esempio, un semplice server "ora esatta" fornisce l'ora attuale ogni volta che un client gli invia una richiesta. I server gestiscono le risorse e i client richiedono servizi alle risorse.

La durata in attività delle componenti client e server è diversa; un server inizia ad essere attivo prima che l'interazione con il client inizi e continua ad accettare richieste e ad inviare risposte senza cessare di essere attivo; la componente client si attiva al momento di inviare una richiesta e si disattiva una volta ricevuta la risposta. Normalmente la componente server inizia ad operare al momento dell'accensione di un host e quando invocata, fornisce alla componente client i mezzi e le procedure o i dati di cui questa ha bisogno.

Molti dei servizi applicativi richiedono un'autorizzazione all'uso: un utente può usare un servizio applicativo se dispone di una utenza ("account"), concessa dall'amministratore del sistema remoto (server), a cui accede mediante una parola d'ordine ("password"). In tal caso, la componente client deve comunicare a quella server il nome dell'utenza e la relativa parola d'ordine, prima di poter richiedere uno specifico servizio.

In questo paragrafo si descriveranno i più importanti e diffusi servizi applicativi che la comunità Internet mette a disposizione dell'utenza. Ognuno di tali servizi opera in accordo ad uno specifico protocollo che governa lo scambio di informazioni tra la componente client e quella server. Inoltre, un dato servizio applicativo può essere implementato nell'ambito di diverse applicazioni, basate su specifiche piattaforme operative e personalizzate o arricchite da interfacce di utente e più o meno sofisticate.

### *TELNET*

La componente client di TELNET è un protocollo di emulazione di terminale che permette agli utenti di un sistema di accedere ad applicazioni in altri sistemi remoti, come se fossero direttamente connessi a tali sistemi. Consente quindi di utilizzare le funzioni di un calcolatore remoto come se il terminale che si usa fosse uno dei terminali locali del sistema remoto. L'host di destinazione deve contenere la componente server di TELNET. Il procedimento può essere iterativo; ovvero ci si può collegare ad un host remoto e, da lì collegarsi ad un altro host e così via.

L'applicativo Telnet consente di aprire sessioni interattive (login remoto) su host ovunque ubicati all'interno della inter-rete purché si disponga di un'autorizzazione all'uso (utenza + parola d'ordine). Il terminale emulato è di tipo testuale.

### *SMTP (SIMPLE MAIL TRANSFER PROTOCOL)*

La posta elettronica (E-mail) è uno dei servizi applicativi più usati in Internet. La sua popolarità deriva dalla facilità di uso che consente di trasferire

velocemente sia piccole note che voluminosi documenti. I vantaggi sono ovvi: il messaggio ricevuto può essere registrato e catalogato, inoltre può essere modificato o usato come base per successivi lavori, al contrario di un messaggio ricevuto su carta o a voce che necessita di dover essere trascritto. Inoltre non richiede che mittente e destinatario siano contemporaneamente presenti. Un utente può inviare un messaggio ed il destinatario può leggerlo in un tempo successivo. Gli applicativi di posta elettronica permettono di scambiare messaggi tra utenti utilizzando il protocollo di comunicazione SMTP.

L'obiettivo che il protocollo SMTP si propone è quello di offrire i meccanismi per trasferire messaggi in forma elettronica in modo affidabile ed efficiente attraverso Internet. A tal fine, SMTP usa il protocollo TCP come protocollo di strato 4 e quindi adotta una modalità di trasferimento con connessione ed affidabile.

A fronte di una richiesta di invio di e-mail da parte dell'utente, il "sender-SMTP" stabilisce una connessione bi-direzionale con il "receiver-SMTP". La connessione è bi-direzionale, in quanto il servizio è di tipo confermato, cioè il destinatario invia al mittente dei messaggi di riscontro per confermare la corretta ricezione del messaggio inviato dall'utente.

Preliminarmente all'invio del messaggio vero e proprio il "sender-SMTP" verifica la disponibilità del "receiver-SMTP" ad accettare il messaggio stesso e, se non vengono riscontrati problemi (indisponibilità del receiver, utente sconosciuto, etc.), si avvia il processo di trasmissione del messaggio generato dall'utente.

Il formato dei messaggi è regolato da uno standard (RFC822). SMTP prevede anche funzionalità di traduzione tra diversi formati e l'uso di gateway verso sistemi di E-mail non basati su TCP/IP. Ha diverse potenzialità: uso di alias, liste di destinatari, risposte automatiche, etc.

### *FTP: FILE TRANSFER PROTOCOL*

E' un altro dei servizi applicativi più usati ed è responsabile di una rilevante quantità del traffico totale. La principale funzione del protocollo FTP è il trasferimento efficiente ed affidabile di archivi (files) tra host remoti. Consente anche di visualizzare il contenuto di sistemi di archiviazione remoti e di modificare o cancellare files e directories ivi residenti.

Supporta il trasferimento di quattro tipi di files: Binari, ASCII, EBCDIC e "paged". Fornisce funzionalità di protezione dei file trasmessi e compressione di dati; non fornisce funzionalità di traduzione o di conversione di files.

Nell'effettuare il trasferimento di un file, la componente client del

protocollo FTP realizza innanzitutto una connessione di controllo con la componente server. Attraverso questa connessione vengono inviati al server FTP i comandi, opportunamente codificati, che originano da richieste dell'utente. Il server FTP, interpreta tali comandi ed effettua la trasmissione dei dati richiesti. Per garantire l'affidabilità dei dati trasmessi usa, come protocollo di strato 4, TCP.

#### *TFTP: TRIVIAL FTP*

E' un protocollo simile a FTP, trasferisce archivi (files) tra host remoti, ma fornisce meno funzionalità di FTP: non consente di proteggere i files trasmessi e di gestire directories. Usa UDP invece di TCP come protocollo di strato di trasporto.

#### *TALK*

Consente il dialogo in tempo reale tra due utenti. Genera, sullo schermo degli utenti, una finestra divisa in due parti; in una compare ciò che si scrive mentre nell'altra compare ciò che scrive l'utente remoto.

#### *X-WINDOW*

E' simile a Telnet, ma fornisce anche sofisticate funzionalità grafiche. Consente ad un utente di accedere contemporaneamente a diverse applicazioni in altri sistemi come se fosse direttamente connesso a tali sistemi.

La principale differenza rispetto a Telnet è che il terminale emulato è di tipo grafico e quindi consente di usare anche applicazioni remote che fanno uso di visualizzazioni non solo testuali.

#### *NFS: NETWORK FILE SYSTEM*

Consente di utilizzare un sistema di archiviazione (file system) remoto allo stesso modo di uno locale. Grazie a NFS è possibile vedere, copiare, modificare e cancellare files residenti in un sistema remoto in modo del tutto trasparente per l'utente. A differenza di FTP, con NFS l'utente può anche ordinare l'esecuzione di un qualunque processo ed usare qualsivoglia file (locale o remoto) per l'input e l'output di tale processo. Un'altra caratteristica di NFS è che consente all'utente finale di eseguire tutte le azioni appena descritte mediante gli stessi comandi del sistema operativo che l'utente usa nel proprio sistema, per le stesse azioni riferite al sistema locale. Non è necessario quindi che l'utente debba usare in modo esplicito il tramite di un'applicazione di rete per interagire con un sistema remoto; NFS gestisce tutte le operazioni

necessarie per tali interazioni senza che l'utente finale ne abbia coscienza ed opera, per usare un termine gergale, "nello sfondo" o "in ombra" (in "background"). Nella Fig. I.4 le directories ed i files dei sistemi remoti "berkeley ed "osaka" sono accessibili come se fossero residenti sul sistema locale, a meno solo di eventuali limitazioni nella velocità di accesso. Sempre in gergo, si dice che i file systems remoti sono "montati" sul sistema locale.

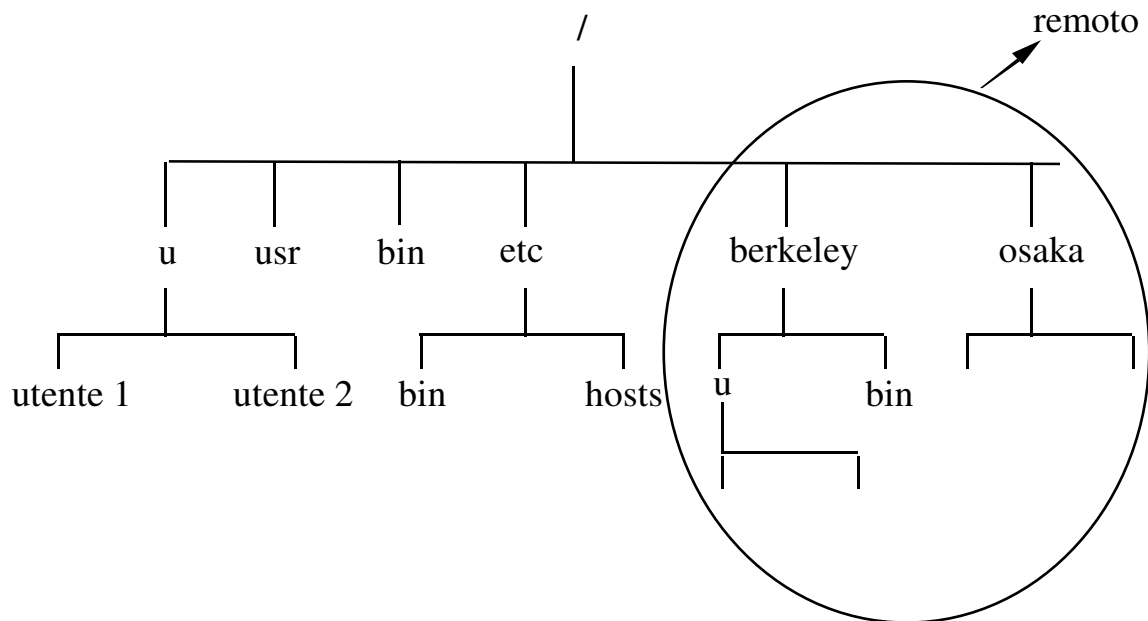


Fig. I.4 – Esempio di funzionamento di NFS

### **WORLD WIDE WEB**

World Wide Web, chiamato anche WWW o W3, è un sistema informativo multimediale basato su ipertesti la cui origine risale al 1989. E' attualmente il servizio applicativo più usato in Internet. Consente ad un utente di acquisire un documento situato in un sistema remoto, usando la componente client di WWW (denominata "browser"). In ogni host in cui siano contenuti dei documenti deve invece operare la componente server. La particolarità di WWW è la tipologia dei documenti acquisibili.

I documenti WWW sono:

- multimediali e cioè contengono informazioni testuali, grafiche e sonore, eventualmente compresse mediante opportuni algoritmi per ridurre il volume di traffico scambiato durante il trasferimento;
- ipertestuali: alcune parole o immagini o zone del documento acquisito contengono dei collegamenti ad altri documenti; l'utente, selezionando uno di questi collegamenti (link), acquisisce il relativo documento.

Quest'ultima caratteristica è quella che ha determinato lo sviluppo attuale del WWW. Un particolare documento può contenere diversi collegamenti, ognuno dei quali è un puntatore ad un altro documento; ogni documento può essere situato in un qualunque host connesso ad Internet. L'utente, sfogliando il documento in questione, può scegliere di seguire uno qualsiasi di questi collegamenti. L'aspetto più comodo, che rende il WWW particolarmente semplice da usare, è che ogni collegamento contiene anche l'indirizzo dell'host in cui il documento cercato si trova. In tal modo, l'utente può acquisire diversi documenti, localizzati ovunque nel mondo, senza bisogno di sapere dove essi siano fisicamente memorizzati. Il processo può divenire iterativo per cui da un documento l'utente può passare ad esaminare un altro documento che contiene collegamenti ad altri documenti ancora e così via.

Di fatto l'intero insieme del WWW (e cioè tutti i documenti accessibili di tipo WWW) appare all'utente, nella sua interezza, come se fosse disponibile nel suo host, a meno di eventuali limitazioni nella velocità di accesso; l'utente non è tenuto a sapere dove sia localizzato uno specifico documento, e non c'è necessità che digiti comandi o interroghi basi di dati alla ricerca di tale locazione.

Ad esempio, quando una nuova parola o un nuovo concetto sono introdotti in un testo, WWW offre i mezzi per collegarsi ad un altro documento in cui vengono forniti maggiori dettagli a riguardo, in modo tale che il lettore può aprire il secondo documento selezionando la parola o il concetto sconosciuto; questo documento può a sua volta contenere ulteriori riferimenti ad altra documentazione residente in altri sistemi. E' ciò che si intende per "navigare" nella rete, seguendo un "albero" di documenti che si può biforcare in ogni momento.

L'accesso ai server WWW avviene tramite applicativi di tipo client. Questi applicativi sono disponibili per diverse piattaforme software. Gli esempi più noti e diffusi di tali applicativi sono Netscape e Internet Explorer. Spesso le applicazioni che implementano la componente client di WWW contengono anche componenti client di altri servizi applicativi, come ad esempio FTP o News. Tali applicazioni contengono inoltre ulteriori moduli software (non servizi di rete, ma applicazioni locali) necessari per:

- convertire il formato di alcuni documenti;
- decomprimere documenti codificati con opportuni algoritmi per ridurre il volume di traffico scambiato;
- presentare in modo adeguato l'informazione (ad es. video o audio) all'utente finale.

Il servizio applicativo WWW si basa su due protocolli:

- Hyper Text Markup Language (html): definisce il formato dei documenti;
- Hyper Text Transfer Protocol (http): è il protocollo che consente di trasferire i documenti.

Il protocollo HTTP usa come protocolli di strato di trasporto sia TCP che UDP, instaurando contemporaneamente anche più di una connessione TCP per il trasporto di diverse parti del documento.

### *GOPHER*

Internet Gopher, o più semplicemente Gopher, è un servizio distribuito per l'acquisizione di documenti, che consente di esplorare, cercare e acquisire informazioni residenti in differenti locazioni. E' simile a WWW ma a questi precedente e sta diventando obsoleto, poiché offre meno funzionalità. L'interfaccia tipica di questo servizio suddivide le informazioni in una serie di menu nidificati, che rispecchiano l'organizzazione delle informazioni in directory, sub-directory e file.

Dal punto di vista dell'utente, tutte le informazioni sembrano essere dislocate nello stesso posto, sebbene alcune sub-directory e il loro contenuto possano trovarsi fisicamente su dei sistemi remoti gestiti da altri servers Gopher. Tale struttura gerarchica delle informazioni porta alla definizione di uno "spazio Gopher", cioè di una struttura logica unitaria composta da informazioni contenute in server gopher ubicati spazialmente in luoghi diversi.

La tipologia delle informazioni gestite da questo servizio comprende file di testo, file binari, immagini e suoni.

### *ARCHIE*

Archie è un sistema informativo nato per offrire un servizio di directory elettronico centralizzato, mediante il quale è possibile localizzare le informazioni nell'ambito della rete Internet; allo stato attuale si contano un centinaio di siti distinti che offrono il servizio.

L'uso principale che viene fatto di tale sistema informativo è quello di interrogare una base di dati di carattere generale, che contiene informazioni su host ad accesso libero presenti nella rete; il risultato dell'interrogazione è una lista di indirizzi, presso i quali è possibile reperire le informazioni oggetto della ricerca.

### *WAIS (WIDE AREA INFORMATION SERVER)*

WAIS è un sistema per il recupero di informazioni distribuite nella rete;

esso aiuta l'utente nella ricerca sulle basi di dati presenti in Internet. Le informazioni contenute in tali basi di dati sono generalmente di tipo testo, ma possono contenere anche suoni, immagini e filmati, con un vasto campo di applicazioni.

L'organizzazione di tali basi di dati varia a seconda dell'ente che gestisce un particolare WAIS, mantenendo comunque una semplice interfaccia d'utente, al quale vengono mascherate le differenze. Tali interfacce fanno uso generalmente del linguaggio naturale, mediante il quale è possibile impostare tutti i parametri necessari alla ricerca.

L'architettura del servizio è strutturata in modo tale che, a seguito di una ricerca avviata dall'utente, il centro servizi a cui tale utente è collegato si incarica di attivare le opportune operazioni di ricerca sulle basi di dati selezionate, eventualmente remote. Il risultato di tale ricerca si concretizza in un insieme di documenti, che contengono le stringhe oggetto della ricerca con i relativi riferimenti necessari per il recupero degli stessi attraverso la rete.

### *WHOIS*

Il servizio Whois fornisce un sistema di directory elettronico per gli utenti registrati nella rete Internet. Fornisce cioè i mezzi per l'identificazione di indirizzi di posta elettronica, indirizzi postali e numeri telefonici.

La base di dati principale con le informazioni generali sulla rete (organizzazioni, siti, reti, persone, etc. ) è gestito da un ente denominato "INTERNET Network Information Center" (INTER-NIC). Allo stato attuale, i nomi dei gestori dei domini registrati sono automaticamente introdotti nella base di dati quando le autorità di coordinamento della rete Internet concedono una numerazione specifica di rete (indirizzo IP).

L'accesso ai vari server Whois attualmente presenti in rete è reso possibile o attraverso sessioni interattive Telnet con siti opportuni, o facendo uso di appositi programmi applicativi.

In aggiunta a tali modalità di accesso INTER-NIC offre per le proprie basi di dati un'interfaccia di tipo E-mail, per consentire l'accesso anche a quella fascia di utenza che è dotata soltanto di posta elettronica, o che non appartiene alla rete Internet.

### *LISTSERV*

Listserv è un'applicazione per la gestione di liste distribuite che permette a gruppi di utenti con interessi comuni di comunicare tra loro tramite posta elettronica, realizzando al tempo stesso un uso efficiente delle risorse di rete.

La funzione principale di Listserv è quella di operare su liste di indirizzi (mailing list), ossia di realizzare la distribuzione di posta elettronica ad un gruppo di utenti stabilendo un vero e proprio forum di utenti su argomenti di interesse comune. Tale tipo di servizio è utile per lo scambio di idee e di informazioni poiché è la stessa applicazione Listserv che si incarica di recapitare a tutti i partecipanti alla mailing list l'informazione inviata da un suo utente, evitando così a quest'ultimo la replica e l'invio dello stesso messaggio a tutti i componenti della lista.

Listserv fornisce inoltre gli strumenti per effettuare monitoraggio ed archiviazione del traffico di posta elettronica, funzioni di file server e ricerca di archivi in basi di dati.

### *NEWS*

Netnews o Usenet o News è un'applicazione per la condivisione di messaggi che permette di scambiarli elettronicamente usando un formato standard.

I messaggi scambiati su Netnews sono organizzati per argomenti all'interno di categorie chiamate newsgroup. I messaggi possono contenere sia esclusivamente testo che informazioni binarie codificate; possono contenere anche intestazioni informative sul mittente, su quando e dove il messaggio è stato messo in rete, dove è transitato ed altre informazioni amministrative.

Esistono oggi migliaia di categorie newsgroup riguardanti i più svariati argomenti. Il contenuto dei messaggi scambiati nei newsgroup non è in genere regolamentato, anche se qualche newsgroup ha un moderatore che esamina i messaggi prima che siano distribuiti e decide quali siano appropriati per la diffusione.

La componente client di News è chiamata "news reader" ed utilizza il protocollo NNTP (Network News Transfer Protocol) per ricevere informazioni da opportune basi di dati.

## II STRATO FISICO

In questo capitolo sono presentate le caratteristiche e le funzionalità dello strato fisico, e cioè dello strato più basso dell'architettura OSI. Tale strato ha il compito fondamentale di rendere indipendente la comunicazione dal particolare mezzo trasmissivo utilizzato nelle varie tratte che compongono la via fisica tra due sistemi terminali. La trattazione che segue è suddivisa in cinque parti.

Nella prima parte (par. II.1 e II.2) è descritto il modello astratto del servizio di strato fisico e sono considerate le funzioni che questo è chiamato a svolgere nel corso dell'evoluzione di una connessione ad esso pertinente.

Nella seconda parte (par. II.3) viene esaminato l'insieme dei trattamenti che è necessario operare in trasmissione e in ricezione sul segnale di dati per poterne trasferire a distanza il relativo contenuto informativo. Tale insieme di trattamenti corrisponde alle cosiddette *funzioni modem*.

Nella terza parte (par. II.4) sono presentati i protocolli specifici dello strato fisico suddivisi in classi. Queste riguardano, rispettivamente, le caratteristiche meccaniche, elettriche, funzionali e procedurali di una connessione di strato fisico. Particolare attenzione sarà rivolta al caso delle raccomandazioni ITU-T (Racc.) V.24 e X.21, che rappresentano degli esempi diffusi e significativi di protocolli di strato fisico.

Altri esempi di funzionalità tipiche dello strato fisico e riguardanti la N-ISDN e le LAN sono trattati rispettivamente nel par. VIII.2 e nel par. VI.1 e VI.3.

In questo capitolo si pone l'obiettivo: 1) di descrivere le possibili alternative per la realizzazione della connessione di strato fisico tra apparecchiature di elaborazione dell'informazione; 2) di inquadrare tale descrizione nel modello generale di funzionamento dello strato.

A questo proposito occorre precisare che tale associazione è spesso piuttosto difficile e non sempre di univoca soluzione. Ciò è dovuto al fatto che la definizione dei protocolli inerenti le funzionalità di strato fisico è, nella maggioranza dei casi, antecedente alla formalizzazione dell'architettura OSI. Possono quindi essere presenti discordanze tra il modello astratto del servizio di strato fisico e le particolari soluzioni sviluppate nei singoli protocolli.

### II.1 Caratteristiche del servizio

Lo strato fisico (strato 1) è lo strato più basso dell'architettura OSI; esso si interfaccia superiormente con lo strato di collegamento e inferiormente con i

mezzi trasmissivi (Fig. II.1).

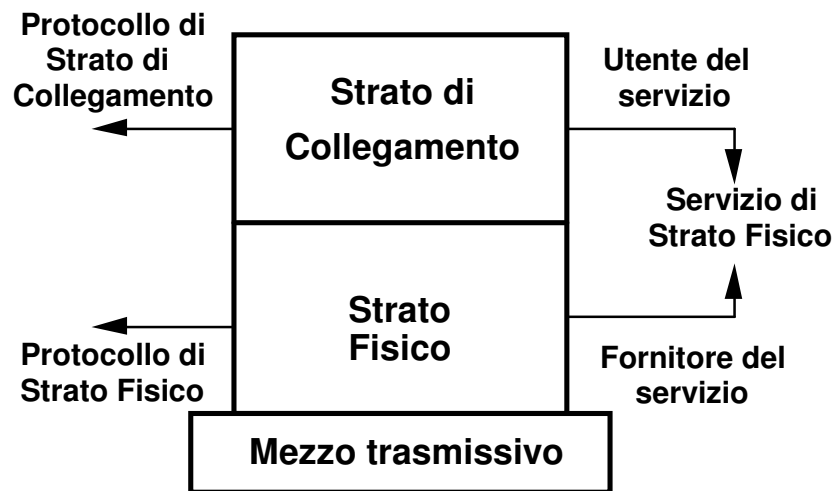


Fig. II.1 - Relazioni tra lo strato fisico e gli strati adiacenti dell'architettura OSI.

Lo scopo principale delle funzioni eseguite all'interno di questo strato e dei servizi offerti allo strato di collegamento è quello di assicurare l'indipendenza della comunicazione dalle caratteristiche fisiche del mezzo trasmissivo utilizzato, fornendo contemporaneamente un trasferimento *trasparente* ed *affidabile* dei dati.

La caratteristica di trasparenza richiede che i dati d'utente siano trasferiti su una connessione di strato fisico senza modifiche e/o limitazioni alla sequenza di cifre binarie che li rappresentano. Per quanto riguarda l'affidabilità di un trasferimento, questa dipende dal grado di fedeltà della sequenza di cifre binarie ricevuta rispetto a quella trasmessa: ciò nonostante effetti vari di natura fisica (disturbi trasmissivi, guasti degli apparati terminali e di linea, ecc.) che possono pregiudicare il conseguimento di questo obiettivo.

Lo strato fisico fornisce i mezzi meccanici elettrici, funzionali e procedurali per instaurare, mantenere e abbattere le connessioni di strato e per trasferire cifre binarie tra due o più entità di collegamento.

Occorre sottolineare che, nel caso dello strato fisico, esiste una stretta relazione tra il concetto di connessione e quello di via fisica in rete. Mentre in tutti gli altri strati dell'architettura OSI il concetto di connessione è esclusivamente di tipo logico, nel caso dello strato fisico questo possiede entrambi i connotati di elemento sia logico che fisico.

Come diretta conseguenza di questa particolarità, nello strato fisico non esiste distinzione tra servizio "orientato alla connessione" e servizio "senza connessione". Il servizio offerto alle entità di collegamento è unico ed è indipendente dal fatto che gli strati superiori operino con l'una o con l'altra delle

due modalità.

Una connessione di strato fisico (*PH-connessione*) coinvolge in generale le entità di strato fisico connesse ai punti terminali tra cui la connessione è instaurata (*entità di origine e di destinazione*); può tuttavia attraversare una o più entità di *rilegamento*. Queste hanno il compito di assicurare la continuità della PH-connessione attraverso la rete, ad esempio eseguendo le funzioni di conversione dei protocolli di strato fisico utilizzati nelle varie tratte della connessione. La Fig. II.2 riporta una schematizzazione di una PH-connessione comprendente una entità di *rilegamento* e evidenzia le funzioni di controllo su di essa operate dallo strato di rete, dato che questo è preposto, nel caso considerato in figura, a svolgere l'instradamento della connessione attraverso la rete.

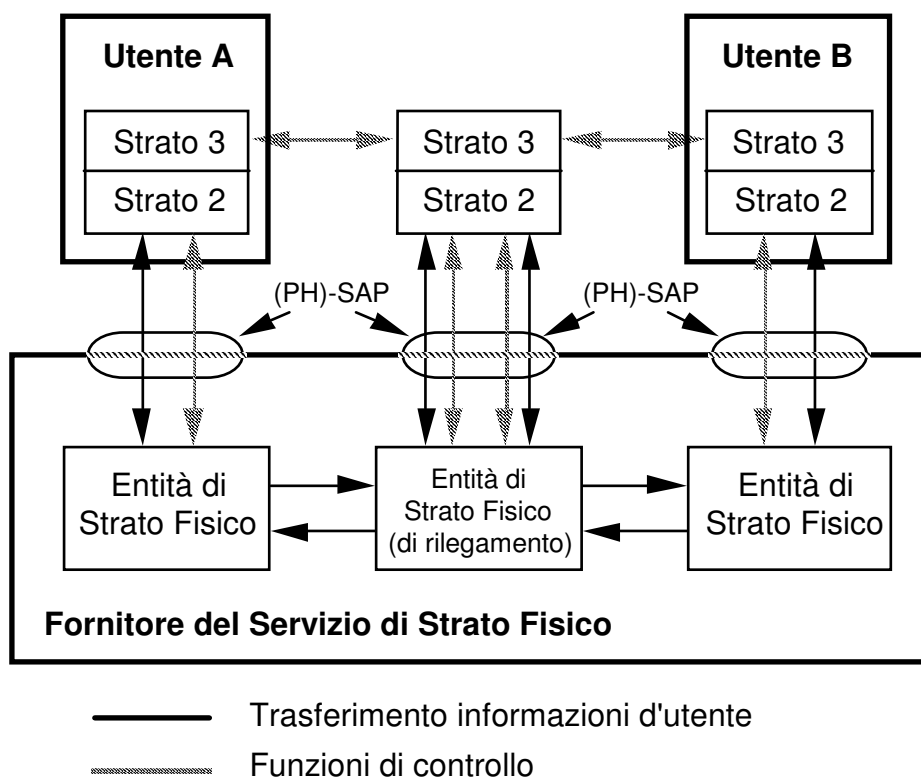


Fig. II.2 - Schematizzazione di una connessione di strato fisico comprendente una entità di *rilegamento* con controllo operato dallo strato di rete.

Ad esempio, nel caso mostrato in Fig. II.3a, la connessione tra due sistemi d'utente A e B avviene mediante due tratte realizzate rispettivamente in coppia coassiale e in fibra ottica. In questo caso il compito delle entità di strato fisico è quello di rendere trasparente agli utenti del servizio (entità di collegamento) l'utilizzazione, nella stessa connessione, di due diversi mezzi trasmissivi. In particolare, ai capi della tratta in coppia coassiale dovranno essere previste

apparecchiature di modulazione e demodulazione (modem), mentre trasduttori elettro-ottici sono necessari agli estremi della tratta in fibra ottica. In questo caso le funzionalità eseguite dai blocchi C e D corrispondono a quelle di una entità di rilegamento, come mostrato nella schematizzazione della connessione mostrata in Fig. II.3b.

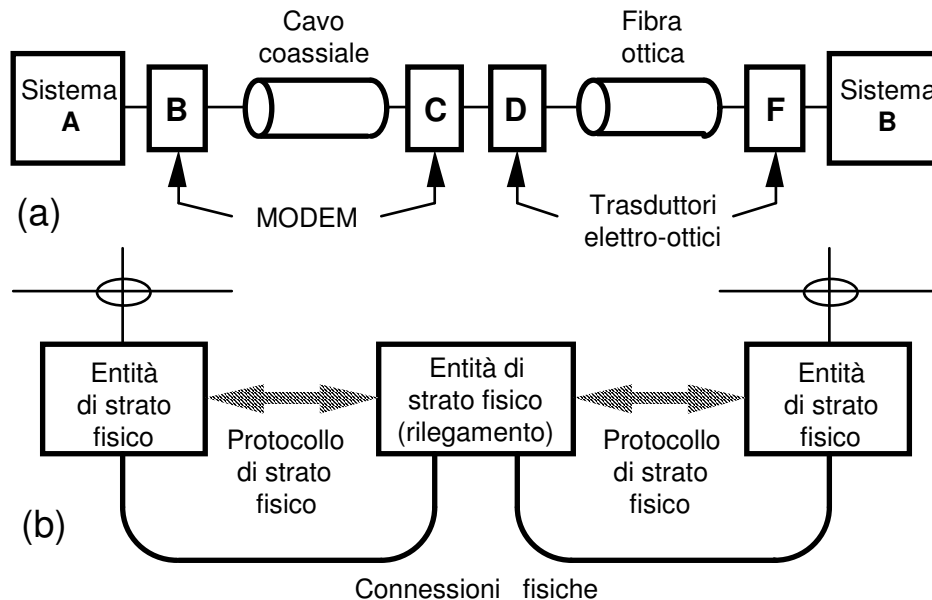


Fig. II.3 - Esempio di connessione fisica:

- (a) *schema fisico;*
- (b) *schema logico.*

Una PH-connessione può essere caratterizzata in base alla *classe di servizio* a cui appartiene e alla *qualità di servizio* che è in grado di offrire. Sia la classe che la qualità di servizio sono determinate dal valore di determinati parametri che sono visibili agli utenti della connessione stessa, e cioè all'interfaccia tra lo strato fisico e quello di collegamento.

In particolare i parametri che concorrono a stabilire la classe di servizio a cui appartiene una specifica PH-connessione sono:

- il *tipo di trasmissione*; questa può essere di tipo *sincrono* o *asincrono* se il segnale di dati emesso dalla entità di origine è, rispettivamente, di tipo sincrono o asincrono;
- la *modalità di funzionamento*; questo può essere di tipo *duplice*, *semi-duplice* o *semplice*; nella modalità duplice la trasmissione può avvenire contemporaneamente nelle due direzioni della connessione (da un estremo della PH-connessione all'altro e viceversa); in quella semi-duplice può essere attiva una sola direzione di trasmissione alla volta, mentre nella modalità semplice la trasmissione è esclusivamente unidirezionale;

- la *topologia*; questa può essere di tipo *punto-punto* o di tipo *punto-multipunto* (Fig. II.4); mentre nella soluzione punto-punto una PH-connezione è utilizzata esclusivamente da solo due entità di origine e di destinazione, nella soluzione punto-multipunto tale connessione è condivisa da una pluralità di entità; ciò comporta un vantaggio in termini di grado di utilizzazione della risorsa trasmissiva, ma pone il problema di coordinare l'accesso a tale risorsa per risolvere eventuali situazioni di contesa.

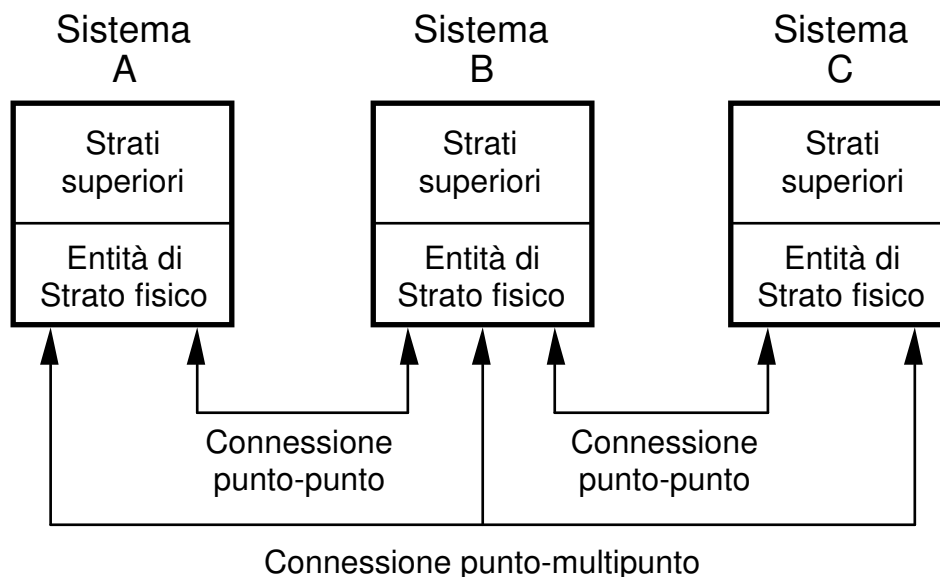


Fig. II.4 -Connessioni fisiche punto-punto e punto-multipunto.

I parametri a cui si fa riferimento nella determinazione della qualità di servizio di una PH-connezione sono:

- 1) la *disponibilità del servizio*, ovvero la quota parte media di tempo in cui la PH-connezione è esente da condizioni di fuori servizio;
- 2) il *tasso d'errore*, definito come la frequenza media secondo cui una singola cifra binaria emessa dall'entità di origine è rivelata non correttamente dall'entità di destinazione;
- 3) la *cadenza di trasferimento*, definita come il valore medio di ritmo binario che la PH-connezione è in grado di garantire nel trasferimento delle relative cifre binarie tra le entità di origine e di destinazione;
- 4) il *ritardo di transito*, definito come l'intervallo di tempo intercorrente tra l'istante di emissione di una cifra binaria dall'entità di origine e quello della sua ricezione da parte dell'entità di destinazione; a tale ritardo contribuiscono il *ritardo di propagazione* sui mezzi trasmissivi e eventuali ritardi introdotti dalle entità di rilegamento.

Ad esempio, nel caso di connessione di strato fisico che utilizzi una tratta

via satellite geostazionario, il ritardo di transito è di circa 280 ms. Infatti, un satellite per telecomunicazioni di questo tipo è posto in un'orbita circolare geostazionaria ad una distanza di 36.000 km dall'equatore e il ritardo di propagazione chilometrico nello spazio libero è di circa 4  $\mu\text{s}/\text{km}$ .

Inoltre, sempre a titolo di esempio, in una rete che utilizza il modo di trasferimento a circuito, una tratta che comprende il transito attraverso un autocommutatore PCM è caratterizzata da un ritardo di transito che è dato dalla somma, oltre che del ritardo di propagazione, anche di tanti contributi quante sono le T-matrici che sono attraversate dalla connessione. Ognuno di questi contributi può avere il valore massimo di 125  $\mu\text{s}$ .

Nella determinazione dei parametri che concorrono a definire la qualità di servizio, occorre tenere conto che tali parametri si riferiscono ad aspetti che sono sotto l'esclusiva responsabilità dello strato fisico. Pertanto il loro valore deve essere misurato in assenza di ulteriori funzioni espletate dall'utente dello strato fisico.

Inoltre, la qualità di servizio di una PH-conessione è una quantità stabilita al momento dell'instaurazione della connessione. Il suo valore può modificarsi, a causa di eventi esterni, durante la sua utilizzazione.

Per chiarire tali affermazioni, si tenga presente, in primo luogo, che per la valutazione del tasso d'errore di una connessione di strato fisico devono essere solo considerate le caratteristiche del mezzo trasmissivo e delle apparecchiature di trasmissione. Le caratteristiche di tasso d'errore, se ritenute non soddisfacenti per una particolare applicazione, saranno migliorate dallo strato di collegamento (strato 2) mediante opportuni controlli di errore (codici di parità, ri-trasmissioni, ecc.). L'applicazione di tali controlli aggiuntivi determina il *tasso d'errore residuo* garantito agli strati superiori.

In secondo luogo, nel caso di PH-conessioni utilizzanti tratte in ponte radio, la qualità di servizio può variare nel tempo poiché dipende dalle condizioni atmosferico-ambientali che sono presenti nella zona circostante la tratta stessa, durante l'intervallo di tempo in cui la PH-conessione è attiva. Ad esempio, per frequenze di portante superiori al GHz, la propagazione delle onde elettromagnetiche è fortemente influenzata dalla presenza di idrometeore (es. pioggia, grandine ecc.). Ciò è dovuto al fatto che la dimensione delle gocce è comparabile con la lunghezza d'onda (dell'ordine del centimetro) del segnale emesso dalle antenne. Tale fenomeno può dar luogo ad attenuazioni molto profonde nel corso della trasmissione e quindi a variazioni della qualità del servizio, in particolare del tasso d'errore, dell'intera connessione.

Infine, occorre ricordare che la velocità di propagazione  $v$  su una linea metallica varia in relazione alle condizioni ambientali. Se la linea è costituita da una coppia simmetrica si ha:

$$v \cong \frac{2}{L \sqrt{\frac{G}{R}} + C \sqrt{\frac{R}{G}}}$$

ove  $R$ ,  $C$ ,  $G$  e  $L$  sono le costanti primarie del mezzo trasmissivo. I valori di tali parametri

variano in funzione delle condizioni ambientali, cioè in funzione della temperatura, dell'umidità, ecc. Corrispondentemente varia anche il ritardo di propagazione sulla PH-connessione che utilizza questo mezzo trasmissivo.

Per facilitare l'interconnessione tra diverse reti pubbliche per dati, l'ITU-T ha normalizzato (Racc. X.1) i possibili ritmi binari secondo cui può avvenire lo scambio di dati tra terminali d'utente nell'ambito dell'impiego di uno specifico modo di trasferimento. In particolare, è stato definito un insieme di *classi di utenza* in relazione al ritmo binario emesso dal terminale sorgente, al tipo di modo di trasferimento consentito dalla rete e all'interfaccia d'accesso utilizzata.

Tali classi di utenza sono descritte in Tab. II.1. Per quanto riguarda la precisazione delle possibili interfacce d'accesso alla rete, queste sono indicate in tabella con la sigla corrispondente alla raccomandazione ITU-T che le riguarda.

Classi di utenza	Modo di trasferimento	Interfaccia di accesso	Ritmo binario (bit/s)
1	A circuito	X.20	300
2		X.20 bis	da 50 a 200
3	A circuito	X.21	600
4		X.21 bis	2400
5			4800
6			9600
7			48000
19			64000
8		A pacchetto	X.25
9	4800		
10	9600		
11	48000		
12	1200		
13	64000		
20	A pacchetto		
21		75/1200	
22		1200	
23		2400	
30	A circ./pacch.	serie I	64000

Tab. II.1 - Classi di utenza nelle reti pubbliche per dati e nell'ISDN in accordo alla Racc. X.1

## II.2 Funzionalità del servizio

Le funzionalità che devono essere assicurate dallo strato fisico alle entità di collegamento che ne facciano richiesta sono:

- a) i *mezzi per attivare una PH-connessione*; connessioni multiple possono esistere tra due entità di collegamento;
- b) i *mezzi per identificare una PH-connessione*; alle entità di collegamento sono forniti i mezzi [*identificatori dei punti terminali di una PH-connessione*] per identificare la particolare PH-connessione utilizzata, tra quelle facenti capo ad uno stesso punto d'accesso al servizio di strato fisico [*PH-SAP*];
- c) i *mezzi per trasferire le unità di dati del servizio di strato fisico (PH-SDU) su una PH-connessione* ; una PH-SDU può essere formata da una singola cifra binaria o da una stringa di cifre binarie; le PH-SDU devono essere trasferite in modo trasparente, senza cioè modifiche o limitazioni al loro contenuto;
- d) i *mezzi per notificare condizioni di malfunzionamento*; alle entità di collegamento sono segnalate eventuali condizioni d'errore interne allo strato fisico;
- e) i *mezzi per disattivare una PH-connessione*; tale disattivazione può essere ordinata dallo strato superiore o avere origine all'interno dello stesso strato fisico in risposta a particolari condizioni di malfunzionamento; la disattivazione forzata di una PH-connessione può comportare la perdita di PH-SDU.

La Fig. II.5 mostra il modello astratto del servizio di strato fisico (*PH-servizio*). Il modello ha lo scopo di definire le interazioni tra le entità dello strato di collegamento (gli *utenti del PH-servizio*) e quelle di strato fisico (i *fornitori del PH-servizio*). Le informazioni necessarie alla gestione di una PH-connessione sono scambiate tra entità appartenenti a strati adiacenti attraverso specifiche *primitive di servizio*. Il funzionamento di una PH-connessione è modellato da due flussi indipendenti di cifre binarie, uno per ogni direzione di trasmissione.

La evoluzione di una PH-connessione può essere suddivisa in tre fasi distinte: la *fase di attivazione*, la *fase di trasferimento dei dati* e la *fase di disattivazione*.

Durante la fase di attivazione della connessione, vengono attivate le direzioni di trasmissione. Nel caso di connessione semi-duplice ogni direzione di trasmissione è attivata in modo indipendente dall'altra, mentre nel caso di connessione duplice entrambe le direzioni sono attivate contemporaneamente.

Durante la fase di trasferimento dei dati vengono trasmesse le PH-SDU,

che sono consegnate a destinazione nello stesso ordine con cui esse sono emesse dall'entità di collegamento di origine.

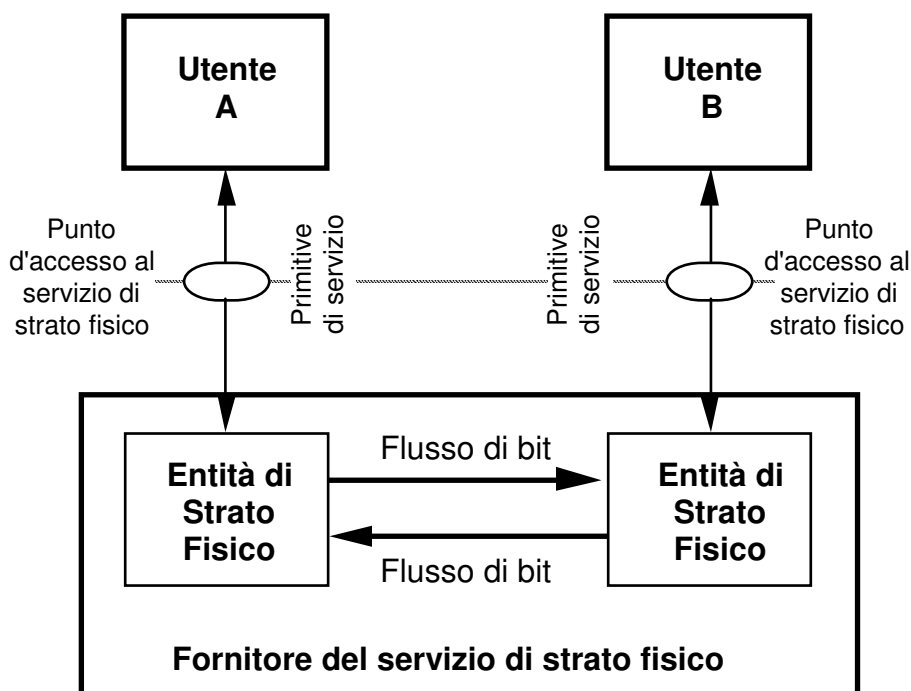


Fig. II.5 - Modello astratto del servizio di strato fisico.

Nella fase di disattivazione della PH-connessione sono disattivate le direzioni di trasmissione. Nel caso di connessione semi-duplice ogni direzione di trasmissione è disattivata in modo indipendente dall'altra, mentre, nel caso di connessione duplice, entrambe le direzioni sono disattivate contemporaneamente.

La Tab. II.2 riporta le primitive di servizio ed i parametri in esse contenuti. Il significato di ogni primitiva è il seguente:

- PH-ACTIVATE: è usata per richiedere o per indicare l'attivazione di una o di entrambe le direzioni di una PH-connessione;
- PH-DATA: richiede o indica il trasferimento delle PH-SDU;
- PH-DEACTIVATE: è usata per richiedere o per indicare la disattivazione di una o di entrambe le direzioni di una PH-connessione; la procedura di disattivazione può essere iniziata sia da un'entità di collegamento, sia dallo stesso strato fisico.

Servizio	Classe di primitive	Tipo di primitive	Tipo di servizio	Parametri
Attivazione di PH connessione	PH-ACTIVATE	Richiesta Indicazione	Non confermato	-
Trasferimento di dati	PH-DATA	Richiesta Indicazione	Non confermato	Dati di utente
Disattivazione di PH connessione	PH-DEACTIVATE	Richiesta Indicazione	Non confermato	-

Tab. II.2 - Primitive di servizio dello strato fisico.

La Fig. II.6 illustra il diagramma di transizione di stato di una PH-connessione facente capo, con uno dei suoi punti terminali, a una interfaccia tra utente e fornitore del PH-servizio. Si distinguono quattro stati, che sono individuati in base a: 1) alla condizione *inattiva* o *attiva* della connessione; 2) alla modalità di funzionamento, che, quando la connessione è attiva, è distinta in duplice da un lato e in semi-duplice o semplice dall'altro; 3) alla direzione attiva di trasmissione quando il modo di funzionamento è semi-duplice o semplice; a tale scopo si utilizzano gli attributi *uscente* e *entrante* con riferimento alla direzione attiva della PH-connessione rispetto al suo punto terminale nell'interfaccia considerata nel diagramma.

Conseguentemente gli stati della connessione sono i seguenti:

- a) *inattiva* (stato 1), quando si è in condizioni di riposo, e cioè senza la possibilità di utilizzare la connessione per un trasferimento di dati;
- b) *attiva duplice* (stato 2), quando è attiva una connessione *bidirezionale* per un funzionamento nel modo duplice;
- c) *attiva uscente, semi-duplice* o *semplice* (stato 3), quando è attiva una connessione *solo uscente* per un funzionamento nel modo semi-duplice o, in alternativa, in quello semplice;
- d) *attiva entrante, semi-duplice* o *semplice* (stato 4), quando è attiva una

connessione *solo entrante* per un funzionamento nel modo semi-duplice o, in alternativa, in quello semplice.

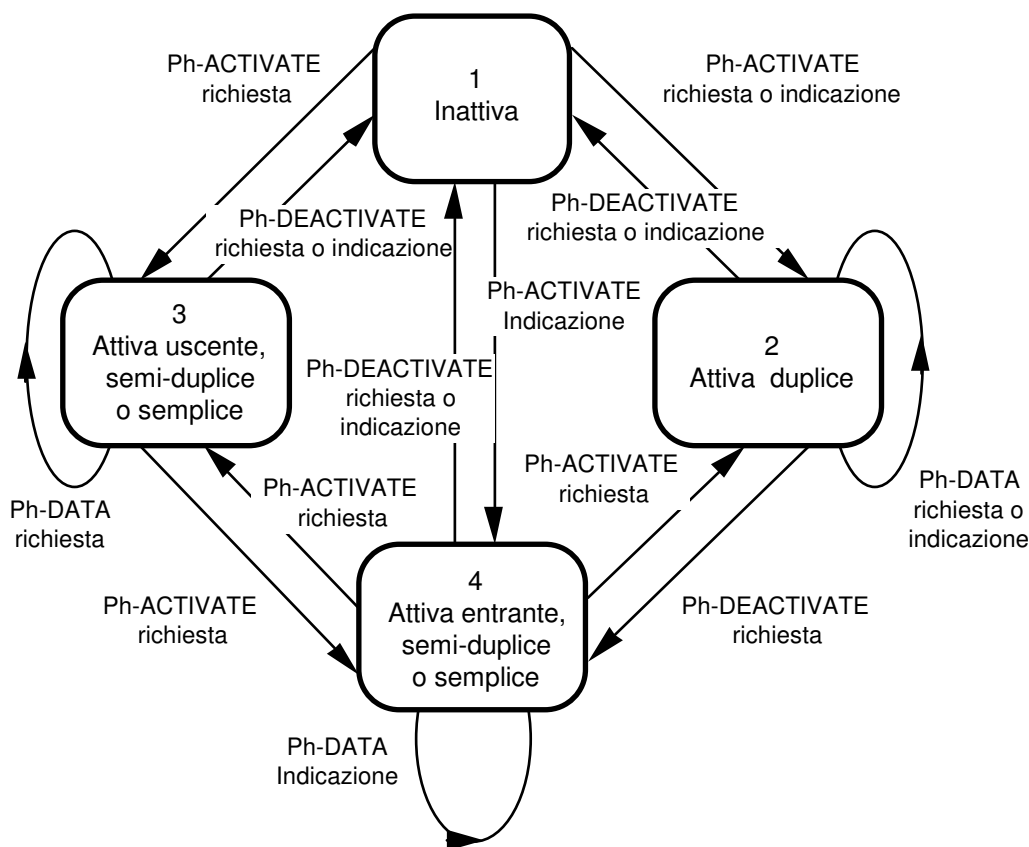


Fig. II.6- Diagramma di transizione di stato per una connessione di strato fisico.

Le transizioni da uno stato all'altro ovvero da uno stato su se stesso avvengono per effetto della emissione o della ricezione, da parte del PH-utente locale, di una delle primitive di Tab. II.2 a seconda dell'elemento di PH-servizio che viene richiesto o indicato dall'utente o dal fornitore. In alcuni casi la transizione avviene solo tramite una primitiva di richiesta ovvero tramite una di indicazione; in altri casi la transizione è consentita a seguito di una primitiva che può essere indifferentemente di richiesta o di indicazione.

Il diagramma di Fig. II.6 descrive le varie possibili evoluzioni dell'interfaccia considerata; è infatti possibile distinguere i tre casi in cui si opera nella modalità duplice, o in quella semplice o in quella semi-duplice; un quarto caso corrisponde ad un funzionamento che è alternativamente duplice e semi-duplice.

#### EVOLUZIONE DI UNA CONNESSIONE DI STRATO FISICO

Nel caso di funzionamento nel modo *duplice*, vengono interessati solo gli stati 1 e 2. La

transizione dallo stato 1 al 2 avviene a seguito di una primitiva PH-ACTIVATE e si manifesta, partendo da una condizione di connessione inattiva a una di connessione attiva bidirezionale. Nello stato 2 è possibile lo scambio bidirezionale simultaneo di dati tramite le primitive PH-DATA. Infine la transizione dallo stato 2 all'1, che conduce ad una disattivazione della PH-connessione bidirezionale, è attuata con la primitiva PH-DEACTIVATE. Si osserva che tutte le tre primitive ora citate possono essere indifferentemente di richiesta o di indicazione: si tratta di richiesta quando l'iniziativa è presa dal PH-utente locale; si ha invece indicazione quando il corrispondente elemento di PH-servizio è invocato da un PH-utente remoto o, direttamente, dal PH-fornitore.

Nel caso di modo *semplice*, l'interfaccia evolve tra gli stati 1 e 3 ovvero tra quelli 1 e 4. Nella prima di queste due alternative, l'attivazione della PH-connessione e il trasferimento dei dati sono per iniziativa del PH-utente locale; invece nella seconda alternativa gli stessi elementi del PH-servizio sono attivati dal PH-utente remoto. In ambedue i casi la disattivazione della PH-connessione avviene indifferentemente per iniziativa del PH-utente locale o di quello remoto ovvero direttamente del PH-fornitore.

Quanto ora detto viene mostrato in Fig. II.6, ove la transizione dallo stato 1 al 3 avviene per effetto di una primitiva di richiesta PH-ACTIVATE, mentre quella dallo stato 1 al 4 si attua tramite una primitiva PH-ACTIVATE, che è del tipo indicazione. Il trasferimento dei dati nello stato 3 o in quello 4 è ottenuto con primitive PH-DATA, che sono rispettivamente dei tipi richiesta o indicazione.

Infine, la disattivazione della PH-connessione si verifica con primitive PH-DEACTIVATE, che possono essere indifferentemente dei tipi richiesta o indicazione in ambedue i casi di partenza dallo stato 3 o dal 4.

Nel caso di modo *semi-duplice*, la PH-connessione assume alternativamente gli stati 3 o 4, passando o meno attraverso lo stato 1.

Si passa dallo stato 1 al 3 o al 4 a seconda che venga attivata la PH-connessione rispettivamente uscente o entrante; le due alternative si attuano con iniziative del tutto analoghe a quelle descritte nel modo semplice. Una volta nello stato 3 o nel 4 il trasferimento dei dati avviene ancora come nel modo semplice. Invece lo scambio tra la condizione di interfaccia emittente (stato 3) a quello di interfaccia ricevente (stato 4) o viceversa si verifica secondo due possibilità.

Nella prima di queste si passa dallo stato di connessione attiva (stato 3 o 4) allo stato 1 e, successivamente, da quest'ultimo al nuovo stato di connessione attiva. Nella seconda si passa direttamente dallo stato 3 al 4 o viceversa esclusivamente per iniziativa del PH-utente locale; in particolare la transizione dallo stato 4 al 3 avviene con l'emissione di una primitiva PH-ACTIVATE di richiesta, mentre la transizione in senso contrario è determinata dalla emissione di una primitiva PH-DEACTIVATE sempre sotto forma di richiesta.

Il diagramma di Fig. II.6 prevede anche la possibilità di una modifica di modo di funzionamento, da semi-duplice a duplice e viceversa. Infatti, dallo stato 4 si passa allo stato 2 o viceversa tramite l'attivazione o la disattivazione della connessione uscente. Le primitive corrispondenti a tali elementi del PH-servizio sono emesse dal PH-utente locale e sono quindi esclusivamente di richiesta.

### II.3 Funzione modem

In questo paragrafo ci riferiamo al modello base di un sistema di

comunicazioni di dati che viene presentato in Fig. II.7, ove la topologia della PH-connessione è di tipo punto-punto.

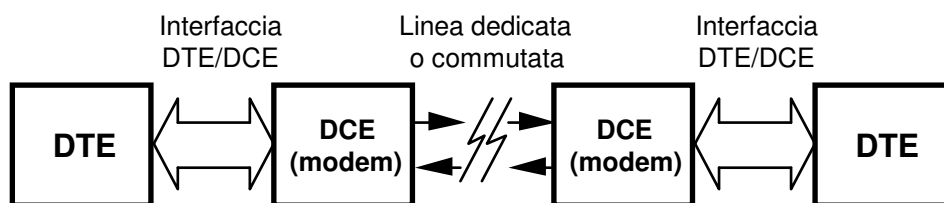


Fig. II.7 - Schema generale di riferimento per una connessione fisica.

Ognuno dei due DTE di estremità in questo modello base può agire come *sorgente* o come *collettore* di dati. Nel primo caso il *DTE sorgente* emette un segnale di dati, che è il supporto di una sequenza di cifre binarie quali risultano dalla codifica dell'informazione insita nei dati da trasferire. Nel secondo caso il *DTE collettore* riceve un segnale di dati, che è il supporto della sequenza di cifre binarie quali risultano a seguito del trasferimento dall'origine alla destinazione e provvede a utilizzare l'informazione corrispondente secondo gli scopi del trasferimento.

Le due apparecchiature DCE, presenti nel modello-base di Fig. II.7, hanno il ruolo di terminazioni della rete, attraverso la quale avviene il trasferimento di informazione.

Pertanto, nell'ambito delle funzioni di strato fisico che debbono fornire i mezzi per trasferire le PH-SDU su una PH-connessione (cfr. par. II.2), ognuno dei due DCE ricopre il ruolo di *terminale trasmittente* o *ricevente* a seconda che sia connesso, rispettivamente, al DTE sorgente o a quello collettore. In ogni DCE si distinguono quindi una sezione trasmittente e una ricevente e l'apparecchiatura che comprende queste due sezioni è chiamata *modem*, secondo un acronimo dei termini *Modulator* e *Demodulator*.

In questo ruolo il compito dei DCE è quello di svolgere la funzione di trattamento terminale del segnale di dati, che, nel seguito, chiameremo *funzione modem*. Deve tuttavia essere sottolineato che, sempre limitatamente alle funzioni di strato fisico, un DCE è anche coinvolto nello svolgimento degli ulteriori compiti richiesti a questo strato (cfr. § I.2.1).

Con riferimento specifico alla funzione modem, l'insieme comprendente, nell'ordine, il DCE trasmittente, il mezzo trasmissivo e il DCE ricevitore costituisce il canale trasmissivo dalla sorgente al collettore dei dati. Questo canale, che è talvolta chiamato anche *circuito di dati*, riceve in ingresso il segnale di dati emesso dalla sorgente e deve restituire in uscita, al collettore, un segnale di dati che sia la replica più fedele possibile di quello di sorgente.

Pertanto il DCE trasmettitore ha il compito di formare il segnale di linea in trasmissione a partire dal segnale di dati emesso dal DTE sorgente, mentre il DCE ricevitore ha il compito di estrarre il segnale di dati ricevuto dal segnale di linea in ricezione.

La formazione del segnale di linea in trasmissione deve essere effettuata in modo che le caratteristiche spettrali di questo segnale siano opportunamente armonizzate con la risposta in frequenza del canale trasmissivo. La estrazione del segnale di dati dal segnale di linea in ricezione deve rispondere al requisito di un contenuto tasso di errore, nonostante i disturbi di natura varia che tendono a mascherarne il relativo contenuto informativo.

In generale il DCE trasmettitore forma i segnali di linea in trasmissione effettuando, sul segnale di dati emesso dal DTE sorgente, una elaborazione che, almeno concettualmente, può essere decomposta nella sequenza di due trasformazioni, e cioè, nell'ordine, di una *codifica di linea* e di una *modulazione*. Lo schema a blocchi del DCE trasmettitore è quindi quello mostrato in Fig. II.8a.

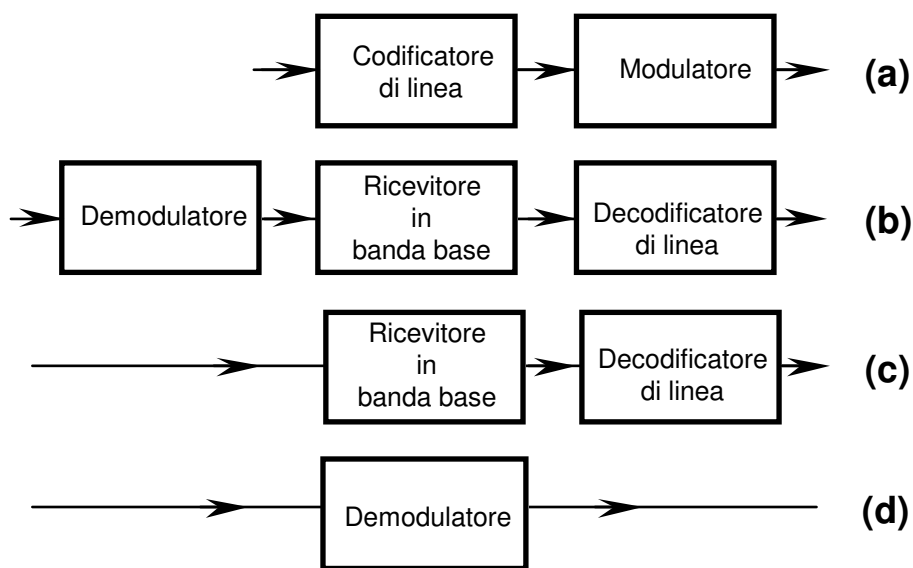


Fig. II.8 - Schemi funzionali di:

- (a) un DCE trasmettitore;
- e di un DCE ricevitore, quest'ultimo nei tre casi di
- (b) trasmissione sincrona in banda traslata;
- (c) trasmissione sincrona in banda base;
- (d) trasmissione asincrona.

Lo scopo e le modalità di attuazione di queste trasformazioni verranno chiarite nel seguito. Occorre però precisare subito che il tipo di modulazione varia a seconda che si operi con una trasmissione in banda base o con una in

banda traslata: in particolare nel secondo caso si impiega usualmente una modulazione armonica tra quelle descritte nel seguito del paragrafo, mentre nel primo caso la modulazione si riduce a una operazione di *formazione di impulso*.

Per ciò che riguarda poi il DCE ricevitore, i trattamenti effettuati sul segnale di linea in ricezione variano, oltre che nei due tipi di trasmissione ora considerati, anche a seconda che la trasmissione sia sincrona o asincrona.

Nel caso in cui la trasmissione sia sincrona e in banda traslata, il DCE ricevitore comprende, nell'ordine, un demodulatore e un ricevitore in banda base. Il risultato di queste due trasformazioni viene poi trattato da un decodificatore, che effettua l'operazione inversa a quella attuata dal codificatore di linea in trasmissione. Questo complesso di operazioni è descritto nello schema funzionale di Fig. II.8b.

Sempre nel caso di trasmissione sincrona, ma in banda base, non è necessaria l'operazione di demodulazione e quindi lo schema funzionale del DCE ricevitore si semplifica come mostrato in Fig. II.8c.

Infine, nel caso di trasmissione asincrona, che è considerato in Fig. II.8d, tutto il trattamento del segnale di banda base (cioè a valle dell'operazione di demodulazione) è effettuato, per ragioni sostanzialmente storiche, dal DTE collettore. Quindi il DCE ricevitore agisce unicamente come demodulatore nel caso di trasmissione in banda traslata. Non ha invece interesse pratico il caso di trasmissione in banda base, dato che questa modalità non ha applicazione conveniente quando la trasmissione è asincrona.

Prima di entrare in maggior dettaglio su quanto si è finora detto sulla funzione modem, è interessante ricordare quale sia il massimo ritmo binario  $C$ , che può essere trasferito in modo affidabile su una PH-connezione. Secondo quanto dimostrato da Shannon in un celebre teorema della teoria dell'informazione, tale massimo (espresso in bit/s) è dato da

$$\text{Eq. II.1 } C = B \lg_2 \left( 1 + \frac{S}{N} \right)$$

ed è direttamente proporzionale alla larghezza di banda  $B$  (espressa in Hz) del canale trasmissivo e dipendente, con legge logaritmica, dal rapporto segnale/rumore  $S/N$  in ricezione.

Per illustrare con un esempio le implicazioni della (Eq. II.1), assumiamo un canale telefonico come supporto di una PH-connezione per dati e teniamo conto che la larghezza di banda utile  $B$  di questo canale è di 3,1 kHz. Se supponiamo, come è ragionevole, che il rapporto segnale/rumore in ricezione

sia uguale a 30 dB, dalla (Eq. II.1) si ottiene  $C =$  circa 30 kbit/s. Questo rappresenta un limite superiore teorico per la capacità trasmissiva di un canale telefonico.

Dopo questo quadro di insieme, esaminiamo ora più in dettaglio le funzioni modem nei due casi di trasmissione in banda base (§ II.3.1) e in banda traslata (§. II.3.2). Il paragrafo verrà concluso (§.II.3.3) dalla trattazione delle modalità secondo cui i DCE trasmettitore e ricevitore possono operare in condizioni di sincronismo.

### II.3.1 *Trasmissione in banda base*

Una trasmissione in banda base è senz'altro adottabile quando il canale trasmissivo ha risposta in frequenza di *tipo passa-basso*. Tale tecnica di trasmissione può però essere impiegata anche con quei tipi di canali trasmissivi con comportamento sostanzialmente passa-basso, ma in cui la banda passante non comprende le frequenze in un intorno sufficientemente ristretto della frequenza zero. Questa limitazione può essere dovuta, ad esempio, alla presenza, lungo il canale trasmissivo, di trasformatori di adattamento o di isolamento.

La limitazione della banda passante alle basse frequenze è uno dei motivi che rendono impossibile trasmettere direttamente il segnale di dati emesso dal DTE sorgente e quindi indispensabile l'operazione di codifica di linea effettuata dal DCE trasmettitore.

Il vantaggio della trasmissione in banda base risiede nella sua semplicità. I DCE trasmettitore e ricevitore (chiamati, nel loro complesso, *modem in banda base*) sono apparecchiature circuitalmente più semplici e quindi economicamente vantaggiose rispetto agli apparati corrispondenti per trasmissione in banda traslata: tale vantaggio è tanto maggiore quanto più elevato è il ritmo binario emesso dalla sorgente.

La trasmissione di dati in banda base è impiegata, ad esempio, in connessioni fisiche supportate da linee metalliche, se queste non presentano limitazioni di banda passante oltre quelle connesse alla legge di incremento dell'attenuazione con la frequenza. Queste linee possono essere rese disponibili dalla rete telefonica in area urbana e possono avere una lunghezza fino a una ventina di chilometri.

Un'ulteriore importante applicazione della trasmissione in banda base si ha con l'impiego di sistemi di trasmissione numerica, che rendono attualmente disponibili canali trasmissivi con capacità di 64 kbit/s. Questa capacità può

essere utilizzata per trasferire direttamente ritmi binari di 64 kbit/s (cfr. la classe d'utenza n° 30 di Tab. II.1) o anche ritmi binari inferiori (ad esempio da 600 bit/s fino a 9600 bit/s secondo le classi di utenza n° 3, 4, 5 e 6 di Tab. II.1). In questo secondo caso occorre fare ricorso a multiplazioni statiche a divisione di tempo per utilizzare la capacità disponibile al servizio di una molteplicità di sorgenti di dati.

Le tecniche di trattamento terminale del segnale di dati impiegate nella trasmissione in banda base hanno importanza fondamentale anche nella trasmissione in banda traslata: infatti, come appare dagli schemi funzionali nelle Fig. II.8a e Fig. II.8b, esse trovano ivi impiego sia nella sezione trasmittente del modem (in particolare, nella formazione del segnale di dati a monte dell'operazione di modulazione), sia in quella ricevente (in particolare, nel trattamento di tale segnale a valle dell'operazione di demodulazione).

Inoltre, sempre nella trasmissione in banda traslata, la sezione di canale trasmissivo, che è compresa tra l'ingresso del modulatore e l'uscita del demodulatore, è di tipo passa-basso. Le proprietà ad essa richieste per contenere il tasso d'errore della connessione sono quindi del tutto analoghe a quelle che verranno qui introdotte nella trattazione della trasmissione su canali di questo tipo.

Nel seguito forniremo alcuni elementi sulle modalità di lavoro delle sezioni trasmittente e ricevente di un modem in banda base.

La sezione trasmittente di un modem in banda base comprende, in accordo allo schema a blocchi di Fig. II.8a, un codificatore di linea e un modulatore.

Il codificatore di linea trasforma il segnale di dati emesso dal DTE sorgente in un segnale numerico di banda base. A tale scopo esso opera sulla sequenza  $[b_i]$  di cifre binarie emesse dal DTE sorgente e la trasforma in una sequenza  $[a_j]$ , i cui elementi sono scelti, in generale, in un alfabeto di  $q$  cifre (o simboli), con  $q \geq 2$  (*cifre q-arie*). La sequenza  $[a_j]$ , chiamata *sequenza di linea*, è poi utilizzata per formare un segnale di banda base a  $q$  livelli.

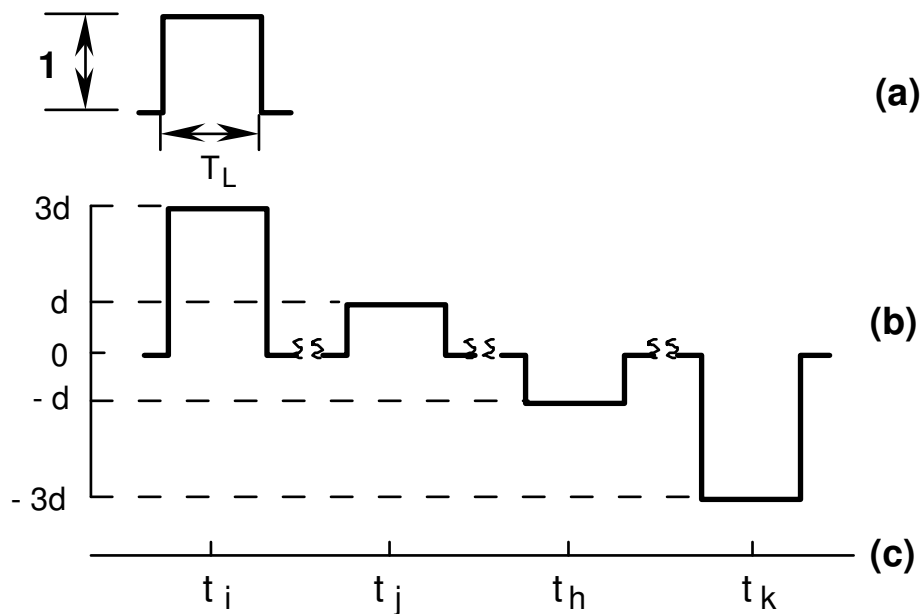


Fig. II.9 - Illustrazione del principio di generazione di un segnale di banda base a quattro livelli, generato utilizzando l'impulso elementare rappresentato in (a), le ampiezze di impulso descritte in (b) e gli istanti significativi riportati in (c).

Tale segnale è composto da una *sequenza di impulsi*, tutti di *forma* uguale a quella di un *impulso elementare* e collocati in coincidenza con gli istanti significativi del segnale. In uno di questi istanti l'ampiezza dell'impulso ivi collocato assume un livello (ad esempio di tensione), che è in corrispondenza biunivoca con l'elemento della sequenza di linea generato nello stesso istante. Quanto ora detto è esemplificato graficamente in Fig. II.9, con riferimento al caso  $q = 4$ .

In Fig. II.10 è mostrato un esempio di codifica di linea con formazione di un segnale di banda base a quattro livelli. L'impulso elementare è di forma rettangolare e le ampiezze degli impulsi possono assumere quattro livelli possibili ( $+3d$ ,  $+d$ ,  $-d$ ,  $-3d$ ) tra loro a distanza  $2d$ .

In particolare la Fig. II.10a descrive il segnale di dati entrante nel codificatore, mentre la Fig. II.10b riporta l'andamento del segnale numerico a quattro livelli.

Lo schema di codifica è descritto dalla seguente legge di corrispondenza tra le sequenze  $[b_i]$  e  $[a_j]$

00	-3
01	-1
10	+1
11	+3

dove, nelle prime due colonne, sono riportate le coppie di cifre binarie  $b_i b_{i+1}$ , che sono poste in corrispondenza con la cifra quaternaria  $a_j$ , indicata nella terza colonna.

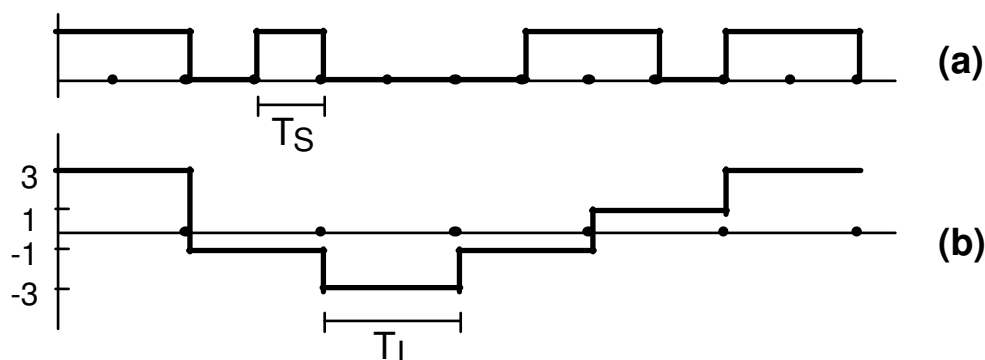


Fig. II.10 - Esempio di codifica a quattro livelli.

Passando poi al modulatore, questo ha lo scopo di formare i segnali di linea in trasmissione a partire dal segnale di banda base uscente dal codificatore di linea. In particolare deve provvedere a modificare, se necessario, la forma dell'impulso elementare e ad attribuire al segnale di linea in trasmissione la potenza richiesta per ottenere un desiderato valore del rapporto segnale/rumore in ricezione.

Le proprietà di un codice di linea per trasmissione in banda base sono svariate. Qui ci limitiamo a considerarne tre tra le più significative.

In primo luogo ha interesse considerare come viene modificato, tra l'ingresso e l'uscita del codificatore, l'intervallo tra due istanti significativi successivi (*intervallo di cifra*). Se, per semplicità come faremo anche nel seguito, ci riferiamo al caso di trasmissione sincrona, questo intervallo è costante nel tempo; indichiamo pertanto con  $T_S$  e  $T_L$  gli intervalli di cifra per il segnale di dati entrante nel codificatore e per il segnale di banda base da esso uscente, rispettivamente (cfr. Fig. II.10).

Il legame tra  $T_S$  e  $T_L$ , che dipende dal tipo di codice adottato, determina, in relazione al *ritmo binario di sorgente*  $F_S = 1/T_S$  (espresso in bit/s), quale sia il ritmo di cifra  $F_L$  uscente dal codificatore. Questa seconda grandezza esprime la frequenza con cui si susseguono gli istanti significativi del segnale di banda

base ed è pertanto definita da  $F_L = 1/T_L$ . Essa è chiamata *ritmo di cifra in linea* ed è misurata in *baud*, quando l'unità di tempo di riferimento è il secondo. Osserviamo che, in relazione al tipo di codifica adottato,  $F_L$  può risultare numericamente uguale a  $F_S$ , ovvero a un suo multiplo, ovvero a un suo sottomultiplo.

In particolare, quando la codifica è a più di due livelli ( $q > 2$ ) e una sola cifra  $q$ -aria viene associata a un gruppo di  $\log_2 q$  cifre binarie, il legame tra  $F_S$  e  $F_L$  è dato da:

$$\text{Eq. II.2 } F_L = F_S / \log_2 q$$

A tale proposito osserviamo che, in generale, la corrispondenza tra le sequenze  $[b_i]$  e  $[a_j]$  è del tipo seguente: gruppi di  $M$  cifre  $q$ -arie ( $M \geq 1$ ) vengono associate a gruppi di  $N$  cifre binarie ( $N \geq 1$ ), con la condizione che i numeri  $q$ ,  $M$  e  $N$  soddisfino la relazione  $q^M \geq 2^N$ .

D'altra parte, per la continuità dei flussi di cifre entranti e uscenti nel codificatore, la durata  $N/F_S$  di ciascun gruppo di cifre binarie deve essere uguale a quella  $M/F_L$  dell'associato gruppo di cifre  $q$ -arie. Conseguentemente il rapporto  $F_S/F_L$  risulta uguale, in generale, al rapporto  $N/M$ .

Si può quindi concludere che:

- 1) se la codifica è a due livelli ( $q = 2$ ,  $M \geq N$ ), si ha sempre  $F_S/F_L \leq 1$ , ove il segno di uguaglianza vale solo nel caso in cui la trasformazione sia un'identità;
- 2) solo aumentando il numero dei livelli ( $q > 2$ ), si può rendere  $M \leq N$  e quindi  $F_S/F_L \geq 1$ , ove il segno di uguaglianza vale nei casi di codifica senza variazione di dimensionalità dei gruppi ( $M = N$ ).

Ad esempio, nel caso della codifica di linea considerata in Fig. II.10, risulta  $q = 4$ ,  $M = 1$ ,  $N = 2$  e quindi  $F_L = F_S/2$ . Conseguentemente, se  $F_S = 9600$  bit/s, si ha  $F_L = 4800$  baud.

Una seconda proprietà di un codice di linea è data dall'andamento dello *spettro di densità di potenza* del segnale numerico che esso genera. Infatti l'intervallo di frequenza in cui tale spettro presenta valori significativi definisce la *larghezza di banda* del segnale di linea e questa larghezza deve essere contenuta per ridurre l'effetto delle distorsioni lineari, dovute al taglio in alta frequenza del canale trasmissivo. Inoltre, per le ragioni chiarite in precedenza, ha interesse il valore che lo spettro assume in corrispondenza della frequenza zero e il suo andamento nell'intorno di questa frequenza.

Come terza proprietà importante occorre infine considerare il numero

massimo di intervalli di cifra in cui il livello di uscita si può mantenere costante in relazione a particolari realizzazioni del segnale di dati (ad esempio in corrispondenza di lunghe sequenze di zeri o di uni consecutivi). Infatti questo numero massimo deve essere sufficientemente piccolo in modo da facilitare il funzionamento dell'estrattore del segnale di temporizzazione (cfr. § II.3.3). I codici di linea per trasmissione di dati in banda base sono svariati. Qui ci limiteremo a descriverne alcuni tra i più significativi a due o a tre livelli. In Fig. II.11 sono mostrati gli andamenti nel tempo di alcuni segnali di banda base ottenuti con codifica di linea a due livelli.

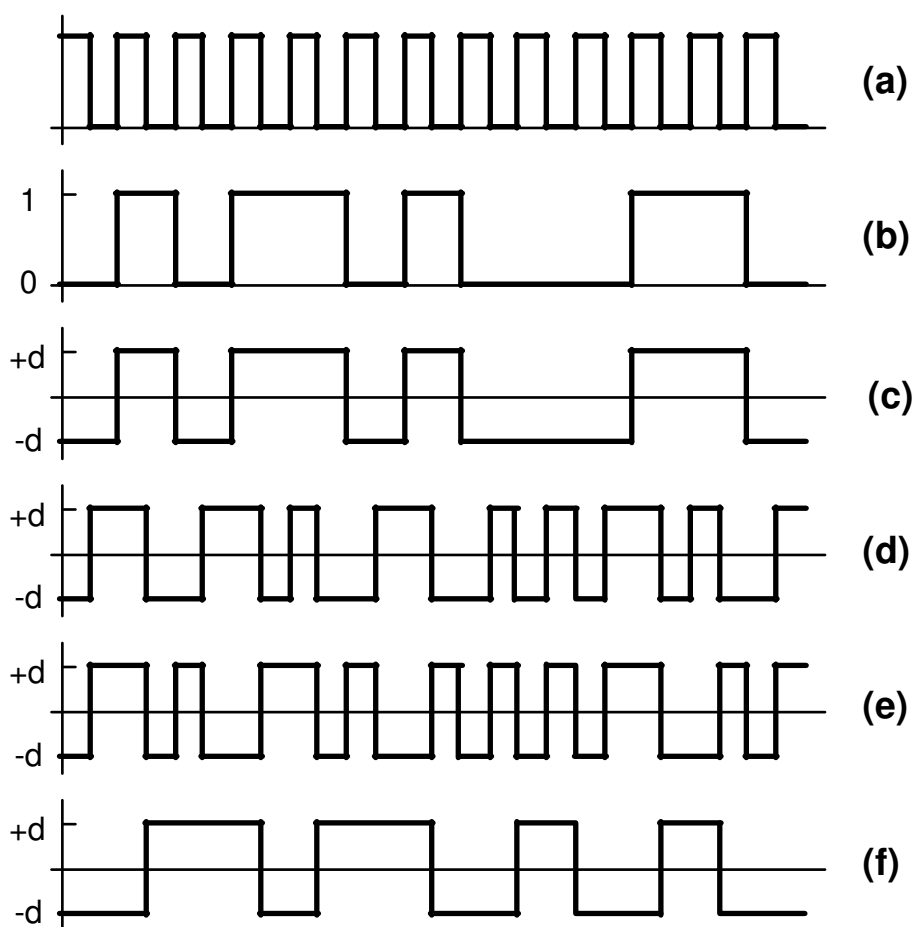


Fig. II.11 - Esempi di segnale di banda base a due livelli:

- (a) *segnale di temporizzazione;*
- (b) *segnale di dati;*
- (c) *segnale di banda base con codice NRZ;*
- (d) *segnale di banda base con codice bifase;*
- (e) *segnale di banda base con codice bifase differenziale;*
- (f) *segnale di banda base con codice di Miller.*

L'impulso elementare è rettangolare e la distanza dei due livelli è uguale a  $2d$ . In particolare in Fig. II.11a è riportato il segnale di temporizzazione che

definisce gli istanti significativi del segnale di dati rappresentato in Fig. II.11b. Gli spettri di densità di potenza dei segnali di banda base considerati in Fig. II.11 sono rappresentati in Fig. II.12.

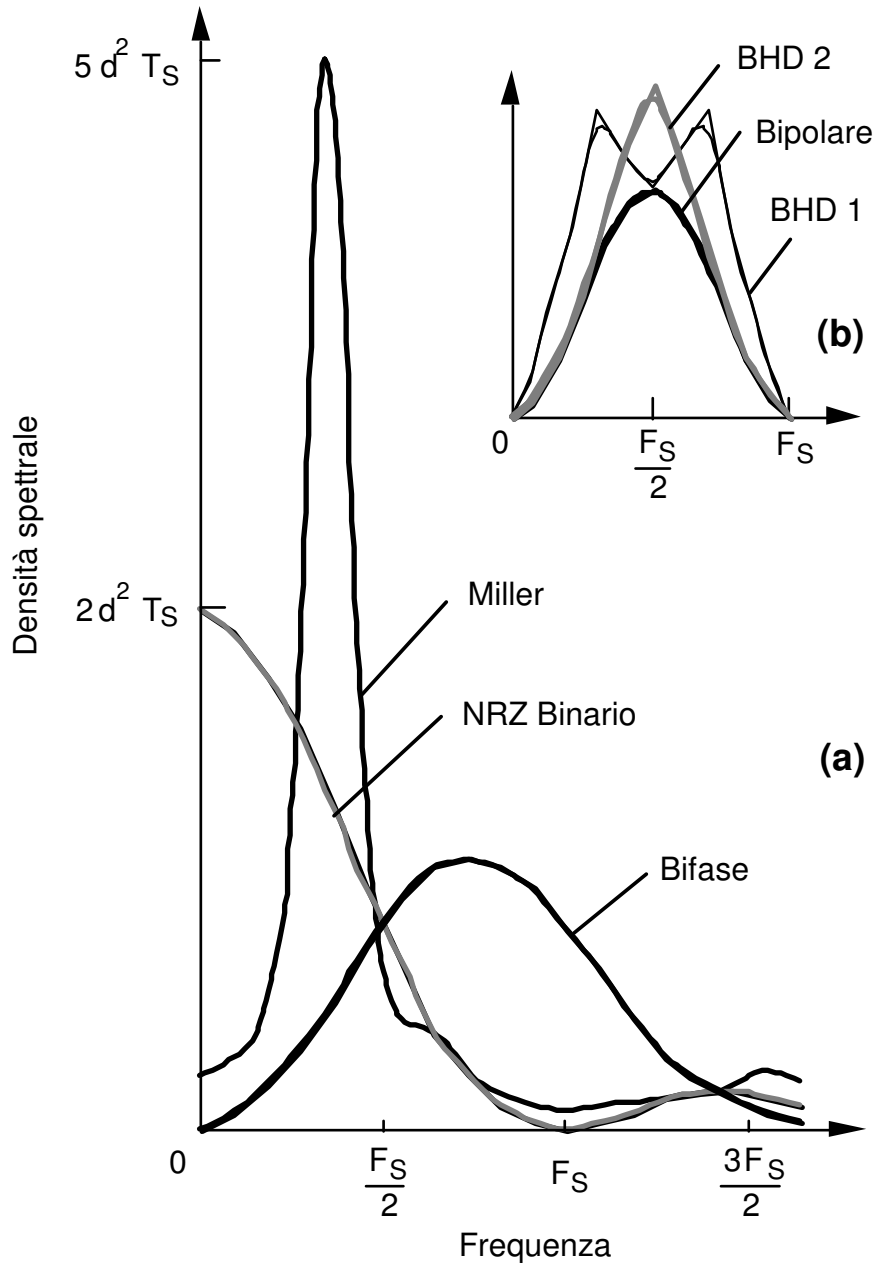


Fig. II.12 - Spettri di densità di potenza di segnali di banda base a due livelli (a) o a tre livelli (b), ottenuti con varie codifiche di linea.

### ESEMPI DI CODICI DI LINEA A DUE E A TRE LIVELLI

Il codice a due livelli più semplice da realizzare è quello denominato NRZ (No Return to Zero), che è così definito: il livello di uscita del codificatore, come illustrato in Fig. II.11c, è uguale a  $-d$  o a  $+d$  a seconda che la cifra binaria d'ingresso sia uguale a 0 o a 1,

rispettivamente. Tuttavia le caratteristiche spettrali del segnale generato da un codificatore NRZ (cfr. Fig. II.12a) mal si prestano a una trasmissione in banda base, dato che lo spettro di densità di potenza è concentrato in vicinanza della frequenza zero. Ulteriore inconveniente di questo codice è dato dal fatto che, quando l'ingresso del codificatore presenta lunghe sequenze di zeri, l'uscita rimane a un livello costante. Osserviamo infine che, con l'uso del codice NRZ, poiché risulta  $T_S = T_L$ , il ritmo di cifra in linea è numericamente uguale al ritmo binario del segnale di dati.

Un codice di linea a 2 livelli, che evita gli inconvenienti del codice NRZ e che quindi è più frequentemente impiegato, è quello *bifase*, noto anche come *codice di Manchester*. In questo caso ciascuna cifra binaria di ingresso  $b_i$  è trasformata nella sequenza di due cifre binarie  $a_i, a_i''$  con la seguente corrispondenza: se  $b_i = 1$ , allora  $a_i = +d, a_i'' = -d$ ; se invece  $b_i = 0$ , allora  $a_i = -d, a_i'' = +d$ . Un esempio di applicazione è dato in Fig. II.11d.

Il codice bifase esige una corretta scelta della polarità dei fili di linea, in quanto un loro scambio provoca il cambiamento di segno del segnale emesso. Per evitare questo inconveniente si preferisce normalmente utilizzare il codice *bifase differenziale*, che è così definito (cfr. Fig. II.11e): se  $b_i = 1$ , allora  $a_i = -a_{i-1}, a_i'' = -a_i$ ; se invece  $b_i = 0$ , allora  $a_i = a_{i-1}, a_i'' = -a_i$ .

Da un punto di vista spettrale i due codici bifase hanno comportamento analogo. Essi presentano infatti (cfr. Fig. II.12a) un contenuto nullo a frequenza zero e occupano una larghezza di banda che è doppia rispetto a quella del codice NRZ: ciò si giustifica tenendo conto che in questo caso risulta  $T_S = 2 T_L$ . Pertanto il ritmo di cifra in linea è numericamente doppio del ritmo binario del segnale di dati.

Un altro codice a due livelli è il *codice di Miller*, che può essere ottenuto da quello bifase sopprimendo in questo una transizione su due, secondo quanto illustrato graficamente in Fig. II.11f. In questo caso lo spettro di densità di potenza (cfr. Fig. II.11a) ha larghezza di banda confrontabile con quella del codice NRZ e quindi decisamente inferiore a quella dei codici bifase; tuttavia, a differenza di questi ultimi, non si ha annullamento dello spettro a frequenza zero.

I codici a tre livelli hanno lo stesso vantaggio del codice di Miller (cioè una ridotta larghezza di banda spettrale), ma, rispetto a questo, hanno l'ulteriore vantaggio di annullare lo spettro a frequenza zero (cfr. Fig. II.12b).

Nel codice *bipolare semplice*, noto anche come *codice AMI* (Alternate Mark Inversion), il segnale emesso ha livello uguale a 0 quando la cifra binaria d'ingresso è 0 ed è uguale alternativamente a  $+d$  o a  $-d$  quando  $b_i = 1$ . Questo codice ha tuttavia lo svantaggio che il segnale emesso rimane a livello zero quando il segnale di dati contiene una sequenza di zeri consecutivi.

Una variante del codice bipolare semplice, che non presenta questo inconveniente, è data dai codici *bipolari ad alta densità* (HDB-High Density Bipolar). Questi codici non si comportano come il codice bipolare semplice solo quando il segnale di dati comprende una sequenza di zeri consecutivi. Nel caso specifico del codice HDBn sono permessi al più  $n$  intervalli di cifra con livello zero. In questa famiglia di codici, che ha comportamento spettrale del tutto analogo a quello del codice bipolare semplice, uno dei più comunemente usati è il *codice HDB3*.

Tra i codici di linea ora descritti solo il codice NRZ non introduce ridondanza; tutti gli altri ne introducono o incrementando il ritmo delle transizioni (come nei codici bifase) o aumentando il numero dei livelli (come avviene nei codici bipolari). Tutte queste tecniche di codifica introducono però, rispetto al codice NRZ, una diminuzione della resistenza ai disturbi del ricevitore. Infatti:

- i codici bifase presentano un raddoppiamento della larghezza di banda spettrale e, conseguentemente, anche della potenza di disturbo ammessa nel ricevitore;
- i codici bipolari hanno la stessa larghezza spettrale, ma dato l'impiego di tre livelli, la loro resistenza ai disturbi è di due volte inferiore;
- il codice di Miller non introduce aumenti della larghezza di banda spettrale e del numero di livelli, ma il mancato annullamento della componente continua si traduce in una diminuzione della resistenza ai disturbi.

Si può quindi concludere che il tipo di codifica di linea deve essere scelto sulla base della risposta in frequenza del mezzo trasmissivo impiegato e del desiderato ritmo binario di sorgente. Con questi elementi è possibile, di caso in caso, ottenere un corretto compromesso tra i vantaggi e gli svantaggi presentati da ogni codice.

Circa quanto ora detto nel confronto tra i codici di linea precedentemente illustrati, è opportuno fornire qualche ulteriore elemento sui due concetti di *ridondanza di codifica* e di *resistenza ai disturbi del ricevitore*. Del secondo parleremo poi nel seguito. Chiariamo invece subito il concetto di ridondanza di codifica.

Chiamiamo *potenzialità informativa* di un segnale numerico il rapporto tra la sua massima quantità di informazione media e la durata dell'intervallo di cifra.

In base a questa definizione, la potenzialità informativa del segnale uscente da un codificatore di linea a  $q$  livelli è data da  $F_L \log_2 q$ . Questa potenzialità è utilizzata in parte per trasferire l'informazione emessa dalla sorgente binaria e in parte per imporre al segnale di linea quei vincoli che gli conferiscono le proprietà desiderate.

La prima parte è equivalente alla potenzialità informativa del segnale di dati emesso dalla sorgente ed è data da  $F_S$ . La seconda parte, uguale alla differenza  $F_L \log_2 q - F_S$ , viene detta *ridondanza assoluta* del codice e costituisce un indice della quantità di vincoli introdotti, nell'unità di tempo, entro il segnale di linea.

Nella sezione ricevente del DCE, un ricevitore in banda base (cfr. Fig. II.8c) comprende un *amplificatore-equalizzatore* e un *decisore*, oltre a un *estrattore di segnale di temporizzazione*, secondo lo schema a blocchi rappresentato in Fig. II.13.

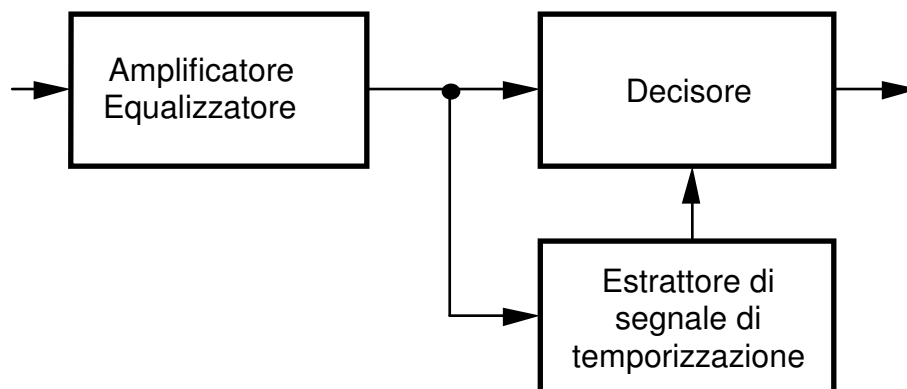


Fig. II.13 - Schema funzionale di un ricevitore in banda base.

Cerchiamo di chiarire i compiti dell'amplificatore-equalizzatore e del decisore. A tale scopo osserviamo che gli impulsi costituenti il segnale di linea in ricezione si presentano: a) attenuati per effetto della *linea*, e cioè del tronco di mezzo trasmissivo interposto tra trasmettitore e ricevitore; b) degradati nella loro forma a causa delle distorsioni lineari, introdotte dalla sezione di canale trasmissivo che precede il ricevitore. Queste degradazioni si manifestano con una deformazione di ogni impulso rispetto a quello in uscita dal trasmettitore e con conseguenti code che invadono gli intervalli di tempo degli impulsi adiacenti.

Compito dell'amplificatore-equalizzatore è allora quello di compensare l'attenuazione di linea e di eliminare (o almeno ridurre) gli effetti di disturbo prodotti dalle deformazioni suddette. Si tratta però di comprendere come queste azioni possano essere svolte e perché lo debbano essere.

A questo scopo spieghiamo: 1) come viene elaborato il segnale a valle dell'amplificatore-equalizzatore per estrarne la sequenza di linea; 2) quali siano le cause di degradazione che possono compromettere il risultato di questa elaborazione; 3) quali siano i parametri del ricevitore che qualificano la fedeltà di ricezione e la sua resistenza ai disturbi.

Circa il punto 1), si tratta dei compiti svolti dal decisore. In questo dispositivo il segnale uscente dall'amplificatore-equalizzatore viene dapprima *campionato* nei suoi istanti significativi. La frequenza e la fase dell'operazione di campionamento debbono essere scelte in modo che siano soddisfatte due condizioni. In primo luogo gli istanti significativi del segnale amplificato e

equalizzato debbono essere in corrispondenza biunivoca con gli analoghi istanti del segnale di linea in trasmissione. Inoltre le ampiezze dei campioni estratti debbono essere di valore tale da facilitare il compito del decisore.

Il rispetto di tali due condizioni impone che la frequenza di campionamento sia numericamente uguale al ritmo di cifra in linea  $F_L$ . Per ciò che riguarda invece la fase temporale di campionamento, questa deve essere scelta, con un opportuno criterio, in un intervallo di durata uguale a  $T_L$  e viene determinata fissando, in detto intervallo, il cosiddetto *istante di decisione*.

Se l'attenuazione di linea fosse completamente compensata, dall'amplificatore, se gli effetti di disturbo legati alla deformazione degli impulsi fossero completamente eliminati dall'equalizzatore e se non fossero presenti altre cause di degradazione, i campioni estratti dal segnale amplificato e equalizzato dovrebbero presentare livelli che coincidono con quelli del segnale di linea in trasmissione e che quindi sarebbero in corrispondenza biunivoca con la sequenza di linea. In realtà, dato che le condizioni ideali sopra delineate non sono completamente conseguibili (soprattutto per ciò che riguarda l'assenza di ulteriori cause di disturbo), il livello di ogni campione si discosta da quello trasmesso e le ampiezze di questi scostamenti hanno caratteristiche aleatorie.

Conseguentemente, per ogni campione estratto, è possibile solo effettuare una *stima* del relativo livello. A questo scopo i campioni estratti dal segnale di linea amplificato e equalizzato vengono confrontati con  $q - 1$  livelli di riferimento (*soglie di decisione*), se  $q$  è il numero dei livelli del segnale di linea in trasmissione. Ciò consente di effettuare una stima della cifra che corrisponde a questo campione. Il punto del ricevitore di banda base in cui viene effettuata questa stima è detto *punto di decisione*.

Il funzionamento del decisore è quindi strettamente legato alla determinazione dei suoi punti di lavoro nel piano tempi-ampiezze. Questi punti, che sono in numero uguale a  $q-1$ , hanno un'ascissa comune, che coincide con l'istante di decisione e ordinate che sono uguali alle ampiezze delle relative soglie; sono quindi allineati su una retta parallela all'asse delle ordinate.

Normalmente i punti di lavoro del decisore sono scelti equidistanti e localizzati nella mezzeria di due livelli adiacenti in trasmissione. Il criterio per la determinazione dell'istante di decisione è basato sul conseguimento della massima resistenza ai disturbi del ricevitore e conduce normalmente a campionare gli impulsi costituenti il segnale amplificato e equalizzato nei punti in cui ogni impulso assume ampiezza massima.

Le alterazioni di livello dei campioni su cui opera il decisore vengono

quanto più possibile ridotte dall'amplificatore-equalizzatore. Alle alterazioni residue (che sono in ogni caso inevitabili) corrispondono degradazioni dell'informazione numerica relativa al segnale di dati emesso dal DTE sorgente.

Infatti, quando, nel punto e nell'istante di decisione, al segnale utile risulta sovrapposta una somma di disturbi di ampiezza complessiva superiore, in valore assoluto, alla semi-differenza tra due consecutive soglie di decisione, si può verificare un non corretto riconoscimento del livello trasmesso. A questo mancato riconoscimento corrisponde uno scostamento della sequenza di linea ricevuta rispetto a quella trasmessa. Tale scostamento va sotto il nome di *errore di linea*.

### IL DECISORE

Consideriamo il caso di un segnale di linea a 3 livelli codificato con cifre uguali a  $-1$ ,  $0$  e  $+1$  e corrispondenti ai livelli  $-2d$ ,  $0$  e  $+2d$ . Le soglie di decisione sono poste a metà ampiezza degli impulsi del segnale di linea, e cioè sono fissate a  $-d$  e a  $+d$ . L'istante di decisione è fissato in corrispondenza al massimo di ogni impulso. Con queste scelte, si può stabilire una semplice regola di stima; questa può essere così formulata:

ampiezza maggiore di $+d$	→	cifra $+1$
ampiezza compresa tra $+d$ e $-d$	→	cifra $0$
ampiezza inferiore a $-d$	→	cifra $-1$

ove, per ogni evento descritto a sinistra della freccia, è indicato, a destra, il risultato dell'operazione di stima.

Con riferimento a queste condizioni, nel punto e nell'istante di decisione possono verificarsi i seguenti casi:

- a un livello qualunque tra i tre possibili si aggiunge o si sottrae una somma di disturbi di ampiezza complessiva inferiore in valore assoluto al livello di soglia  $d$ ;
- a un livello  $+2d$  si aggiunge, o a un livello  $-2d$  si sottrae, una somma di disturbi di ampiezza complessiva arbitraria;
- a un livello  $+2d$  si sottrae, o a un livello  $-2d$  si aggiunge, una somma di disturbi di ampiezza complessiva superiore, in valore assoluto, al livello di soglia  $d$ ;
- a un livello  $0$  si aggiunge o si sottrae una somma di disturbi di ampiezza complessiva superiore, in valore assoluto, al livello di soglia  $d$ .

Allora nei primi due casi il riconoscimento è corretto, mentre nei secondi due casi si verifica errore di linea.

Passando ora alle cause di disturbo che possono produrre errori di linea, si può affermare che i principali disturbi di cui può essere affetto il segnale amplificato e equalizzato sono l'*interferenza intersimbolica* e i *rumori* di varia origine. Concentriamo qui la nostra attenzione sull'interferenza intersimbolica. Questa è dovuta alle distorsioni lineari subite dagli impulsi di cui è costituito il

segnale di linea e alle conseguenti deformazioni subite da ogni impulso elementare, secondo quanto già detto in precedenza.

Il risultato di questa deformazione è mostrato, ad esempio, in Fig. II.14a, ove è riportato l'andamento di un impulso elementare e ove, sull'asse dei tempi, sono indicati gli istanti di campionamento; la fase di questi istanti è stata fissata, come d'uso, con il criterio che, nell'istante di decisione, l'impulso assuma ampiezza massima. Con questa assunzione e per fissare le idee, indichiamo con  $t = 0$  l'istante di decisione dell'impulso in esame.

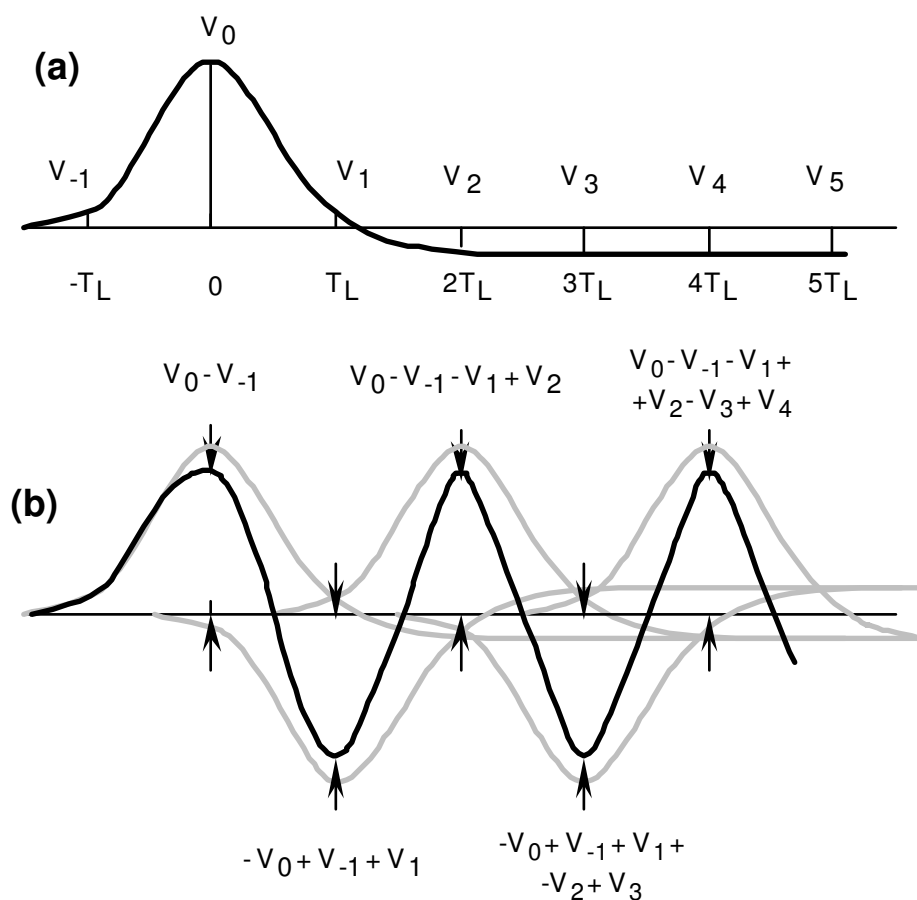


Fig. II.14 - Esempio di interferenza intersimbolica:

- (a) andamento dell'impulso elementare;
- (b) andamento del segnale di banda base associato alla sequenza di linea  $+1, -1, +1, -1, +1$ .

Si vede allora che, negli istanti di campionamento in anticipo ( $t = -kT_L, k = 1, 2 \dots$ ) o in ritardo ( $t = kT_L, k = 1, 2 \dots$ ) rispetto all'istante  $t = 0$ , le ampiezze dell'impulso non sono nulle. Tra tali ampiezze distinguiamo quelle che precedono (*pre-cursori*) e che seguono (*post-cursori*) l'ampiezza massima. Mentre i pre-cursori sono in numero limitato, i post-cursori possono essere teoricamente in numero infinitamente grande.

A causa dei fenomeni ora descritti e come mostrato in Fig. II.14b, l'ampiezza di ciascun impulso nel suo istante di decisione viene alterata (in eccesso o in difetto) dai post-cursori di tutti gli impulsi precedenti e dai precursori di qualche impulso susseguente. Questa alterazione costituisce l'interferenza intersimbolica.

Facciamo notare che, sebbene le distorsioni lineari siano deterministiche, l'interferenza intersimbolica è un disturbo aleatorio, che dipende dalla statistica delle cifre che compongono il segnale di linea, oltre che, ovviamente, dalla risposta in frequenza del canale trasmissivo e dalla forma dell'impulso elementare.

Quanto si è ora detto circa l'origine dell'interferenza intersimbolica consente di spiegare anche quali possono essere i criteri per eliminare questa forma di disturbo.

A tale scopo è sufficiente che l'impulso elementare a valle dell'amplificatore-equalizzatore presenti pre-cursori e post-cursori che sono esattamente nulli. Questo è il *primo criterio di Nyquist*, che, riferito al dominio della frequenza, può essere così formulato: un canale trasmissivo presenta interferenza intersimbolica nulla se si comporta come un filtro passa-basso con una risposta d'ampiezza che è simmetrica, nel piano frequenze-ampiezze, rispetto al punto di coordinate  $(F_L/2, T_L/2)$  e con una risposta di fase che è lineare con la frequenza.

Un esempio di risposta d'ampiezza per un canale trasmissivo che presenta interferenza intersimbolica nulla è mostrato in Fig. II.15a. Si tratta di una risposta del tipo a *coseno rialzato*, che ha valore costante uguale a  $T_L$  tra  $f = 0$  e  $f = (1 - \beta) F_L/2$  e che assume valore nullo per  $f \geq (1 + \beta) F_L/2$ . Il parametro  $\beta$ , che compare in queste espressioni e che è chiamato *coefficiente di arrotondamento*, può assumere tutti i valori compresi tra 0 e 1.

Al variare di  $\beta$  si possono avere vari tipi di risposte, le quali passano tutte per il punto di coordinate  $(F_L/2, T_L/2)$  e presentano un fronte di attenuazione simmetrico rispetto alla frequenza  $F_L/2$ . La frequenza di taglio della risposta è compresa fra un minimo uguale a  $F_L/2$  e un massimo dato da  $F_L$ .

Per  $\beta = 0$  (Fig. II.15b) la risposta di ampiezza è quella di un filtro passa-basso ideale, mentre, al crescere di  $\beta$ , i fronti di attenuazione diventano meno ripidi. In Fig. II.15c è presentato il caso  $\beta = 1$ .

Dal primo criterio di Nyquist possiamo trarre un'importante conclusione: la trasmissione di un segnale di banda base con ritmo di cifra  $F_L$  può essere effettuata, senza le degradazioni introdotte dall'interferenza intersimbolica, se il

canale trasmissivo presenta una banda passante con frequenza di taglio  $B$  non inferiore a  $F_L/2$ .

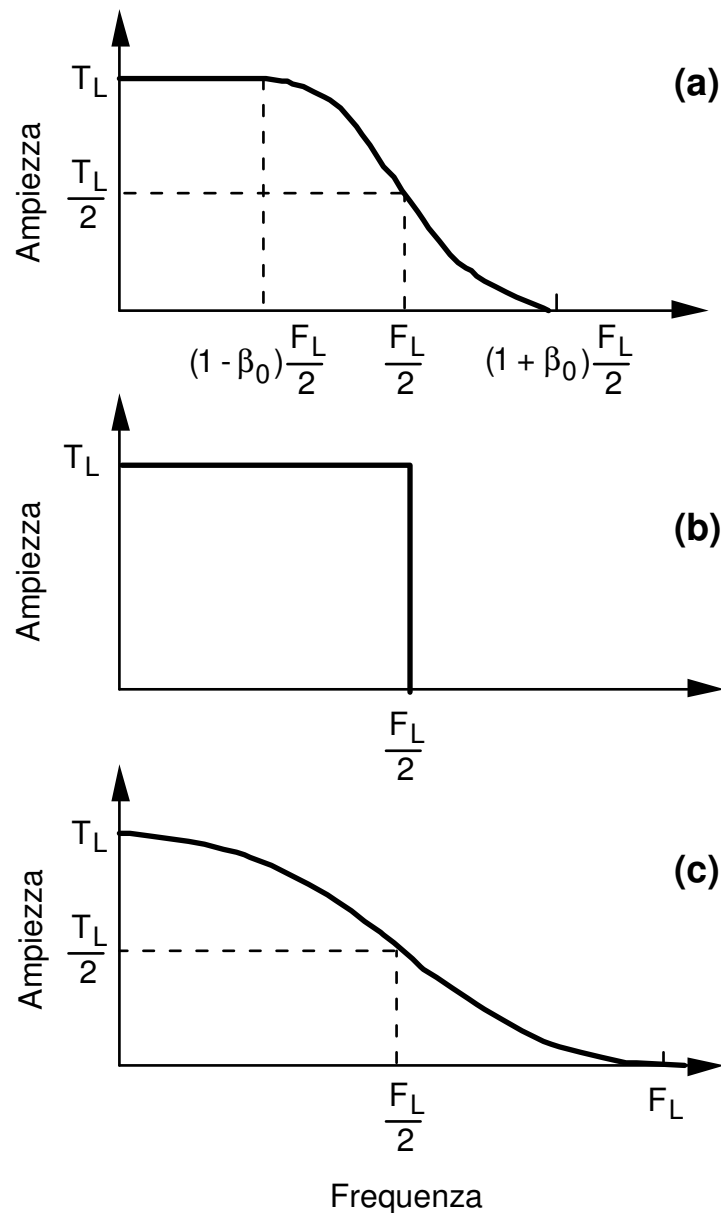


Fig. II.15 - Esempi di risposta in ampiezza a coseno rialzato con coefficiente di arrotondamento uguale a:

- (a) un valore generico  $0 < \beta < 1$ ;
- (b) 0;
- (c) 1.

Possiamo aggiungere che quando tale frequenza di taglio  $B$  è proprio uguale a  $F_N = F_L/2$ , e cioè a un valore che è chiamato *frequenza di Nyquist del segnale di linea*, il canale trasmissivo deve avere una risposta in frequenza che è di tipo passa-basso ideale (Fig. II.15b) e che quindi non è fisicamente

realizzabile. Tuttavia condizioni di realizzabilità fisica possono approssimarsi aumentando la larghezza di banda del canale, come mostrato nelle Fig. II.15a e Fig. II.15c.

Pertanto la frequenza di Nyquist del segnale di linea è la minima larghezza di banda di un canale trasmissivo che sia in grado di trasferire, senza interferenza intersimbolica, un segnale di linea con ritmo di cifra uguale a  $F_L$ .

Per concludere rimane da definire quale sia il parametro del ricevitore che qualifica la fedeltà di ricezione. Tale parametro è il *tasso di errore in linea*, e cioè la frequenza media con cui si verifica un evento di errore di linea.

È da sottolineare che il tasso d'errore in linea non coincide in generale con quello visibile dal DTE collettore, anche se il rapporto tra questi due tassi di errore è costante. Per distinguerli, il secondo verrà chiamato nel seguito *tasso di errore binario*.

Mostriamo come un errore di linea non determina necessariamente un errore di cifra binaria. Se il segnale di linea è ottenuto da una codifica bipolare semplice e se il livello in ricezione, corrispondente all'emissione di una cifra +1, viene stimato come il livello associato alla cifra -1, si verifica errore di linea. Non si ha invece errore sulla sequenza binaria emessa dal decodificatore in quanto, in base alla corrispondenza della codifica bipolare, le cifre +1 e -1 sono decodificate entrambe in una cifra binaria 1.

In generale il legame tra tasso di errore in linea e tasso di errore binario dipende dal rapporto del numero di errori in decodifica. Tale rapporto è ovviamente legato alla corrispondenza tra le sequenze di sorgente e di linea stabilita dal codice di linea.

Si può dimostrare che il tasso di errore in linea aumenta con legge monotona al diminuire del rapporto segnale/rumore nel punto di decisione. Questo risultato ci consente di affermare che la funzione specifica dell'equalizzatore è quella di eliminare (o almeno ridurre) l'interferenza intersimbolica, massimizzando al tempo stesso il rapporto segnale/rumore nel punto di decisione. Questi due requisiti sono in generale tra loro contrastanti in quanto la eliminazione (o la riduzione) dell'interferenza intersimbolica richiede all'equalizzatore una banda passante larga, mentre per ridurre il rumore sarebbe necessario avere una banda passante stretta. Il soddisfacimento di queste esigenze contrastanti può essere ottenuto solo con una soluzione di compromesso.

Un metodo grafico che permette di valutare l'entità massima dell'interferenza intersimbolica nel punto di decisione e di scegliere la posizione più conveniente dei punti di lavoro del decisore è il cosiddetto *diagramma ad occhio*. Questo può essere tracciato, nel piano tempi-ampiezze, sovrapponendo tutte le possibili forme d'onda del segnale di linea amplificato e equalizzato, che si presentano nei successivi intervalli temporali di durata uguale all'intervallo di cifra in linea  $T_L$ .

In Fig. II.16 è riportato un esempio di diagramma ad occhio relativo a un segnale di linea in codice bipolare e a condizioni di interferenza intersimbolica nulla. Nel piano tempi-ampiezze si distinguono, tramite campitura, due regioni che sono gli *occhi* visibili nel punto di decisione.

Nel caso in cui fosse presente interferenza intersimbolica, ognuno dei due occhi presenterebbe una riduzione di area, ma continuerebbe a rappresentare il luogo dei punti di lavoro del decisore, che assicurano, in assenza di altri disturbi, una immunità da errori di linea nella discriminazione tra due livelli adiacenti in ampiezza.

Scegliamo allora un punto di lavoro all'interno di un occhio e valutiamo le distanze minime di questo punto dai bordi dell'occhio, misurate nelle direzioni delle ordinate e delle ascisse. Queste due distanze rappresentano, rispettivamente, il margine contro errori di linea legati alla presenza di altri disturbi e a errati posizionamenti dell'istante di decisione. Ciò mette in evidenza che il miglior punto di lavoro è quello posizionato nel centro dell'occhio, e cioè nel punto in cui le distanze suddette assumono valore massimo.

Con i concetti precedentemente introdotti possiamo infine chiarire come si può valutare la resistenza ai disturbi presentata dal ricevitore. A tale scopo utilizziamo il diagramma ad occhio e misuriamo la distanza tra i bordi dell'occhio lungo la retta verticale passante per il punto di lavoro. Questa distanza, che è chiamata *ampiezza dell'occhio*, può essere assunta come misura della resistenza ai disturbi, nel senso che quanto più l'ampiezza dell'occhio è piccola, tanto minore è detta resistenza.

Rispetto alle condizioni ideali considerate in Fig. II.16, una riduzione della resistenza ai disturbi del ricevitore si può avere sostanzialmente per due motivi principali:

- 1) a causa di imperfezioni nel processo di decisione (dovute, ad esempio, a spostamenti dei punti di lavoro rispetto ai centri degli occhi);
- 2) per effetto dell'interferenza intersimbolica o di altre fonti di distorsione che provocano una riduzione dell'area dell'occhio.

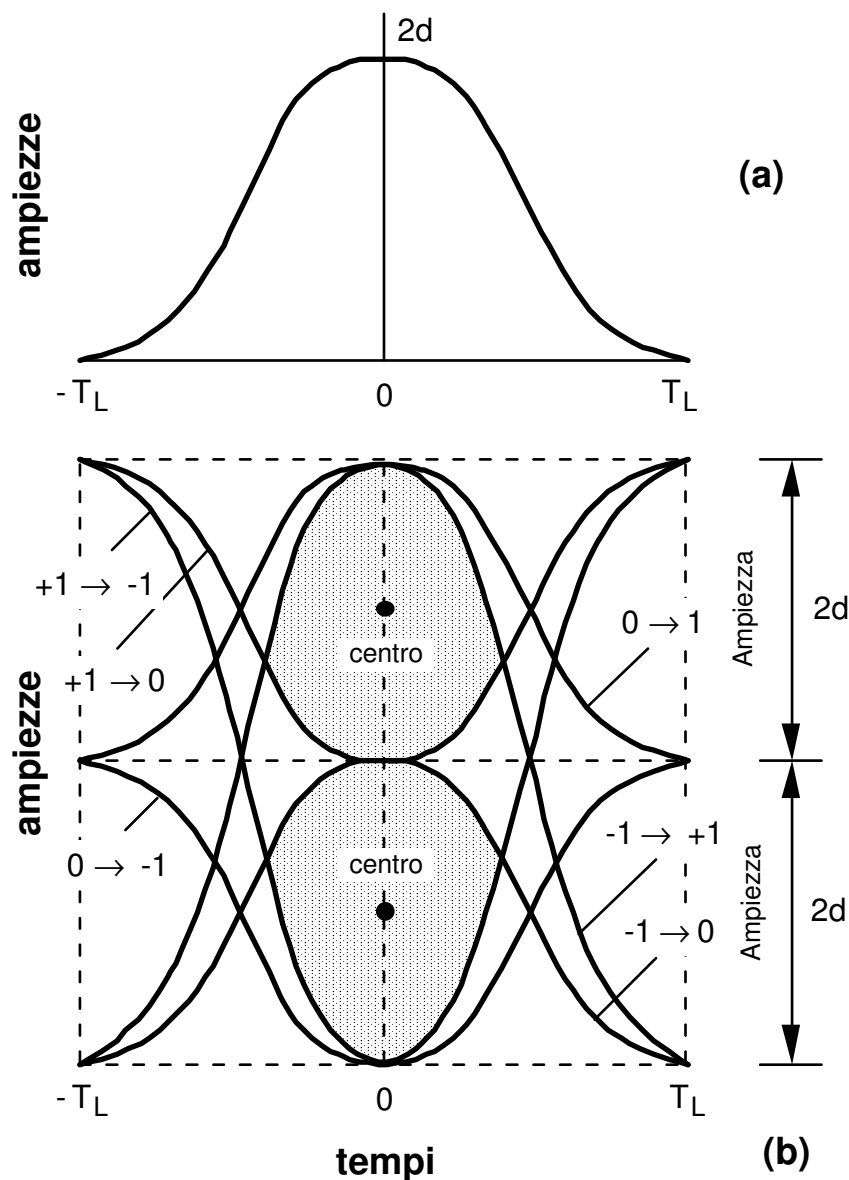


Fig. II.16 - Diagramma ad occhio (b) di un canale trasmissivo ad interferenza intersimbolica nulla con segnale di linea a tre livelli e con impulso elementare rappresentato in (a).

Circa il punto 1), sottolineiamo che spostamenti dell'istante di decisione, rispetto alla scelta ipotizzata in Fig. II.16, sono inevitabili a causa delle fluttuazioni spurie di fase con cui può essere ricostruito il segnale di temporizzazione in ricezione. Queste fluttuazioni spurie sono denominate *jitter di temporizzazione*.

Come appare dalla Fig. II.16 l'ampiezza ideale dell'occhio  $A_I$  è uguale a  $2d$ . In presenza di interferenza intersimbolica e di altre cause di deformazione degli impulsi, tale ampiezza si riduce a un valore  $A_R$  minore di  $A_I$ . In queste condizioni l'ulteriore causa di disturbo è rappresentata dal rumore. Per mantenere invariato il tasso di errore rispetto al caso ideale considerato in Fig.

II.16, occorre che il rapporto segnale/rumore nel punto di decisione venga incrementato di una quantità  $\Delta S/N$  esprimibile, in unità logaritmiche, con

$$\frac{\Delta S}{N} = -20 \lg_{10} \frac{A_R}{A_I}$$

Da questa relazione si vede allora che, al crescere dell'interferenza intersimbolica e quindi al diminuire di  $A_R$ , l'incremento  $\Delta S/N$  aumenta, tendendo asintoticamente all'infinito quando  $A_R$  tende a zero.

Questo comportamento *a soglia* è distintivo della trasmissione numerica rispetto a quella analogica; in questo secondo caso, infatti, un incremento di rapporto segnale/rumore rende sempre possibile, almeno in linea teorica, il conseguimento di una desiderata qualità di trasmissione.

### II.3.2 *Trasmissione in banda traslata*

La tecnica di trasmissione in banda traslata è largamente impiegata per il trasferimento di dati su canale telefonico analogico (*dati in banda fonica*) e su quello del gruppo primario FDM. Mentre nel primo caso la larghezza di banda lorda del canale è di 4 kHz, nel secondo è di 48 kHz. Le classi di servizio sono quelle riportate in Tab. II.1 fino al n° 7: in particolare per ritmi binari fino a 9600 bit/s (classi dal n° 1 al n° 6) si utilizza il canale telefonico, mentre per il ritmo binario di 48 kbit/s è previsto l'impiego del gruppo primario FDM.

Nel caso di trasmissione in banda traslata, intervengono, come già detto, vere e proprie operazioni di modulazione e di demodulazione.

Con riferimento agli schemi delle sezioni trasmittente e ricevente di un modem operante in banda traslata, riportati nelle Fig. II.8a, Fig. II.8b e Fig. II.8d, il segnale inviato all'ingresso del modulatore (*segnale modulante*) e quello uscente dal demodulatore sono segnali numerici di banda base.

Il primo è ottenuto dal segnale di dati emesso dal DTE sorgente con una codifica di linea; questa opera nello stesso modo descritto in § II.3.1, anche se con codici che presentano proprietà sostanzialmente diverse da quelle richieste nel caso di trasmissione in banda base.

Il segnale uscente dal demodulatore viene trattato da un ricevitore di banda base, che agisce con principi analoghi a quelli descritti in precedenza, e da un decodificatore che effettua l'operazione inversa a quella del codificatore di trasmissione. Pertanto in questa sezione ci limiteremo a considerare i principi di funzionamento del modulatore e del demodulatore.

Nelle tecniche di modulazione impiegate per la trasmissione dei dati, la

portante è generalmente di tipo sinusoidale, mentre il segnale modulante è un segnale numerico a due o più livelli. Un possibile codice a due livelli è quello NRZ, mentre i codici multilivello possono essere del tipo descritto in precedenza. Talvolta, per le ragioni che saranno chiarite nel seguito, si impiega una *codifica differenziale* in luogo di una *codifica diretta*.

La codifica diretta è quella precedentemente considerata, che pone in corrispondenza gruppi di  $\log_2 q$  cifre binarie con una cifra  $q$ -aria. La codifica differenziale si ottiene da quella diretta generando la cifra attuale  $a_j$  come *somma modulo  $q$*  della cifra precedente  $a_{j-1}$  e della cifra  $a_j'$  ottenuta da una codifica diretta: l'informazione è quindi associata alla differenza tra  $a_j$  e  $a_{j-1}$ .

Il ritmo di cifra in linea  $F_L$ , che è quello del segnale modulante e che, in questo contesto, è anche chiamato *velocità di modulazione*, è di norma numericamente uguale al ritmo binario di sorgente  $F_S$  nel caso di codifica a due livelli ed è a questo legato dalla (Eq. II.2) nel caso di codifica multilivello.

Le varie tecniche di modulazione impiegabili per trasmissione di dati possono essere confrontate sulla base di svariati elementi. Ne citiamo due, e cioè la *sensibilità ai disturbi* e la *efficienza spettrale*.

Per fornire una definizione di questi due parametri facciamo riferimento al tasso di errore binario e al rapporto segnale/rumore all'ingresso del demodulatore, che chiameremo, per brevità, *rapporto  $S/N$  in ricezione*. Questo è definito convenzionalmente come il rapporto tra la potenza media del segnale di linea in ricezione e la potenza di rumore in una banda di larghezza numericamente uguale alla metà del ritmo binario di sorgente (e cioè alla frequenza di Nyquist del segnale numerico binario).

Con questi riferimenti la sensibilità ai disturbi di una tecnica di modulazione è valutata comunemente in base al rapporto  $S/N$  in ricezione, che consente di ottenere un prefissato tasso di errore binario.

L'efficienza spettrale è invece misurata dal ritmo binario di sorgente che può essere trasferito per unità di larghezza di banda del canale trasmissivo, con un prefissato rapporto  $S/N$  in ricezione. Detta efficienza è tanto maggiore quanto minore è la larghezza di banda del segnale modulato, a parità di ritmo binario di sorgente e di rapporto  $S/N$  in ricezione.

La modulazione consiste nell'associare a ogni livello del segnale modulante un distinto valore di un parametro della portante, e cioè della sua ampiezza o della sua frequenza o della sua fase quando la modulazione è di tipo *semplice* ovvero un valore di due di questi parametri quando la modulazione è di tipo *misto*. Per ognuna di queste tecniche si parla di *livello di ampiezza*, di *frequenza* o di *fase* a seconda di quale sia il parametro della portante che viene

associato a un livello del segnale modulante.

Qui, per ragioni di brevità, ci limitiamo a considerare i casi di modulazione semplice in cui il parametro della portante su cui si agisce è la frequenza o la fase. Nel primo caso, corrispondente a una modulazione di frequenza, si parla di *modulazione a spostamento di frequenza* (Frequency Shift Keying o *FSK*). Il secondo caso corrisponde invece a una modulazione di fase e si riferisce alla cosiddetta *modulazione a spostamento di fase* (Phase Shift Keying o *PSK*). Per ciò che riguarda poi le modulazioni miste, daremo un cenno sulla modulazione mista di ampiezza e di fase.

La modulazione FSK, impiegata nella trasmissione di dati su canale telefonico, utilizza un segnale modulante a due livelli e associa ad ognuna delle due cifre binarie un determinato e diverso valore di frequenza. In corrispondenza della cifra "1" viene trasmessa una frequenza  $F_I$ , mentre in corrispondenza di un "0" è inviata una frequenza  $F_M$ , normalmente maggiore di  $F_I$ . Nel passaggio da una frequenza all'altra è preferibile mantenere la continuità di fase: ciò allo scopo di limitare l'occupazione spettrale del segnale modulato.

La scelta delle due frequenze  $F_M$  e  $F_I$ , che sono simmetriche rispetto alla frequenza  $F_p$  della portante, determina l'*indice di modulazione* del modulatore. Tale indice è espresso da

$$\text{Eq. II.3 } \mu = \frac{F_M - F_I}{F_S}$$

ove  $F_S$  è la frequenza numericamente uguale al ritmo binario di sorgente.

Lo spettro di densità di potenza di un segnale modulato FSK a due livelli è simmetrico rispetto alla frequenza  $F_p$  della portante e presenta un andamento che dipende dall'indice di modulazione. In Fig. II.17 è mostrato tale andamento per quattro diversi valori di  $\mu$ . Si può notare che la potenza del segnale modulato è ripartita su tutto l'asse delle frequenze; se tuttavia si definisce come larghezza di banda spettrale  $W$  quella che comprende il 95% della potenza del segnale modulato, si può dimostrare che  $W$  presenta un minimo per valori di  $\mu$  prossimi a 0,64.

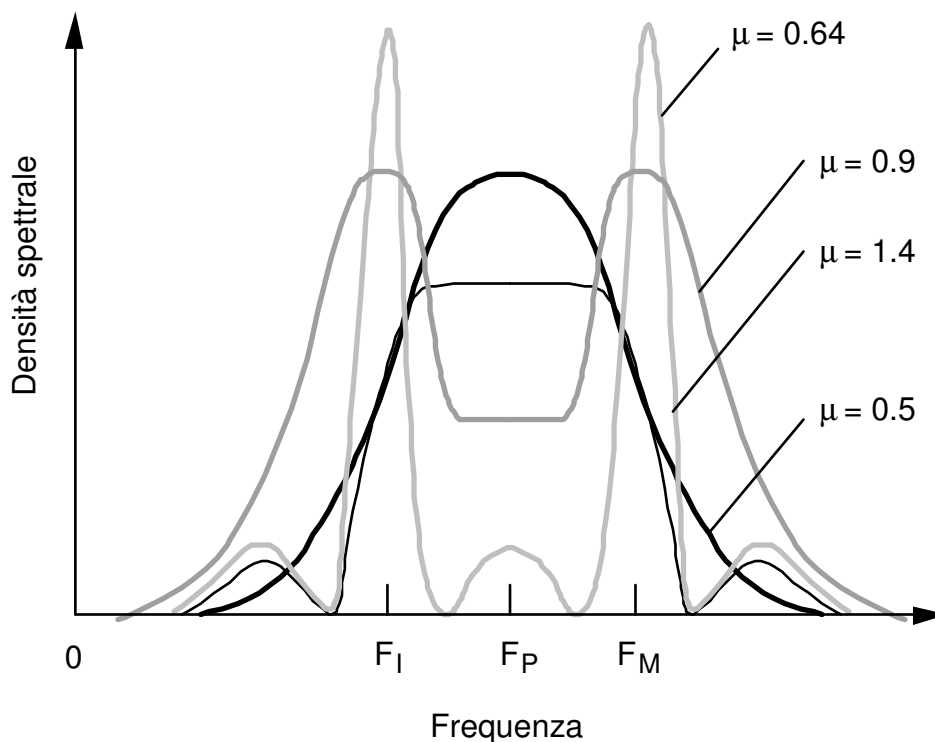


Fig. II.17 - Spettro di densità di potenza di un segnale modulato FSK a due livelli, per vari valori dell'indice di modulazione  $\mu$ .

### MODEM FSK NORMALIZZATI

Nella Racc. V.21 è normalizzato un modem FSK che opera con una velocità di modulazione uguale a 300 baud e con trasmissione asincrona. Esso consente una modalità di funzionamento duplice su linea a due fili. Ciò è reso possibile applicando il principio della divisione di frequenza: cioè il canale telefonico è suddiviso in due sub-canali, chiamati *canale inferiore* e *canale superiore* e centrati rispettivamente sulle frequenze 1080 Hz e 1750 Hz.

Circa la corrispondenza tra le cifre binarie del segnale di dati e le frequenze del segnale di linea, un "1" è trasmesso con le frequenze 980 Hz e 1600 Hz sul canale inferiore e su quello superiore, rispettivamente, mentre uno "0" è associato, nello stesso ordine, alle frequenze 1180 Hz e 1850 Hz.

Dato che in questo caso risulta  $F_M - F_I = 200$  Hz e poiché  $F_S = 300$  Hz, l'indice di modulazione  $\mu$  è uguale a  $2/3$ .

Il modem FSK normalizzato nella Racc. V 23 utilizza per le frequenze  $F_M$  e  $F_I$  i valori di 2100 Hz e di 1300 Hz. La frequenza di portante è uguale a 1700 Hz, mentre la velocità di modulazione è di 1200 baud. Dato che  $F_M - F_I = 800$  Hz e  $F_S = 1200$  Hz, risulta  $\mu = 800/1200 = 0,66$ , e cioè un valore dell'indice di modulazione che è prossimo a quello ottimo per ridurre la larghezza di banda spettrale  $W$  del segnale modulato.

La demodulazione di un segnale modulato FSK a due livelli può essere di tipo *incoerente*. Ad esempio si può utilizzare un *demodulatore a limitatore-discriminatore*. Il limitatore elimina la modulazione spuria di ampiezza presente nel segnale di linea in ricezione e mantiene un livello praticamente costante all'ingresso del discriminatore. Quest'ultimo trasforma poi le variazioni di frequenza in variazioni di ampiezza, dalle quali può essere estratto il segnale di dati ricevuto.

Il principale vantaggio di un modem FSK risiede nella sua semplicità realizzativa, dovuta sostanzialmente alla possibilità di una rivelazione incoerente. È inoltre possibile evitare talvolta il recupero del segnale di temporizzazione.

Gli svantaggi più rilevanti sono la sua ridotta efficienza spettrale (legata alla larghezza di banda spettrale del segnale modulato, che, a parità di condizioni, è normalmente maggiore di quella ottenibile con altre tecniche di modulazione), e alla elevata sensibilità ai disturbi, che si traduce, rispetto ad altre tecniche di modulazione (ad esempio, al PSK a due livelli con ricezione differenzialmente coerente) in un aumento di circa 3 dB del rapporto  $S/N$  in ricezione. Questi inconvenienti fanno sì che la modulazione FSK a due livelli venga impiegata nel trasferimento di dati con ritmi binari non superiori a 1200 bit/s, almeno per ciò che riguarda l'utilizzazione dei canali telefonici.

Nella modulazione PSK, si associa ad ogni cifra  $q$ -aria  $a_j$  un valore di fase  $\theta_j$  della portante (*livello di fase*) scelto tra  $q$  possibili. Per facilitare il compito del ricevitore, questi valori di fase sono distribuiti con regolarità nell'intervallo  $[0, 2\pi]$ ; cioè ad esempio

$$\theta_k = k \frac{2\pi}{q} \quad (k = 1, \dots, q) \quad \text{ovvero} \quad \theta_k = k \frac{2\pi}{q} + \frac{\pi}{q} \quad (k = 1, \dots, q)$$

La codifica necessaria per ottenere la sequenza di linea a  $q$  livelli da quella binaria può essere diretta o differenziale. Normalmente si assume un numero di livelli uguale a una potenza intera di 2.

Per descrivere gli schemi di modulazione PSK si può rappresentare ciascun livello di fase  $\theta_k$  con un punto  $P_k = A \exp(j\theta_k)$  del piano complesso. Si ottengono allora i *diagrammi di costellazione* mostrati in Fig. II.18, ove sono considerati i casi di modulazione a 2, 4 e 8 livelli.

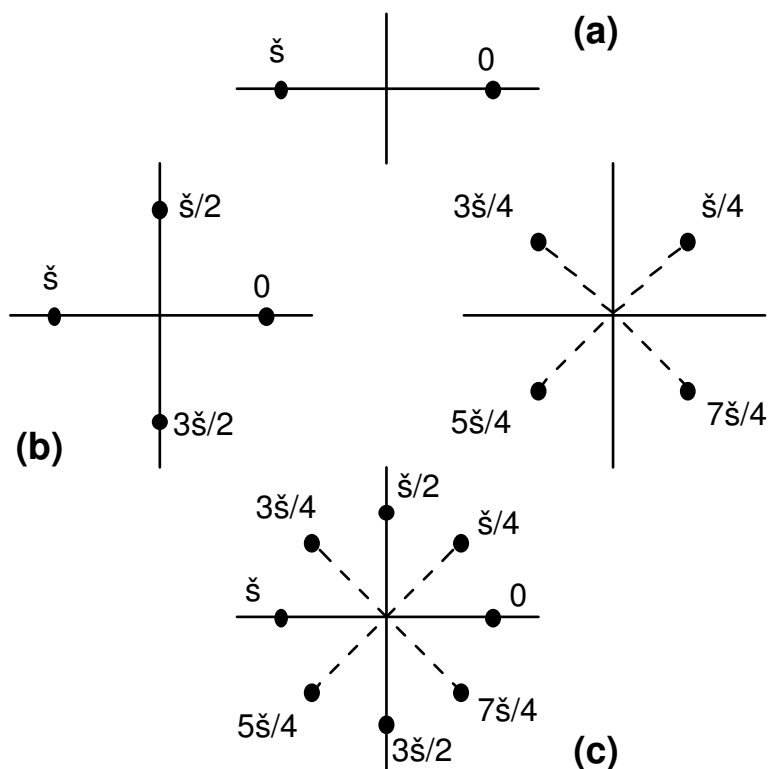


Fig. II.18 - Diagrammi delle costellazioni per un segnale modulato PSK a due livelli (a), a quattro livelli (b) e a otto livelli (c).

La demodulazione di un segnale PSK con codifica diretta deve essere di tipo *coerente*; cioè per la sua attuazione è necessario che in ricezione la portante venga ricostruita in frequenza e in fase. Questa esigenza crea tuttavia complicazioni notevoli dovute alla difficoltà di acquisire e di mantenere stabile la fase della portante.

Circa la difficoltà di acquisizione, occorre tenere conto che la portante può essere ottenuta dal segnale modulato mediante operazioni non lineari, che permettono di eliminare la modulazione. Questa tecnica di acquisizione conduce tuttavia a una ambiguità nella fase ricostruita. Nella trasmissione di dati si preferisce allora eliminare questa ambiguità mediante l'impiego della codifica differenziale.

Infatti ciascuna cifra della sequenza di linea originaria a  $q$  livelli viene individuata da un *cambiamento di fase* anziché da un valore assoluto di questa. Poiché il salto di fase viene sempre riconosciuto correttamente qualunque sia la posizione del segnale di riferimento, una demodulazione coerente di un segnale modulato PSK con codifica differenziale non pone problemi di ambiguità nell'acquisizione della fase della portante.

Per ciò che riguarda il problema legato alla stabilità della fase della portante, questo può essere superato sempre nel caso di impiego della codifica

differenziale. In questo caso infatti la demodulazione può essere *differenzialmente coerente*, e cioè può essere effettuata mediante una rivelazione per confronto fra le fasi associate a due cifre successivamente ricevute.

La differenza sostanziale tra la rivelazione coerente e quella differenzialmente coerente sta nel fatto che, nel primo caso, si paragona la fase del segnale ricevuto con un riferimento da considerare privo di rumore, mentre nel secondo caso vengono paragonate le fasi di due segnali entrambi corrotti dal rumore. Questa circostanza produce una degradazione nelle prestazioni del sistema.

Lo spettro di densità di potenza di un segnale modulato PSK a due livelli (cfr. Fig. II.18a) può essere ottenuto osservando che questo tipo di modulazione equivale a una modulazione di ampiezza a banda laterale doppia e a portante soppressa. Per utilizzare le proprietà spettrali di quest'ultima occorre determinare lo spettro del segnale modulante binario.

A tale scopo, per ragionare in termini intuitivi, ci riferiamo al caso in cui tale segnale presenta massima larghezza di banda: ciò si verifica quando si ha alternanza regolare di "1" e di "0". In queste condizioni la frequenza fondamentale del segnale modulante coincide numericamente con la metà del ritmo binario di sorgente ed è cioè uguale a  $F_S/2$ . Si può allora concludere che lo spettro del segnale modulato PSK a due livelli è centrato intorno alla frequenza  $F_p$  della portante e presenta una larghezza di banda che è numericamente uguale a  $F_S$ .

Passando a un segnale modulato PSK multilivello, questo presenta una larghezza di banda spettrale che è inversamente proporzionale al numero  $\log_2 q$  di cifre binarie poste in corrispondenza con una cifra  $q$ -aria. Così, ad esempio, l'occupazione di banda di un segnale PSK a 4 livelli o a 8 livelli è circa la metà o un terzo di quella di un PSK a 2 livelli.

Circa la sensibilità ai disturbi, nel caso di rivelazione differenzialmente coerente il rapporto  $S/N$  in ricezione necessario per ottenere un prefissato tasso di errore aumenta al crescere del numero dei livelli. Questo spiega perché questa tecnica di modulazione non viene impiegata per un numero di livelli di fase maggiore di 8.

Per aumentare ulteriormente il numero di livelli, si procede con una modulazione mista di ampiezza e di fase, nella quale ogni fase può assumere, ad esempio, due o più ampiezze.

### MODEM NORMALIZZATI DI TIPO PSK E A MODULAZIONE MISTA

Nella Racc. V.26 è normalizzato un modem PSK a 4 livelli, con ricezione differenzialmente coerente e per trasmissione sincrona.

Il ritmo binario di sorgente è di 2400 bit/s e, conseguentemente, la velocità di modulazione è di 1200 baud. La frequenza di portante è di 1800 Hz.

Nella codifica differenziale la sequenza binaria di sorgente è divisa in coppie di cifre binarie consecutive (*dibit*). La corrispondenza fra dibit e cambiamenti di fase può essere dei seguenti due tipi

Dibit	Cambiamento di fase	
	Soluzione A	Soluzione B
00	0	$+\pi/4$
01	$+\pi/2$	$+3\pi/4$
10	$+\pi$	$+5\pi/4$
11	$+3\pi/2$	$+7\pi/4$

Conseguentemente i diagrammi di costellazione del segnale modulato sono quelli mostrati in Fig. II.18b.

Un esempio di modem PSK a 8 livelli con ricezione differenzialmente coerente e per trasmissione sincrona è offerto dalla Racc. V.27.

In questo caso il ritmo binario di sorgente è di 4800 bit/s, mentre la velocità di modulazione è di 1600 baud. La frequenza di portante è di 1800 Hz.

La codifica differenziale pone in corrispondenza terne di cifre binarie (*tribit*) con cambiamenti di fase, secondo la corrispondenza seguente

Tribit	Cambiamento di fase	Tribit	Cambiamento di fase
001	0	111	$+\pi$
000	$+\pi/4$	110	$+5\pi/4$
010	$+\pi/2$	100	$+3\pi/2$
011	$+3\pi/4$	101	$+7\pi/4$

Il diagramma di costellazione del segnale modulato è quello mostrato in Fig. II.18c.

Una modulazione mista di ampiezza e di fase è impiegata dal modem normalizzato dalla Racc. V.29, da utilizzare per trasmissione sincrona.

Il ritmo binario di sorgente è al massimo di 9600 bit/s, ma può essere ridotto a 7200 bit/s e a 4800 bit/s. In ogni caso la velocità di modulazione è di 2400 baud. Nel caso in cui  $F_S = 9600$  bit/s, la modulazione è a 16 livelli, di cui 8 di fase e 2 di ampiezza, mentre per  $F_S = 7200$  bit/s o 4800 bit/s, il numero di livelli si riduce a 8 e a 4 rispettivamente.

Per realizzare questo schema di modulazione, nel caso  $F_S = 9600$  bit/s la sequenza binaria di sorgente è suddivisa in gruppi di 4 cifre binarie (*quadribit*). Una prima cifra binaria (in ordine di tempo) di ogni quadribit determina l'ampiezza, mentre le altre 3 cifre binarie

definiscono un cambiamento di fase secondo la corrispondenza precisata dalla Racc. V.27. I valori che possono essere assunti dalla fase e dall'ampiezza della portante sono riportati nel diagramma di costellazione in Fig. II.19a. Nella stessa figura sono anche riportati gli analoghi diagrammi nel caso in cui il ritmo binario di sorgente sia ridotto a 7200 bit/s (Fig. II.19b) e a 4800 bit/s (Fig. II.19c).

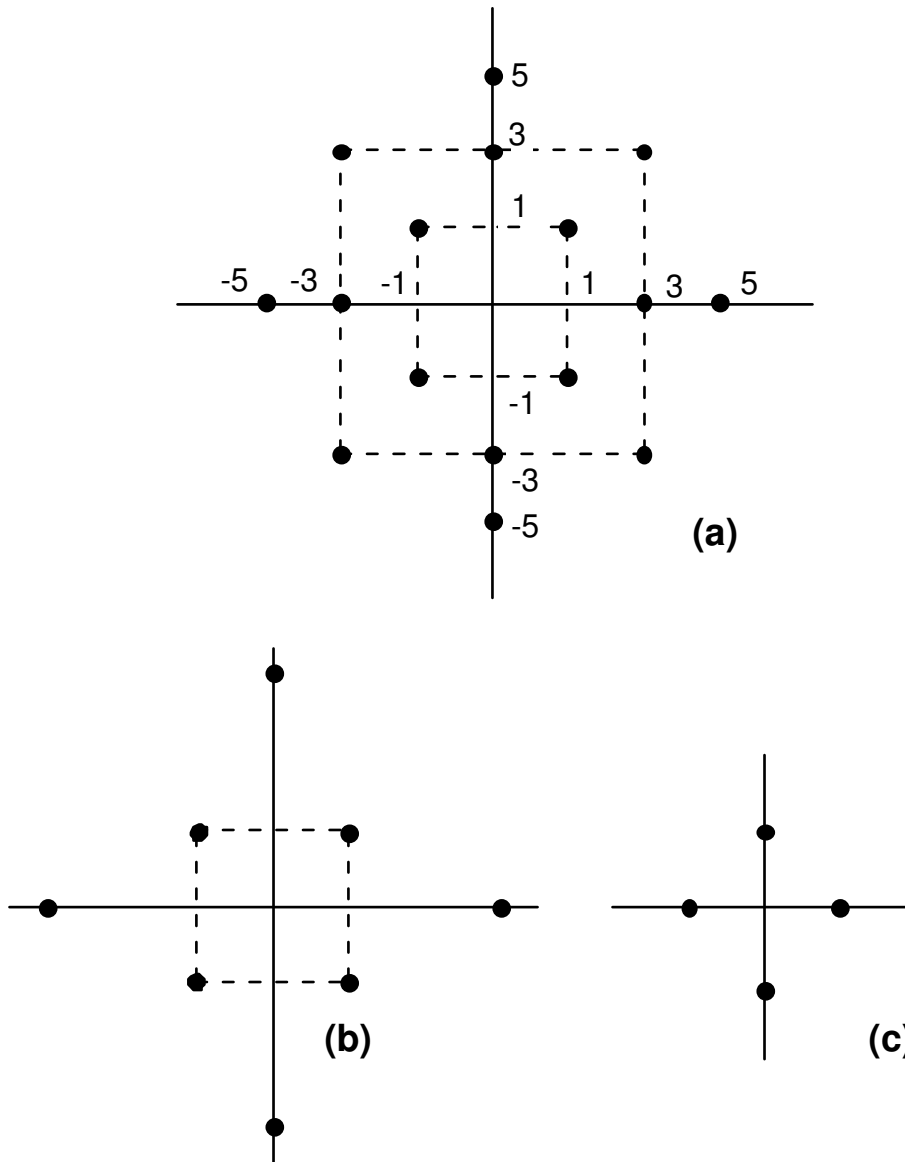


Fig. II.19 - Diagramma delle costellazioni per un segnale modulato misto in ampiezza e in fase a 16 livelli secondo la Racc. V.29 (a) e diagrammi ridotti con 8 livelli (b) e con 4 livelli (c).

### II.3.3 Condizione di sincronismo

Nella trasmissione dei dati, l'informazione emessa dal DTE sorgente e ricevuta dal DTE collettore è in forma numerica. Analoga affermazione può essere fatta per ciò che riguarda i segnali di banda base nelle sezioni

trasmittente e ricevente di un modem per trasmissione in banda base o in banda traslata.

Conseguentemente sia i DTE che i DCE debbono funzionare sotto il controllo di coppie di segnali di temporizzazione (*sincrosegnali*), uno in trasmissione e l'altro in ricezione. Con riferimento specifico ai DCE, nella sezione trasmittente del modem il sincrosegnale in trasmissione è utilizzato per definire gli istanti significativi in cui il segnale di banda base in trasmissione assume i livelli associati alla sequenza di linea in trasmissione. Il sincrosegnale in ricezione è invece utilizzato per campionare il segnale di banda base in ricezione e per consentire quindi la stima dei livelli assunti da questo solo in detti istanti significativi.

Affinché questi istanti, nel trasmettitore e nel ricevitore, siano in corrispondenza biunivoca, come è necessario per un corretto trasferimento dell'informazione associata alla sequenza numerica emessa dalla sorgente, i due sincrosegnali debbono avere la stessa frequenza. Questa condizione si esprime dicendo che i due apparati modem debbono essere in *sincronismo di cifra*.

In una trasmissione sincrona, l'asse dei tempi su cui sono cadenzate le operazioni del trasmettitore è suddiviso in intervalli, tutti di egual durata, tra due istanti significativi successivi. Il ricevitore deve allora ricostruire la frequenza di questa suddivisione.

Questa ricostruzione di sincrosegnale operata nel ricevitore di banda base viene effettuata da un dispositivo apposito, chiamato *estrattore del segnale di temporizzazione* (cfr. Fig. II.13). Questo agisce sul segnale amplificato/egualizzato e utilizza le transizioni del segnale di banda base in ricezione. Per assicurare che queste siano sufficientemente frequenti, in trasmissione si opera con la codifica di linea (nel modo precedentemente illustrato) e con uno *sparpagliamento* (scrambling) della sequenza di cifre binarie emessa dalla sorgente.

Questa seconda funzione ha lo scopo di modificare la sequenza di cifre binarie entrante nel codificatore di linea rispetto a quella emessa dal DTE sorgente. La modifica ha come risultato la casualizzazione delle cifre della prima sequenza rispetto a quelle della seconda. Ovviamente a valle del decodificatore di linea in ricezione deve essere presente un dispositivo che effettua l'operazione complementare a quella effettuata in trasmissione.

Oltre alla condizione di sincronismo di cifra, deve anche essere assicurata la *sincronizzazione di carattere*, in modo da consentire al DTE utilizzatore di riconoscere l'inizio e la fine di ogni parola di codice che, nella codifica dell'informazione effettuata dal DTE sorgente, rappresenta un carattere.

Nella trasmissione sincrona, la sincronizzazione di carattere si effettua al livello di gruppi di caratteri (chiamati *blocchi* o *trame*) e si fonda sull'utilizzazione di particolari combinazioni di cifre binarie. Tale funzione può essere considerata appartenente allo strato 2 e verrà quindi meglio chiarita nel successivo cap. III.

In una trasmissione asincrona, la sincronizzazione di carattere si effettua con il riconoscimento dei segnali di START e di STOP che delimitano ciascun carattere. La sincronizzazione di cifra è immediata in quanto il trasmettitore e il ricevitore dispongono di orologi locali di ugual frequenza nominale. Nel ricevitore il segnale di START sgancia, all'inizio di ciascun carattere, l'emissione di un sincrosegnale, che permette di campionare le cifre binarie contenute nel carattere. Anche in presenza di una leggera deriva di questo sincrosegnale, non viene compromesso il risultato dell'operazione, data la breve durata del carattere.

#### **II.4      Protocolli di strato fisico**

Nei par. II.1 e II.2 sono stati delineati il modello funzionale dello strato fisico e l'insieme delle funzioni che questo è chiamato a svolgere. In questa sezione saranno esposti i concetti principali che sono alla base della definizione dei protocolli di strato fisico, e cioè di quell'insieme di regole che devono essere seguite per permettere a due entità di strato fisico di attivare, di mantenere e di disattivare una PH-connessione.

È il caso di sottolineare che, a differenza di quanto avviene negli altri strati dell'architettura in cui al termine protocollo si associa esclusivamente un insieme di regole procedurali a cui le entità di strato devono attenersi, nello strato fisico il termine protocollo ha un'accezione più ampia. Questa comprende, infatti, anche la definizione delle caratteristiche strutturali della connessione fisica, e cioè, ad esempio, delle dimensioni e della struttura dei connettori, nonché dei valori di tensione e di corrente dei segnali scambiati dalle entità di strato fisico.

In base alla definizione generale dello strato fisico, i protocolli appartenenti a questo strato possono essere classificati in quattro classi che si riferiscono rispettivamente alle caratteristiche *meccaniche*, *elettriche*, *funzionali* e *procedurali* di una interfaccia tra entità di strato fisico (*interfaccia fisica*). In particolare nella definizione di tali caratteristiche, ci si riferirà allo schema generale di connessione fisica mostrato in Fig. II.7. I concetti che verranno ora esposti si riferiscono quindi alla definizione delle caratteristiche meccaniche,

elettriche, funzionali e procedurali dell'*interfaccia fisica tra DTE e DCE*.

Nel seguito, con riferimento a tale interfaccia, saranno trattati separatamente i quattro aspetti ora introdotti e presentati gli esempi più significativi di protocolli oggi in uso.

#### II.4.1 *Caratteristiche meccaniche*

Le caratteristiche meccaniche dell'interfaccia fisica tra DTE e DCE riguardano la definizione delle particolarità costruttive dei connettori che consentono l'accoppiamento fisico tra tali apparecchiature.

La necessità di definire le caratteristiche dei connettori risulta evidente se si considera che un requisito fondamentale da soddisfare è quello della *portabilità* di un apparato di telecomunicazione. Secondo tale requisito, uno stesso apparato, ad esempio un terminale, deve poter essere connesso di volta in volta ad un diverso punto d'accesso della rete, senza per questo dover mutare le modalità di accoppiamento fisico.

Oltre al precedente requisito, la definizione di un insieme limitato di connettori ha lo scopo di facilitare l'uso, nello stesso sistema di comunicazione, di apparati prodotti da costruttori diversi. In assenza di regolamentazioni a riguardo sarebbe necessario, infatti, disporre di un opportuno adattatore per ogni possibile tipo di accoppiamento.

È il caso di sottolineare che, comunque, la definizione di connettori di tipo normalizzato, se da un lato facilita l'accoppiamento tra apparati diversi, dall'altro rappresenta solo il primo e più semplice passo verso la possibilità di instaurare un effettivo colloquio tra tali apparati. A questo scopo sono necessari infatti sia altri elementi dello strato fisico sia le funzioni tipiche degli strati superiori.

In generale, un connettore è costituito da un certo numero di *piolini* (pin) a cui corrispondono altrettanti fili, che individuano i cosiddetti *circuiti di interfaccia*. Come sarà illustrato più in dettaglio nel seguito (par. II.4.3), ogni circuito ha lo scopo di espletare, durante la durata della connessione, una o più funzioni (ad es. quelle di trasferimento dei dati d'utente, di controllo e di temporizzazione). In altre parole, su ogni circuito viaggiano segnali che trasportano informazioni sullo stato dell'interfaccia e sull'evoluzione della connessione.

L'ITU-T non ha emesso direttamente alcuna raccomandazione inerente gli aspetti meccanici di una interfaccia fisica, ma ha assunto come riferimento le norme sviluppate in ISO. Tali norme (*standard*) sono riassunte in Tab. II.3.

Standard ISO	Numero piedini	Descrizione
2110	25	Modem in banda fonica per trasmissione seriale e parallela. Interfacce per reti pubbliche per dati. Reti telegrafiche e telex. Apparecchiature a chiamata automatica.
2593	34	Modem a larga banda (48 kbit/s)
4902	37/9	Modem in banda fonica o larga banda per trasmissione seriale
4903	15	Interfacce con reti pubbliche per dati

Tab. II.3 - Standard ISO relativi alla definizione delle caratteristiche meccaniche dei connettori.

Come si può osservare, i connettori si differenziano per il numero di piedini di cui sono composti, e cioè per il numero di circuiti che formano la relativa interfaccia. Questa differenza, come sarà illustrato più in dettaglio nel seguito, deriva dall'affermarsi nel tempo di una tecnica costruttiva che prevede la progressiva semplificazione delle interfacce fisiche attraverso la diminuzione del numero di circuiti. Tale diminuzione è ottenuta affidando ad uno stesso circuito funzioni prima espletate da fili diversi.

La Fig. II.20 - Caratteristiche meccaniche dei connettori: riporta, a titolo di esempio le caratteristiche costruttive dei connettori ISO 2110 a 25 piedini ed ISO 4903 a 15 piedini. Il connettore ISO 2110 è quello oggi utilizzato per l'accoppiamento tra terminali e modem in banda fonica e per l'accoppiamento di terminali con reti pubbliche per dati. Il connettore ISO 4903 è invece usato esclusivamente in questo secondo caso.

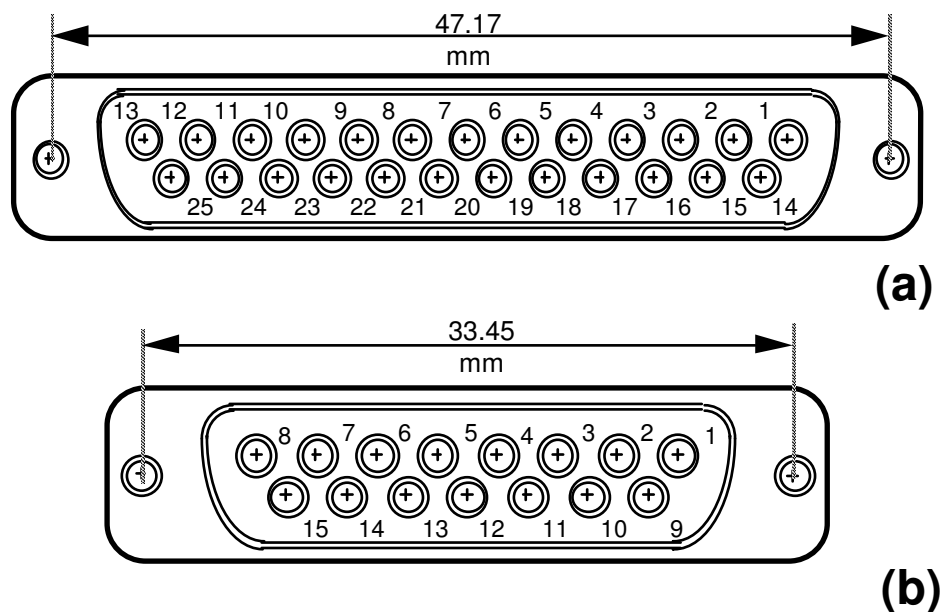


Fig. II.20 - Caratteristiche meccaniche dei connettori:

- (a) ISO 2110 (25 piedini);  
 (b) ISO 4903 (15 piedini).

#### II.4.2 Caratteristiche elettriche

In § II.4.1 è stato introdotto il concetto di interfaccia fisica costituita da un insieme di circuiti. Le caratteristiche elettriche di una connessione fisica hanno lo scopo di definire, da un lato, lo schema di riferimento di ogni circuito e, dall'altro, i valori delle grandezze elettriche (tensione, corrente, impedenza) misurate in opportune sezioni di questo.

I principali schemi di riferimento utilizzati nella definizione delle caratteristiche elettriche di un interfaccia sono presentati in Fig. II.21.

In tali schemi le due terminazioni del circuito sono costituite da un *generatore* e da un *ricevitore*. Il generatore identifica quell'entità (DTE o DCE) che emette il segnale su quel particolare circuito, mentre il ricevitore corrisponde all'entità che ha il compito di rivelare i segnali presenti sul circuito. Il filo che individua un circuito di interfaccia è detto *uscente* od *entrante* a seconda che l'entità a cui esso fa capo agisca rispettivamente come generatore o come ricevitore per quel particolare circuito.

La necessità di introdurre un generatore nello schema di riferimento di un circuito di interfaccia è giustificata dal fatto che i livelli di tensione interni alle apparecchiature elettroniche usati per rappresentare gli stati logici "0" e "1" sono nella generalità dei casi non sufficientemente alti da sopportare l'attenuazione del circuito di interfaccia. Il generatore, quindi, ha il compito di

convertire il segnale interno a bassa tensione in un segnale esterno a tensione più alta. Il ricevitore avrà ovviamente il compito di eseguire la funzione inversa.

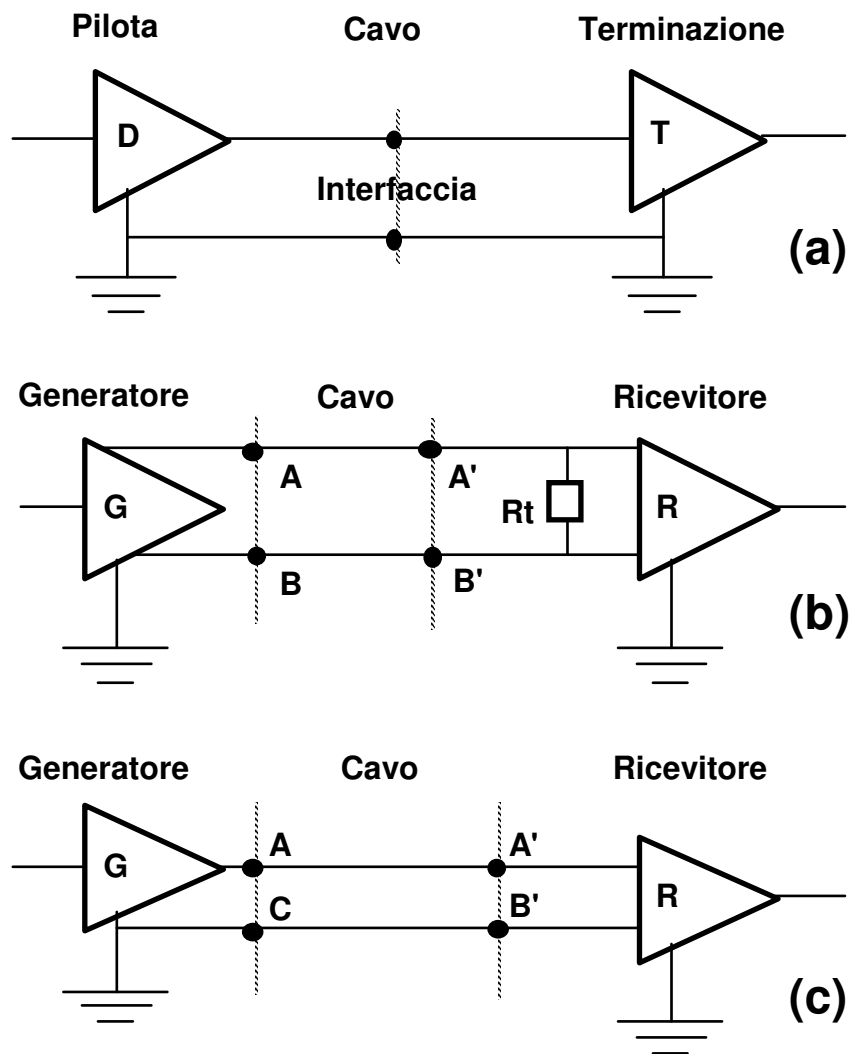


Fig. II.21 - Schemi di riferimento per la definizione delle interfacce elettriche:

- (a) *interfaccia sbilanciata (Racc. V.28);*
- (b) *interfaccia con generatore bilanciato e ricevitore differenziale (Racc. V.10);*
- (c) *interfaccia con generatore sbilanciato e ricevitore differenziale (Racc. V.11).*

I livelli di tensione emessi dal generatore per ogni stato logico e quelli corrispondenti riconoscibili dal ricevitore individuano il valore massimo di attenuazione che può essere tollerato nel relativo circuito di interfaccia.

Per chiarire ulteriormente il ruolo del generatore in un circuito di interfaccia, facciamo riferimento al caso di un'apparecchiatura realizzata a circuiti integrati in tecnologia TTL (Transistor - Transistor Logic). In questo caso, per limiti connessi a tale tecnologia, i livelli che rappresentano gli stati logici "0" e "1" sono dati rispettivamente da valori compresi tra 0,2 e 0,8 V e tra

2 e 5 V. Ovviamente, valori di tensione compresi tra i livelli che identificano uno "0" e un "1" inducono una indeterminazione nel riconoscimento dello stato logico. In particolare, supponendo di dover trasmettere un "1" e nell'ipotesi che il segnale emesso sia di poco superiore al livello minimo che identifica questo stato logico, sarebbe sufficiente anche una piccola attenuazione sul circuito di interfaccia per rendere equivoco il segnale ricevuto.

Come si può osservare negli schemi di Fig. II.21, i circuiti di interfaccia possono assumere due configurazioni diverse: la configurazione *sbilanciata* e quella *bilanciata*. Poiché uno stesso schema è adottato per tutti i circuiti di un interfaccia, gli stessi attributi sono usati per caratterizzare quest'ultima.

In entrambe le configurazioni ogni circuito è costituito da due fili. Nella configurazione sbilanciata uno dei due fili è connesso alla massa di entrambe le apparecchiature accoppiate ed è comune a tutti i circuiti dell'interfaccia. Il valore del potenziale di questo filo comune costituisce, per il ricevitore, il riferimento necessario alla rivelazione dei segnali presenti sull'altro filo di ogni circuito.

Nella configurazione bilanciata le coppie di fili di ogni circuito di interfaccia sono distinte e isolate rispetto alla massa del generatore e del ricevitore. In assenza di segnale i due fili sono a potenziale uguale, mentre la presenza di segnale viene rivelata dal ricevitore misurando la differenza di potenziale tra di essi.

La configurazione bilanciata è decisamente più onerosa di quella sbilanciata da un punto di vista realizzativo, ma consente il conseguimento di migliori prestazioni. Infatti, la differenza sostanziale nel comportamento dei due tipi di configurazioni consiste nel diverso grado di protezione che esse offrono nei confronti dei disturbi di *diafonia*, e cioè dei disturbi indotti, per effetto di accoppiamenti elettromagnetici, su di un circuito dagli altri circuiti dell'interfaccia.

In particolare, la configurazione bilanciata è più resistente ai disturbi di diafonia rispetto a quella sbilanciata. Ciò può essere spiegato dal fatto che lo stesso disturbo si induce su entrambi i fili del circuito di interfaccia e che, quindi, la differenza di potenziale tra di essi è, entro certi limiti, indipendente dalla presenza di tale disturbo.

La resistenza ai disturbi di diafonia e la massima attenuazione tollerabile nei circuiti di interfaccia determinano le prestazioni dell'intera interfaccia ai fini della sua utilizzabilità in condizioni operative diverse. In particolare i parametri più significativi per definire tale utilizzabilità sono il *massimo ritmo binario*

ammissibile e la *massima distanza* a cui DTE e DCE possono essere collocati.

Nello schema di Fig. II.21a l'interfaccia è di tipo sbilanciato. Poiché questa è più sensibile ai disturbi di diafonia, essa consente un ritmo binario non superiore a 20 kbit/s ed una distanza massima di 15 m. Tale configurazione è definita nella Racc. V.28.

Nello schema di Fig. II.21b l'interfaccia è di tipo bilanciato. Con tale configurazione la diafonia risulta inferiore al caso sbilanciato e quindi i valori di ritmo binario e di distanza massima consentiti salgono rispettivamente a 10 Mbit/s e a 100 m. Tale configurazione è definita nella Racc. V.11.

Il terzo schema (Fig. II.21c) è utilizzato per l'accoppiamento di apparecchiature operanti ognuna secondo una delle due precedenti alternative. In questo caso il generatore è di tipo sbilanciato mentre il ricevitore è bilanciato. Le prestazioni di tale configurazione sono intermedie tra quelle relative ai due casi precedenti; in particolare il ritmo binario ammesso è di 300 kbit/s, mentre la distanza massima consentita è dell'ordine delle decine di metri. Tale configurazione è definita nella Racc. V.10.

In ognuno degli schemi di riferimento di Fig. II.21 sono indicate particolari sezioni in cui sono definite le grandezze elettriche tipiche dell'interfaccia. A titolo di esempio in Tab. II.4 sono riportati i valori di tensione, di corrente e di impedenza tipici delle interfacce fisiche descritte nelle Racc. V.10, V.11 e V.28. Questi valori sono qui riportati al solo scopo di dare un'idea degli ordini di grandezza per i parametri elettrici in gioco, mentre la definizione operativa di tali parametri esula dagli scopi di questa trattazione.

Grandezza	Racc. V.10	Racc. V.11	Racc. V.28
Impedenza d'uscita	< 50 $\Omega$	< 100 $\Omega$	< 300 $\Omega$
Tensione a circuito aperto	4 - 6 V	< 6 V	< 25 V
Tensione su carico di prova	> 0.9 V	> 2 V	5 - 15 V
Corrente di corto circuito	< 150 mA	< 150 mA	-
Corrente di power-off	< 100 $\mu$ A	< 100 $\mu$ A	-

Tab. II.4 - Specifiche elettriche del generatore definite  
nelle Racc. V.10, V.11 e V.28

#### II.4.3 Caratteristiche funzionali

Come si è accennato in precedenza, ad ogni circuito di interfaccia corrispondono una o più funzioni che lo stesso circuito deve espletare nel corso della connessione fisica. Le caratteristiche funzionali di una connessione fisica riguardano appunto l'assegnazione delle funzioni ai singoli circuiti di un interfaccia. Nel seguito, per rendere più semplice ed efficace la descrizione delle caratteristiche funzionali e procedurali, in luogo della dizione circuito d'interfaccia, useremo la denominazione *filo d'interfaccia*.

In generale, le funzioni che i fili di interfaccia sono chiamati a svolgere possono essere classificate in quattro categorie: a) *trasferimento dei dati*; b) *controllo*, c) *temporizzazione*; d) *massa*.

Su i fili di interfaccia dedicati al trasferimento dei dati viaggiano i segnali elettrici che trasportano le informazioni scambiate tra DTE e DCE. I fili di interfaccia dedicati alle funzioni di controllo sono utilizzati per lo scambio di informazioni inerenti l'attivazione, la supervisione e la disattivazione della connessione fisica. I fili di interfaccia preposti alla temporizzazione trasportano i sincrosegnali che permettono agli apparati riceventi di rivelare correttamente i segnali presenti sui fili dedicati al trasferimento dei dati. I fili di massa hanno infine lo scopo di fissare il livello di riferimento per la rivelazione delle tensioni sugli altri conduttori.

Il criterio seguito in passato nella definizione delle caratteristiche funzionali di una connessione fisica consisteva nell'assegnare una singola funzione ad ogni filo dell'interfaccia. Seguendo tale criterio si è arrivati a definire un'interfaccia DTE/DCE composta da 41 fili. Tale interfaccia è definita nella Racc. V.24 e si applica alla trasmissione dei dati, di tipo sia sincrono che asincrono, su linee telefoniche commutate o dedicate a due o quattro fili.

È il caso di sottolineare che tale raccomandazione definisce un'interfaccia in modo completamente generale, mentre nelle applicazioni pratiche solo un sotto-insieme di tali fili di interfaccia è effettivamente impiegato.

In Fig. II.22 è riportato l'insieme dei fili di interfaccia della Racc. V.24 che è impiegato nelle fasi di attivazione e di disattivazione della connessione fisica. Nella stessa figura è indicata l'associazione tra ogni filo e il corrispondente piedino del connettore ISO 2110.

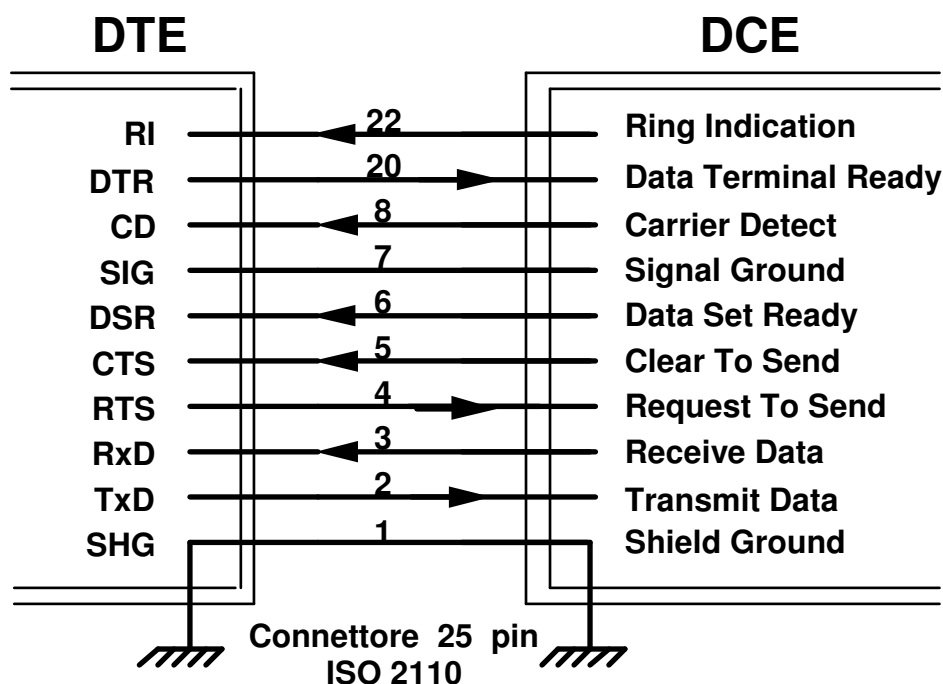


Fig. II.22 - Definizione dei fili dell'interfaccia V.24 e loro corrispondenza con i piedini del connettore ISO 2110.

In accordo alla classificazione dei fili d'interfaccia sopra introdotta, i fili riportati in Fig. II.22 possono essere così classificati: a) i *fili di trasferimento dei dati* (Transmit Data, Receive Data); b) i *fili di controllo* (Ring Indication, Data Terminal Ready, Carrier Detect, Data Set Ready, Clear To Send, Request To Send); c) i *fili di massa* (Signal Ground, Shield Ground). Accanto a questi vanno considerati anche i fili di temporizzazione. Comunque, poiché tali fili non intervengono direttamente nelle procedure di gestione dell'interfaccia non sono riportati in figura.

Le funzioni espletate da ciascuno dei fili di interfaccia ora introdotti sono:

- *Transmit Data* (TxD) (dal DTE al DCE): è utilizzato dal DTE per trasmettere i dati verso il DCE;
- *Receive Data* (RxD) (dal DCE al DTE): è utilizzato dal DCE per trasmettere i dati verso il DTE;
- *Ring Indication* (RI) (dal DCE al DTE): indica al DTE una richiesta entrante di attivazione di connessione (chiamata entrante);
- *Data Terminal Ready* (DTR) (dal DTE al DCE): indica la disponibilità del DTE ad accettare una eventuale chiamata entrante;
- *Carrier Detect* (CD) (dal DCE al DTE): indica al DTE, che ha effettuato una richiesta di attivazione di connessione (chiamata uscente), l'avvenuta connessione con il DTE chiamato.
- *Data Set Ready* (DSR) (dal DCE al DTE): indica al DTE chiamante che il DCE è pronto per l'inizio della fase di trasferimento dei dati;
- *Request To Send* (RTS) (dal DTE al DCE): è utilizzato dal DTE per richiedere

- l'autorizzazione a trasferire i propri dati;
- *Clear To Send (CTS)* (dal DCE al DTE): è utilizzato dal DCE per abilitare il DTE al trasferimento dei dati, in risposta ad una sua richiesta presentata tramite il filo di interfaccia RTS.
  - *Signal Ground (SG)*: rappresenta il livello di riferimento di tensione per i segnali scambiati su tutti i fili dell'interfaccia;
  - *Shield Ground (SHG)*: è utilizzato per portare allo stesso potenziale di riferimento entrambe le apparecchiature ai capi dell'interfaccia.

In § II.4.4 verrà mostrato come tali fili sono impiegati nel corso delle fasi di attivazione e di disattivazione della PH-conessione.

Rispetto al criterio applicato nella definizione dell'interfaccia V.24, l'evoluzione della tecnologia ha in seguito permesso di creare fili di interfaccia *multifunzionali*, in grado cioè di espletare una molteplicità di funzioni. L'applicazione di tale criterio ha permesso di definire interfacce semplificate, composte da un numero assai minore di fili.

A tale riguardo un esempio in questo senso è offerto dall'interfaccia normalizzata nella Racc. X.24, che definisce le funzioni dei fili dell'interfaccia DTE/DCE nelle reti pubbliche per dati. Essa comprende i 9 fili indicati in Fig. II.23.

Le funzioni svolte da ogni filo sono:

- *Ground (G)*: nel caso di configurazione elettricamente bilanciata, interconnette tra loro i punti a potenziale nullo del generatore e del ricevitore e funziona da ritorno comune per tutti i fili dell'interfaccia; nel caso di configurazione elettricamente sbilanciata, sono necessari due diversi ritorni per ognuna delle due direzioni di funzionamento ( $G_a$  nella direzione verso il DTE,  $G_b$  nella direzione verso il DCE);
- Transmit (T)*: è utilizzato per la trasmissione dal DTE al DCE; su questo filo viaggiano sia segnali di controllo della connessione, sia informazioni di utente;

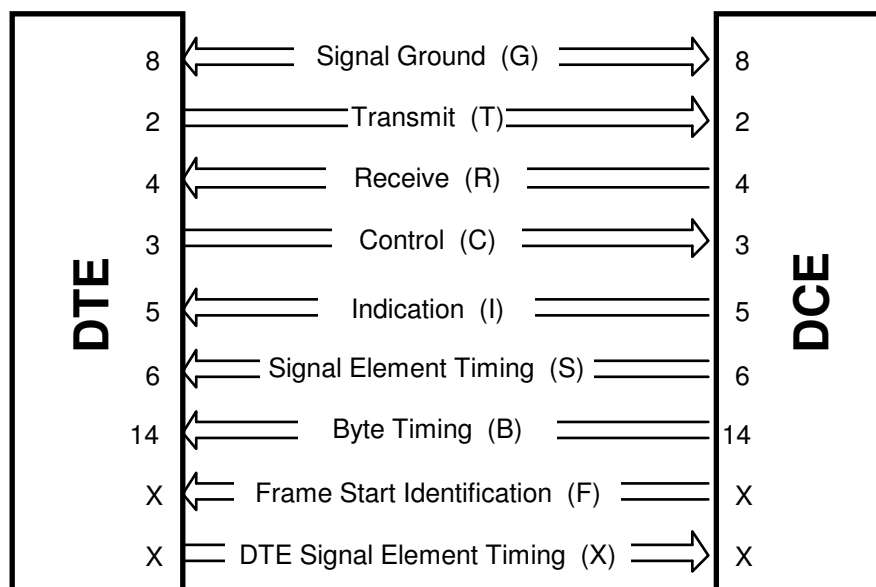


Fig. II.23 - Configurazione dell'interfaccia X.24.

- *Receive (R)*: è utilizzato dal DTE per la ricezione dei segnali trasmessi dal DCE; come per il filo T, anche in questo caso sullo stesso filo viaggiano sia segnali di controllo della connessione sia informazioni di utente;
- *Control (C)*: su questo filo il DTE invia al DCE segnali di controllo in determinate fasi della connessione; il filo può assumere solo due stati logici, chiamati ON e OFF; durante la fase di trasferimento dei dati il filo C è mantenuto stabilmente nello stato ON;
- *Indication (I)*: i segnali ricevuti su questo filo indicano al DTE lo stato del processo di controllo della chiamata; il filo può assumere solo due stati logici in analogia a quanto avviene sul filo C; se si ha trasferimento dei dati sul filo R, il filo I assume lo stato ON;
- *Signal element timing (S)*: il segnale presente su questo filo fornisce al DTE le informazioni di temporizzazione necessarie alla corretta rivelazione dei segnali sul filo R (*sincronizzazione di cifra*);
- *Byte timing (B)*: il DCE invia su questo filo le informazioni di temporizzazione concernenti la durata di un gruppo di otto cifre binarie (*sincronizzazione di byte*);
- *Frame start identification (F)*: indica al DTE l'inizio di una trama, nel caso in cui l'interfaccia DTE/DCE sia di tipo multiplato;
- *DTE transmit signal element timing (X)*: il segnale presente su questo filo fornisce al DCE le informazioni di temporizzazione necessarie alla corretta rivelazione dei segnali sul filo T; le sue funzioni sono quindi analoghe a quelle del filo S.

Nella Fig. II.23 sono indicate inoltre le modalità di associazione dei fili di interfaccia, previsti nella Racc. X.24, ai piedini del connettore ISO 4903. Occorre anche in questo caso sottolineare che la definizione dell'interfaccia X.24 è completamente generale; in casi particolari è sufficiente l'impiego di un sotto-insieme di tali fili.

#### II.4.4 *Caratteristiche procedurali*

In questa sezione sono introdotti i concetti fondamentali riguardanti le modalità di utilizzazione di una interfaccia fisica, ovvero l'insieme delle regole che devono essere seguite da DTE e DCE nella gestione dei fili di interfaccia per realizzare la corretta trasmissione delle cifre binarie informative tra gli estremi della connessione fisica. Tale insieme di regole definisce le *caratteristiche procedurali* dei protocolli di strato fisico.

Le procedure per la gestione di una interfaccia fisica sono contenute in varie norme internazionali, ognuna delle quali è relativa ad un particolare tipo di interfaccia. La Tab. II.5 riporta l'elenco delle raccomandazioni emesse dall'ITU-T a tale scopo. Spesso, tali raccomandazioni riguardano non solo funzioni di strato fisico, ma anche funzioni logicamente collocabili negli strati superiori dell'architettura OSI; inoltre, la demarcazione tra i due insiemi funzionali è spesso difficilmente individuabile. Ciò è conseguenza del fatto che la normativa relativa alle interfacce fisiche è temporalmente antecedente rispetto allo sviluppo del modello OSI e quindi l'organizzazione delle stesse spesso si discosta dall'architettura definita in questo modello.

Al riguardo, come vedremo più in dettaglio nel seguito, le procedure di attivazione e di disattivazione definite nel protocollo X.21 contengono elementi funzionali (definizione del tipo di chiamata, servizi addizionali, aspetti di tassazione, ecc.) che più correttamente andrebbero considerati appartenenti allo strato di rete.

<b>Racc.</b>	<b>Procedure per la gestione di</b>
V.24	Circuiti d'interconnessione tra DTE e DCE.
V.25	Apparecchiature a chiamata automatica
X.20	Un'interfaccia asincrona per reti pubbliche per dati
X.20 bis	Un interfaccia asincrona per reti pubbliche per dati, nel caso di utilizzazione dei modem asincroni della serie V
X.21	Un interfaccia sincrona per reti pubbliche per dati
X.21 bis	Un interfaccia sincrona per reti pubbliche per dati, nel caso di utilizzazione dei modem sincroni della serie V
X.22	Un interfaccia sincrona per reti pubbliche per dati nel caso di utilizzazione a divisione di tempo dei circuiti

Tab. II.5 - Raccomandazioni ITU-T relative alla definizione delle caratteristiche procedurali dello strato fisico.

Nel seguito di questo paragrafo saranno trattati in dettaglio i due esempi più significativi di procedure di gestione dell'interfaccia DTE/DCE, più esattamente quelle definite nelle Racc. V.24 ed X.21.

Mentre la Racc. V.24, ha il campo di applicazione già precisato in § II.4.3, la Racc. X.21 definisce, invece, le procedure di controllo dei fili di interfaccia, definiti nella Racc. X.24 per la gestione della connessione fisica tra DTE e DCE in reti pubbliche per dati.

La Racc. V.24 definisce sia la configurazione dei fili di interfaccia tra DTE e DCE, sia le modalità di gestione di questi ultimi nel caso di connessione attraverso la rete telefonica pubblica. Come già accennato, la configurazione dell'interfaccia V.24 comprende un elevato numero di fili. Solo un ristretto sotto-insieme di essi è però utilizzato in condizioni normali di funzionamento. Per rendere più semplice e significativa la descrizione delle procedure di gestione dell'interfaccia V.24 esamineremo nel seguito, a titolo di esempio,

quelle necessarie all'attivazione e alla disattivazione di una connessione fisica tra DTE e DCE. Ciò permette di considerare solo il sotto-insieme di fili indicato in Fig. II.22 a cui sarà fatto esclusivo riferimento nella descrizione successiva.

In Fig. II.24 sono riportate, sotto forma di un diagramma spazio-temporale, le procedure relative alle fasi di attivazione, di trasferimento dei dati e di disattivazione di una connessione tra due DTE remoti interconnessi tra loro attraverso una rete telefonica.

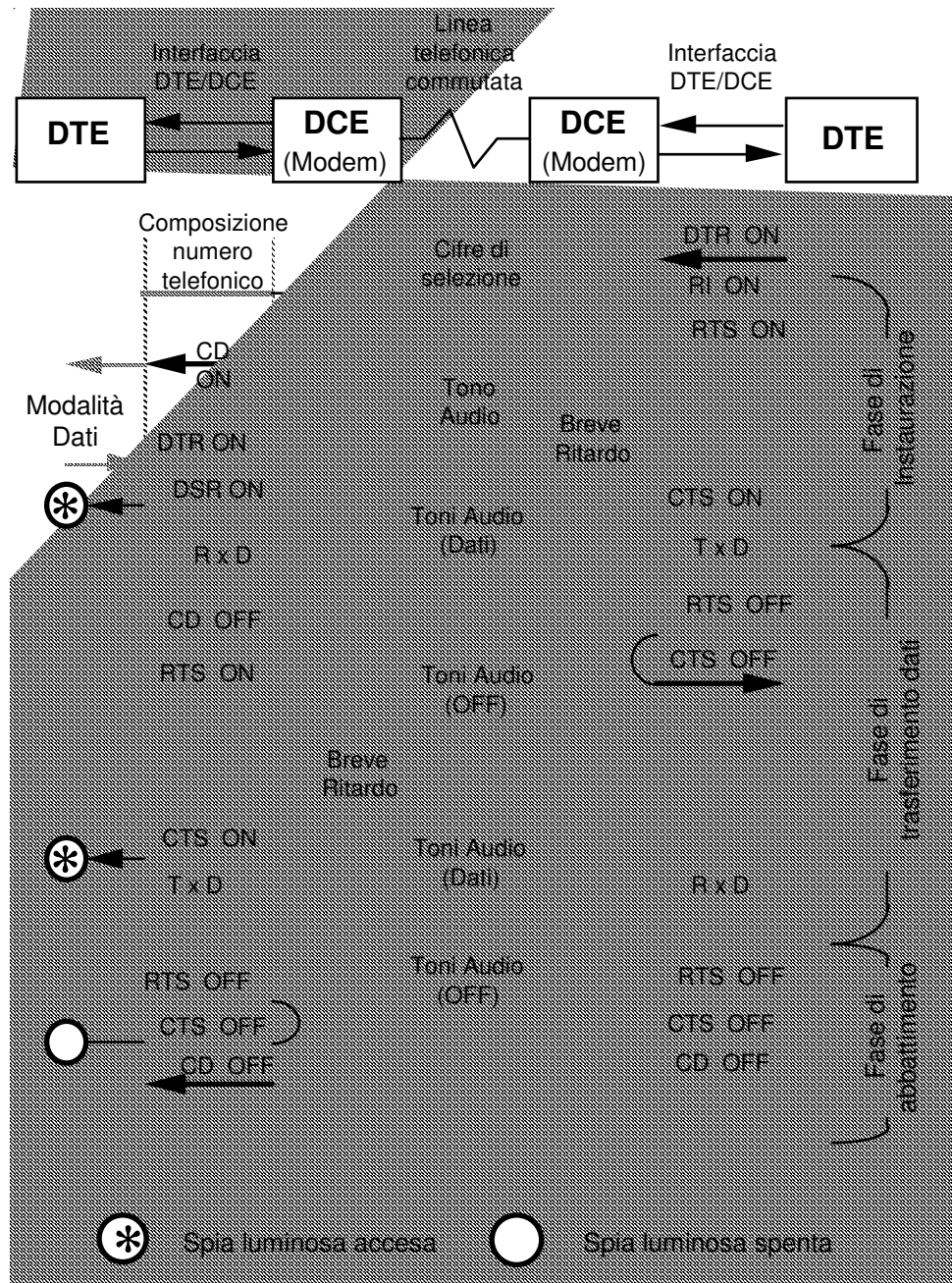


Fig. II.24 - Esempio di procedura relativa alla Racc. V.24.

Nella parte alta della figura è raffigurato lo schema di riferimento della

connessione. In particolare, nell'esempio si assume che il DTE chiamante sia un terminale comandato da un operatore, mentre il DTE chiamato è un elaboratore a risposta automatica. La linea di comunicazione è infine supposta di tipo semi-duplice.

Al di sotto dello schema di riferimento sono indicate, invece, nella loro sequenza temporale, le azioni che le varie apparecchiature interessate alla connessione intraprendono nel corso della procedura. Distinguendo le procedure relative alle fasi di attivazione, di trasferimento dei dati e di disattivazione, queste sono descritte in dettaglio qui di seguito

#### *PROCEDURE RELATIVE ALLA RACC. V. 24*

La *fase di attivazione* della connessione inizia quando l'operatore compone il numero telefonico del DTE chiamato. A seguito di tale azione il DTE ed il DCE chiamanti si pongono in uno stato di attesa della risposta da parte dell'elaboratore remoto. Il DCE chiamato, nel momento in cui riceve la richiesta di connessione, esamina per prima cosa lo stato del DTE ad esso connesso. Se il DTE è nello stato di "pronto a ricevere" (filo DTR nello stato ON), il DCE avverte, tramite il filo RI, il DTE chiamato che è presente una richiesta di connessione entrante. L'accettazione della richiesta di connessione da parte del DTE è segnalata ponendo il filo RTS nello stato ON.

Quest'ultima azione provoca due effetti: il primo consiste nell'invio verso il DCE chiamante di un tono audio indicante che la chiamata è stata accettata; il secondo, invece, riguarda lo stato del DCE chiamato che si pone in attesa di ricevere dati dall'elaboratore. Tale disponibilità a ricevere i dati è segnalata al proprio DTE ponendo, dopo un breve ritardo, il filo CTS nello stato ON.

All'altro estremo della connessione, non appena il DCE chiamante riceve il tono audio indicante l'accettazione della chiamata avverte il proprio DTE ponendo nello stato ON il filo CD. Poiché abbiamo supposto che il DTE chiamante sia un terminale non automatico, questo evento sarà avvertito dall'operatore quando il tono audio sarà percepito attraverso il ricevitore del microtelefono. A seguito di tale evento, è compito dell'operatore commutare l'interruttore presente sul modem dalla posizione "audio" alla posizione "dati". Tale operazione corrisponde a porre il filo DTR nello stato ON. La fase di attivazione della connessione è chiusa dal DCE chiamante ponendo nello stato ON il filo DSR; ciò è segnalato dall'accensione di un'opportuna spia di segnalazione sul modem.

Ci riferiamo ora alla *fase di trasferimento dei dati*. Poiché nel caso che stiamo commentando è ipotizzata una PH-connessione di tipo semi-duplice, durante la fase di trasferimento dei dati gli intervalli di emissione dei DTE impegnati nella connessione si alterneranno. La gestione di tali intervalli di emissione è effettuata mediante i fili RTS e CTS.

Nella Fig. II.24 il primo intervallo di emissione è assegnato al DTE chiamato. Non appena l'elaboratore riceve l'abilitazione a trasmettere (filo CTS nello stato ON), inizia, sul filo TxD, l'invio del segnale di dati. Alla fine della emissione, l'elaboratore e il proprio DCE riporteranno, in successione, i fili RTS e CTS nello stato OFF; ciò indicherà al DCE remoto che la trasmissione del messaggio è conclusa e si attende un'eventuale risposta.

Il DTE chiamante riceverà i dati emessi dall'elaboratore. Il termine della trasmissione sarà indicato dalla transizione da ON ad OFF dello stato del filo CD. Se il DTE chiamante ha

dati da trasmettere, richiederà a sua volta il diritto alla trasmissione ponendo nello stato ON il filo RTS. Non appena il DCE abiliterà la trasmissione (filo CTS nello stato ON), il DTE potrà iniziare l'invio dei dati sul filo TxD.

Procedure analoghe alla precedente saranno svolte ai due estremi della connessione ogni qual volta dovrà essere effettuato il cambio del turno di emissione.

La *fase di disattivazione* della connessione inizia quando entrambe le parti hanno terminato lo scambio informativo. La procedura seguita dai due DTE consiste semplicemente nel porre entrambi i fili RTS nello stato OFF. A tale evento seguirà, da parte dei DCE, l'operazione di porre i fili CD e DSR anch'essi nello stato OFF.

Come abbiamo sottolineato più volte, quanto è considerato in Fig. II.24 si riferisce al caso di PH-connessione di tipo semi-duplice. In tale tipo di connessione il tempo necessario per effettuare le operazioni di cambio del turno di trasmissione è detto *tempo di inversione* (turn-around time). Tale tempo, in molti casi, rappresenta una percentuale non trascurabile della durata totale della fase di trasferimento dei dati e quindi può essere un elemento notevolmente peggiorativo della qualità e dell'efficienza della connessione fisica.

Al riguardo, la perdita di efficienza nell'utilizzazione della connessione fisica è tanto più elevata quanto più è limitato l'ammontare dei dati che debbono essere trasferiti e tanto più è grande il numero di volte in cui deve essere effettuato il cambio di turno. Trascurando la durata delle fasi di attivazione e di disattivazione della connessione stessa ed il tempo di propagazione dei segnali sul mezzo trasmissivo, l'efficienza  $E$  della connessione, definita come la quota parte di tempo in cui la connessione è effettivamente utilizzata per il trasferimento di dati informativi, dipende dal numero di inversioni di turno  $M$  secondo la relazione:

$$E = \frac{T_t}{T_t + MT_i}$$

dove  $T_t$  è il tempo totale di trasmissione e  $T_i$  è il tempo di inversione. Supponendo che debbano essere scambiati dati per un totale di 300 bit su una connessione di tipo semi-duplice operante al ritmo binario di 1200 bit/s e che il tempo di inversione sia uguale a 10 ms, l'efficienza della connessione è uguale a 0.96 se  $M=1$ . Se il numero di inversioni sale a 5, l'efficienza scende a 0.83; se occorrono invece 10 inversioni, il valore di  $E$  risulta uguale a 0.71.

Per tali ragioni è preferibile usare PH-connessioni di tipo duplice, in cui i DTE possono emettere informazioni anche contemporaneamente. Se è utilizzato

il modo di funzionamento duplice, dal punto di vista della procedura di gestione dell'interfaccia fisica, le differenze rispetto al caso considerato in Fig. II.24 consistono esclusivamente nel fatto che i fili RTS, CTS e CD di entrambe le interfacce DTE/DCE sono mantenuti permanentemente nello stato ON.

Nel caso di funzionamento dell'interfaccia V.24, abbiamo supposto che terminale e elaboratore fossero posti in posizioni geograficamente distanti l'una dall'altra e comunque tali da richiedere l'uso di una connessione fisica attraverso una rete di comunicazione.

La stessa interfaccia V.24 può essere però utilizzata anche per connessioni a brevissima distanza, dell'ordine al massimo di qualche decina di metri, quali ad esempio si incontrano nel caso di connessione tra terminali e elaboratore nell'ambito dello stesso centro di calcolo.

In questo caso, data la distanza limitata, l'uso dei modem non è necessario e la connessione tra le due apparecchiature può avvenire direttamente. Poiché però l'interfaccia V.24 è stata concepita per il colloquio tra DTE e modem, occorre risolvere il problema di stabilire quale delle due parti (terminale o elaboratore) deve assumere il compito di emulare la presenza dei modem stessi agli estremi della connessione.

La soluzione che si adotta normalmente consiste nel modificare opportunamente l'interconnessione dei fili di interfaccia, in modo che ognuna delle due apparecchiature si comporti, nei confronti dell'altra, esattamente come se fosse presente una coppia di modem.

Il modo in cui sono modificati i fili dell'interfaccia è quello indicato in Fig. II.25.

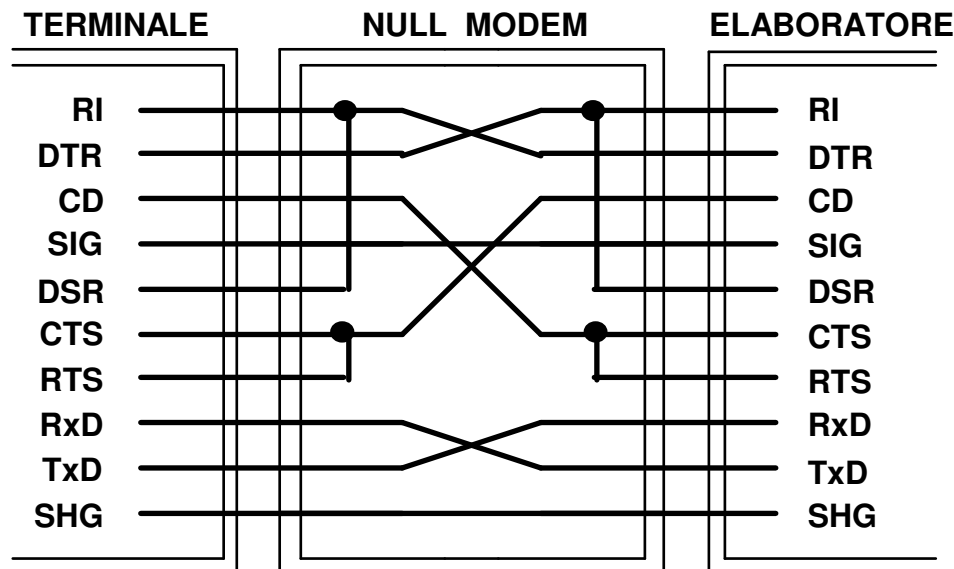


Fig. II.25 - Connessione terminale-elaboratore tramite "null modem".

Tale nuova configurazione viene detta *null modem* e rende possibile una connessione fisica di tipo duplice tra due dispositivi asincroni.

Come è immediato notare dalla Fig. II.25, le caratteristiche della configurazione null modem sono:

- a) i fili TxD e RxD sono invertiti tra loro ai due capi della connessione in modo che i dati trasmessi da una apparecchiatura sul filo TxD siano ricevuti dall'altra sul filo RxD e viceversa;
- b) i fili RI e DTR sono invertiti tra loro ai due capi della connessione; ciò equivale a simulare, nel caso in cui entrambe le apparecchiature siano attive, la presenza di una chiamata entrante;
- c) i fili DTR e DSR sono interconnessi tra loro ad ogni estremo della connessione; in tal modo viene segnalata la condizione di attività della connessione;
- d) i fili CTS e RTS sono interconnessi tra loro ad ogni estremo della connessione; ciò equivale a simulare l'abilitazione a trasmettere concessa dal modem al DTE;
- e) i fili RTS e CD sono invertiti tra loro ad ogni estremo della connessione; ciò segnala alle due parti la presenza della portante sulla connessione;
- f) i fili di massa sono connessi direttamente; il potenziale di riferimento è quindi lo stesso per entrambe le apparecchiature.

La Racc. X.21 definisce le caratteristiche fisiche e le procedure di controllo di chiamata, relative all'interfaccia tra DTE e DCE nel caso di

trasmissione sincrona in reti per dati a commutazione di circuito.

Dal punto di vista delle caratteristiche fisiche, la Racc. X.21 prevede l'utilizzazione dell'interfaccia definita nella Racc. X.24, di cui abbiamo parlato in § II.4.3, mentre il connettore fisico è quello definito nello standard ISO 4903 e riportato in Fig. II.20.

Dal punto di vista delle caratteristiche procedurali, la Racc. X.21 descrive in dettaglio le procedure di controllo e di rilascio della connessione fisica tra DTE e DCE. Il controllo dello scambio dati durante la fase informativa è invece compito dei protocolli di strato superiore.

Illustreremo le procedure relative alla gestione dell'interfaccia descrivendo l'evoluzione del suo stato. Questo è determinato dalla particolare configurazione che, istante per istante, assumono i fili T, R, C ed I. L'evoluzione è sinteticamente descritta da opportuni diagrammi di transizione di stato. In Fig. II.26 sono illustrati lo schema di riferimento della connessione tra DTE e DCE e le convenzioni adottate per la rappresentazione degli stati dell'interfaccia e delle possibili transizioni tra di essi.

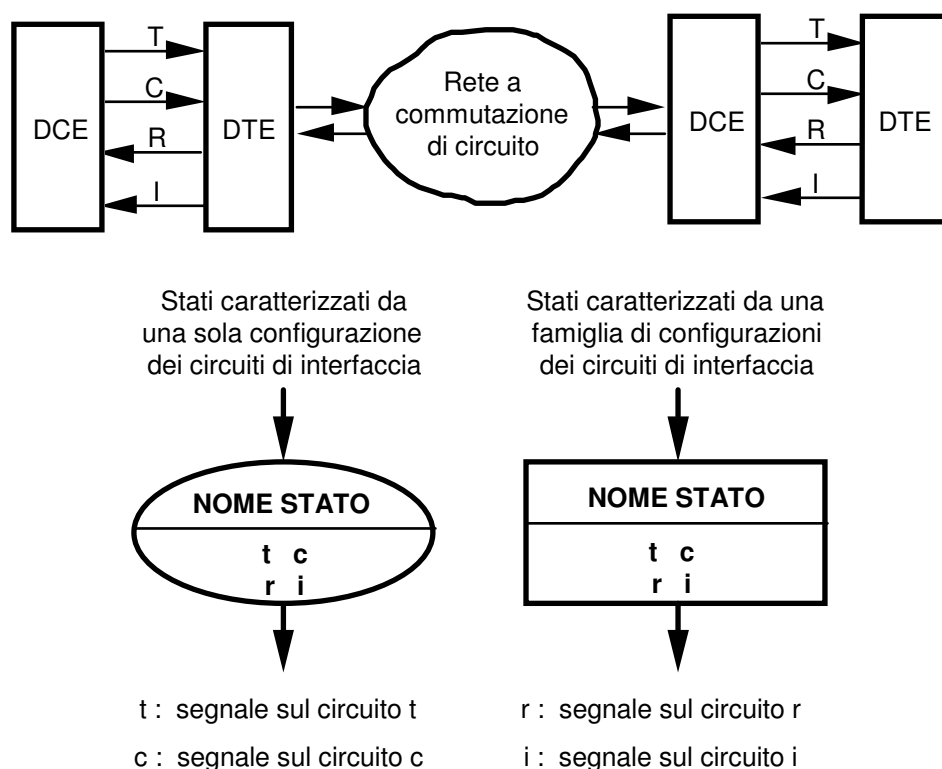


Fig. II.26 - Convenzioni per la rappresentazione dei diagrammi di transizione di stato relativi alla Racc. X.21.

In termini generali l'evoluzione dello stato dell'interfaccia può essere suddivisa in due fasi principali: la *fase di riposo* e la *fase operativa*. Durante la

fase di riposo nessun colloquio è in corso tra DTE e DCE, mentre nella fase operativa DTE e DCE si scambiano le informazioni per la gestione della connessione fisica. La fase operativa può essere ulteriormente suddivisa in tre sottofasi: 1) la *fase di controllo della chiamata*; 2) la *fase di trasferimento dei dati*; 3) la *fase di disattivazione della connessione*.

Nel seguito descriveremo dunque le fasi di una connessione attraverso il commento dei relativi diagrammi di transizione di stato. Saranno prese inoltre in considerazione esclusivamente le condizioni normali di funzionamento; non si considereranno cioè condizioni anomale dovute a malfunzionamenti delle apparecchiature od ad altre cause esterne.

Durante la fase di riposo, il DTE e DCE segnalano la loro disponibilità o indisponibilità ad entrare in una eventuale fase operativa. Il diagramma relativo alla fase di riposo è quello descritto in Fig. II.27.

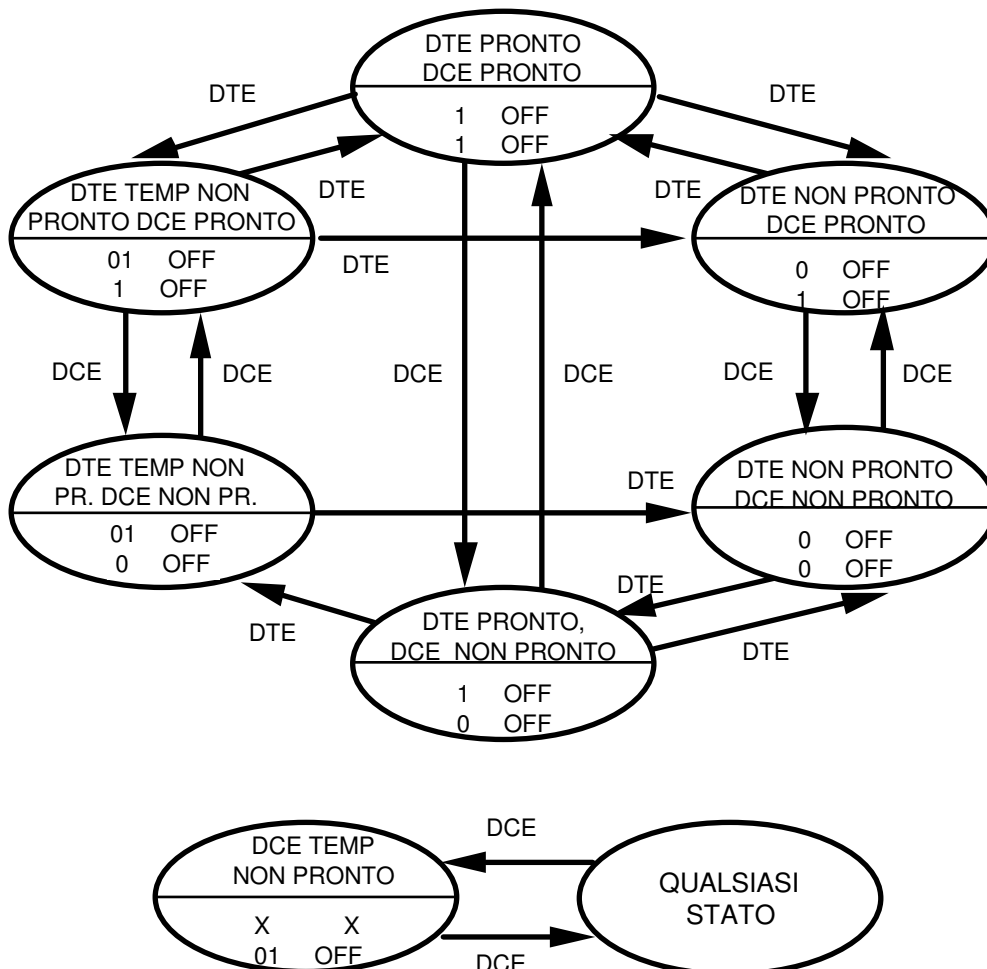


Fig. II.27 - Diagramma di transizione di stato della Racc. X.21 relativo alla fase di riposo.

Nella fase di controllo di chiamata si eseguono le funzioni relative all'instaurazione di una connessione tra DTE e DCE. Nel diagramma di

transizione di stato, riportato in Fig. II.28, si distinguono due diversi casi: a) *chiamata uscente dal DTE*, segnalata dal DTE ponendo a "0" il filo T e nello stato ON il filo C (stato di *richiesta di connessione*); b) *chiamata entrante al DTE*, segnalata dal DCE con la trasmissione consecutiva sul filo R di caratteri "BEL" dell'alfabeto internazionale n° 5 (stato di *chiamata entrante*).

Una condizione particolare dell'interfaccia è espressa dallo stato di *collisione di chiamata*. Una collisione di chiamata può avvenire in due casi: 1) quando il DTE riceve una segnalazione di chiamata entrante in risposta ad una richiesta di chiamata da lui emessa; 2) quando il DCE riceve una richiesta di connessione in risposta ad una segnalazione di chiamata entrante. Nello stato di collisione di chiamata è sempre privilegiata la chiamata uscente.

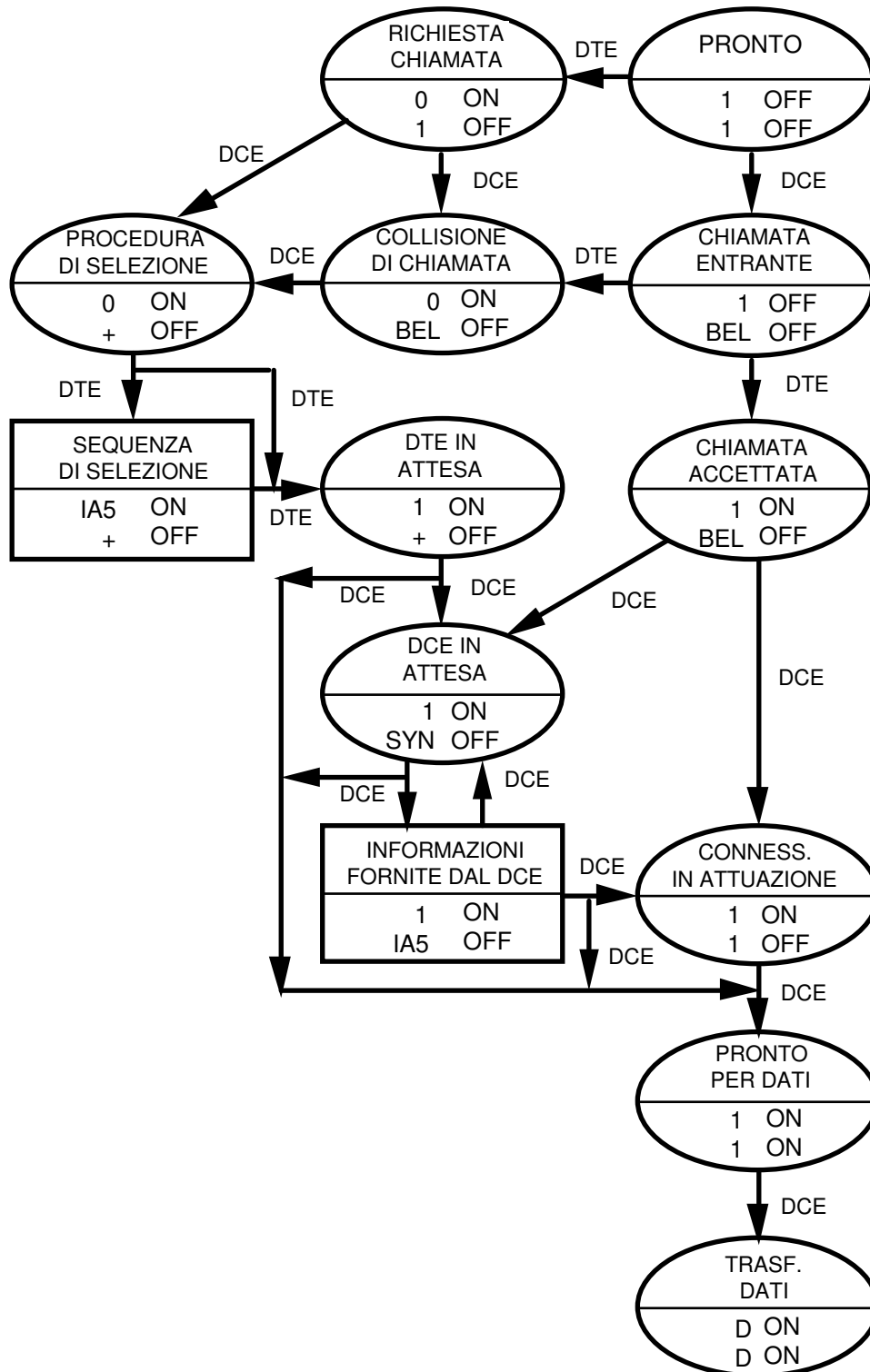


Fig. II.28 - Diagramma di transizione di stato della Racc. X.21 relativo alla fase di controllo di chiamata.

La parte finale della procedura di controllo di chiamata è comune ai due casi di chiamata entrante ed uscente. Essa è in generale composta da quattro stati che comunque, come si nota dal diagramma, possono, in casi particolari,

non essere attraversati durante la procedura.

La fase di trasferimento dei dati è delimitata dall'apparire della condizione di pronto per dati e termina con l'invio, da parte del DTE o del DCE, di una richiesta di disconnessione. Durante questa fase il DTE può inviare o ricevere una qualsiasi sequenza di bit.

La fase di rilascio ha lo scopo di riportare l'interfaccia nelle condizioni di riposo. Può avere inizio da un qualsiasi stato assunto dall'interfaccia, sia nella fase di trasferimento dei dati sia nella fase di controllo di chiamata. L'evoluzione dello stato durante questa fase è rappresentata in Fig. II.29. Si possono distinguere due casi: a) *disconnessione iniziata dal DTE*; b) *disconnessione iniziata dal DCE*.

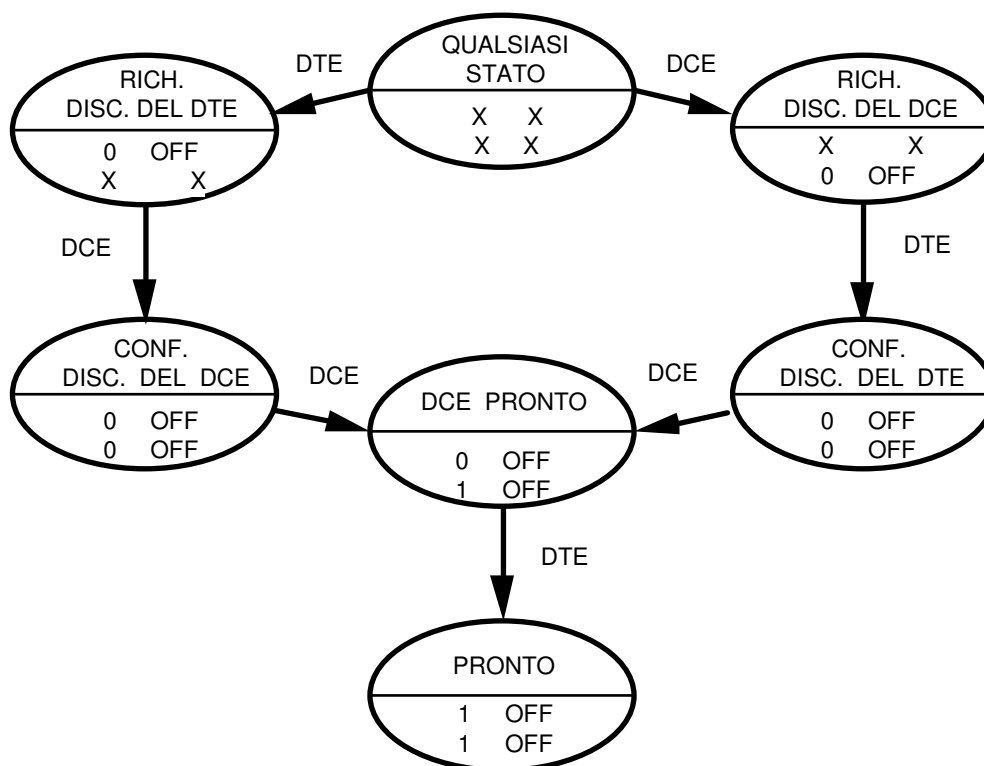


Fig. II.29 - Diagramma di transizione di stato della Racc. X.21 relativo alla fase di disconnessione.

### PROCEDURE RELATIVE ALLA RACC. X.21

Con riferimento alla Fig. II.27, lo stato di pronto del DTE è segnalato al DCE mediante la configurazione (1,OFF) i fili T e C, viceversa Il DCE segnala la sua condizione di pronto ponendo nella stessa configurazione i fili R ed I. Gli stati di indisponibilità di una delle apparecchiature ad entrare nella fase operativa è invece segnalato ponendo a "0" il filo T od R nel caso del DTE o del DCE, rispettivamente.

Il diagramma di stato dell'interfaccia prevede inoltre che sia il DTE che il DCE possano

trovarsi temporaneamente in uno stato di non pronto; questa condizione è segnalata all'altra entità mediante l'alternarsi degli stati "0" ed "1" sui fili T ed R.

Relativamente alla Fig. II.28, nel caso di chiamata uscente, il DCE, in risposta alla richiesta presentata dal DTE, trasmette il carattere "+" dell'alfabeto IA5 sul filo R per sollecitare il DTE ad emettere le informazioni di selezione (stato di *procedura di selezione*). Le informazioni di selezione (indirizzi del DTE chiamato) vengono trasmesse dal DTE sul filo T (stato di *sequenza di selezione*). Al termine dell'invio di tali informazioni, il DTE si pone in uno stato di attesa della conferma dell'avvenuta connessione (stato di *DTE in attesa*).

Nel caso di chiamata entrante, il DTE ha il compito di accettare o meno tale richiesta di connessione proveniente dal DCE. Nel caso in cui la richiesta venga accolta (stato di *chiamata accettata*) il DTE pone ad "1" il filo T.

Sempre relativamente alla Fig. II.28, in generale, nel momento in cui il DTE ha ultimato l'invio delle informazioni di selezione e si è posto nello stato di attesa (chiamata uscente) o ha accettato la richiesta di connessione (chiamata entrante), il DCE si pone anch'esso in uno stato di attesa (stato di *DCE in attesa*) trasmettendo consecutivamente sul filo R una sequenza di caratteri "SYN". Tale stato di attesa corrisponde al tempo necessario alla ricezione da parte del DCE remoto di ulteriori informazioni riguardanti la connessione.

Al termine del periodo di attesa, il DCE fornisce al DTE ulteriori informazioni sulle caratteristiche generali della chiamata (stato di *informazioni fornite dal DCE*). Tali informazioni possono riguardare: l'identificazione del DTE chiamante e del DTE chiamato, la tariffazione, la data e l'ora della connessione, ecc.

Al termine di tale trasmissione l'interfaccia assume lo stato di *connessione in attuazione* segnalato dalla configurazione (1,OFF) dei fili R ed I. Quando la connessione è effettivamente disponibile per il trasferimento dei dati, il DCE forzerà la transizione dell'interfaccia nello stato di *pronto per dati*, segnalato dallo stato ON del filo I. Questo stato è l'ultimo della fase di controllo di chiamata e si può dare inizio alla fase di trasferimento dei dati.

Passando infine all'esame della Fig. II.29, nel caso di disconnessione iniziata dal DTE, questo pone nella configurazione (0, OFF) i fili T e C dell'interfaccia (stato di *richiesta di disconnessione del DTE*). Nel caso invece di disconnessione iniziata dal DCE, in modo analogo, il DCE pone nella configurazione (0, OFF) i fili R e I (stato di *richiesta di disconnessione del DCE*). A tali sollecitazioni verrà risposto con una conferma di disconnessione (stati di *conferma di disconnessione del DCE o DTE*) e successivamente l'interfaccia si riporterà nella condizione di riposo (stati di *DCE pronto e DTE pronto*).

Infine ci riferiamo al diagramma temporale relativo alle azioni intraprese dai *DTE* e *DCE* chiamanti e chiamati, nel corso della fase di instaurazione di una chiamata, nell'ipotesi di utilizzazione del protocollo X.21. Tale diagramma è mostrato in Fig. II.30 e può essere tracciato a partire dal diagramma di transizione di stato raffigurato in Fig. II.28.

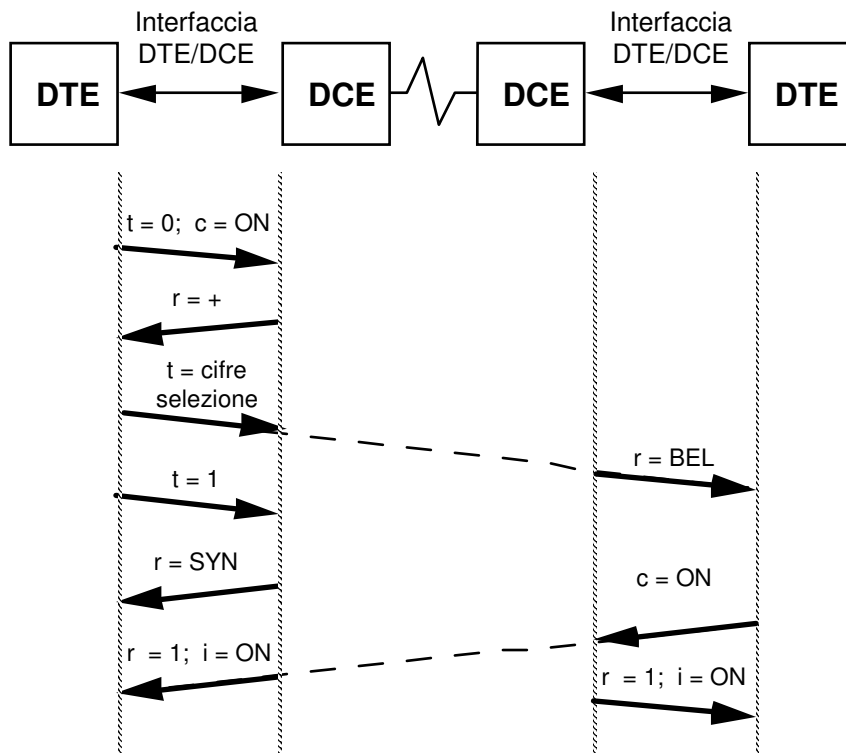


Fig. II.30 - Diagramma temporale relativo alla fase di instaurazione di una chiamata mediante il protocollo X.21.



### III STRATO DI COLLEGAMENTO

Questo capitolo è dedicato alla trattazione del secondo strato dell'architettura OSI, e cioè dello strato di collegamento. Conseguentemente, mentre nel precedente cap. II si è fatto riferimento a un contesto strettamente legato alla *fisica* del problema del trasferimento, qui si affrontano aspetti nei quali le funzionalità e i mezzi coinvolti hanno significato solo su un piano *logico*.

Vengono dapprima (par. III.1) illustrate le caratteristiche del servizio di strato di collegamento. Successivamente (par. III.2) si passa ad un'analisi delle funzioni che lo strato è chiamato a svolgere nel corso dell'evoluzione della comunicazione. Vengono poi (par. III.3) considerate le procedure attraverso le quali ogni strato può conseguire le sue finalità (*protocolli di strato*). Segue infine la descrizione di un protocollo di strato di collegamento avente significatività in contesti specifici (par. III.4, LAP-B).

Altri esempi di protocolli di strato di collegamento sono trattati nel par. VIII.4 (LAP-D nella N-ISDN) e nel par. VI.9 (strato LLC nelle LAN).

In questo insieme di argomenti, si approfondiranno soprattutto i temi con maggiore contenuto concettuale; si demanderà agli esempi il duplice compito di chiarire questioni di maggior dettaglio che derivano dall'applicazione dei concetti di base e di fornire una più agevole chiave di interpretazione a impostazioni inevitabilmente astratte.

#### III.1 Caratteristiche del servizio

Lo strato di collegamento (*Data Link - DL*) è, partendo dal basso, il secondo strato dell'architettura OSI; esso si interfaccia superiormente con lo strato di rete e inferiormente con lo strato fisico (Fig. III.1). Il servizio di strato di collegamento (*DL-servizio - DLS*) viene fornito dalle entità di questo strato (*DLS-fornitore*) alle entità dello strato di rete (*DLS-utenti*), utilizzando le risorse messe a disposizione dallo strato fisico.

La finalità dello strato di collegamento è quella di garantire il trasferimento trasparente ed affidabile delle unità di dati del DL-servizio (*Data Link Service Data Units - DLSDU*) tra le entità di rete. Lo strato di collegamento deve rendere, quindi, non visibile ai DLS-utenti il modo in cui sono utilizzate le risorse di comunicazione.

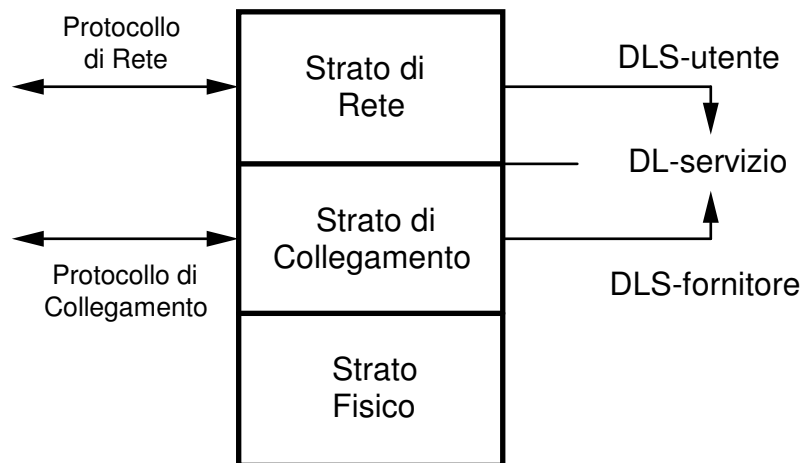


Fig. III.1 - Relazione tra lo strato di collegamento e gli strati adiacenti dell'architettura OSI.

In particolare, il DL-servizio assicura:

- *l'indipendenza dal sottostante strato fisico*; il DL-servizio rende il funzionamento del DLS-utente indipendente dalla configurazione della connessione di strato fisico o dalle sue modalità operative;
- *la trasparenza delle informazioni trasferite*; il trasferimento delle DLSDU viene assicurato indipendentemente dal loro contenuto, dal loro formato o dal loro tipo di codifica;
- *il trasferimento affidabile delle informazioni*; il DL-servizio ha lo scopo di rendere non percepibile allo strato di rete eventuali perdite, errori, duplicazioni e, se richiesto, alterazioni nella sequenza delle DLSDU da trasferire;
- *la selezione di una qualità di servizio*; il DL-servizio mette a disposizione del DLS-utente un servizio di trasferimento delle informazioni caratterizzato da una concordata *qualità del servizio*; identificata da valori di determinati parametri, quali la portata (throughput), il ritardo di transito, ecc.
- *la possibilità di indirizzamento dei dati*; nel caso in cui lo strato di collegamento supporti più di due punti d'accesso al DL-servizio (*DLSAP*), lo strato di rete ha la possibilità di identificare verso quale DLSAP un DL-servizio deve essere reso disponibile.

Una delle funzioni di base del DL-servizio è quella di rendere non percepibile allo strato superiore le modalità di funzionamento dello strato fisico. Se, ad esempio, la connessione di strato fisico opera in modalità semi-duplice (cfr par. II.2), lo strato di collegamento deve fornire i mezzi necessari alla gestione di tale connessione, in modo da rendere trasparente all'utente l'alternanza delle opportunità di trasmissione nei due versi. L'unico effetto

visibile all'utente del servizio deve essere l'aumento del tempo di transito delle DLSDU rispetto al caso di connessione duplice.

Nelle reti a lunga distanza, lo strato di collegamento opera normalmente tra due sole entità di rete (es. nodi di commutazione, multiplatori, ecc.). In tal caso, la funzione di indirizzamento non è necessaria poiché i punti di accesso al servizio di collegamento sono univocamente determinati.

Diverso è il caso di reti ad accesso multiplo, come, ad esempio, le reti in area locale (LAN). In queste reti, lo strato di collegamento offre servizio contemporaneamente a una molteplicità di entità di rete, caratterizzate da DLSAP diversi. La funzione di indirizzamento è allora necessaria per regolare lo scambio informativo tra tali entità.

Il DL-servizio può essere fornito secondo due modalità: a) *orientata alla connessione*; b) *senza connessione*.

Nella modalità orientata alla connessione, lo scambio delle DLSDU attraverso lo strato di collegamento è subordinato all'instaurazione di una connessione (*DL-connessione*) tra i DLSAP associati alle entità di rete. L'utilizzazione del DL-servizio orientato alla connessione avviene attraverso tre fasi distinte: a) la *fase di instaurazione* della DL-connessione; b) la *fase di trasferimento di dati*; c) la *fase di abbattimento* della DL-connessione.

Una DL-connessione è instaurata quando si hanno dati da trasferire e rimane attiva solo sino a quando il trasferimento dei dati non è stato completato. Il suo tempo di tenuta è quindi finito e chiaramente determinabile. Inoltre, una connessione determina un'associazione implicita tra le DLSDU scambiate durante la sua durata. Tale legame logico è sfruttato per l'esecuzione delle funzioni di controllo di flusso e di sequenza.

Nella modalità senza connessione, un DLS-utente può iniziare la emissione di una singola DLSDU verso qualsiasi altro DLS-utente senza la preventiva instaurazione di una connessione. In questo caso lo scambio informativo ha le seguenti caratteristiche:

- richiede una pre-esistente associazione tra le entità impegnate nello scambio informativo, che stabilisca a priori le caratteristiche delle DLSDU da scambiare;
- il DLS-utente deve comunicare al DLS-fornitore, insieme ai dati da trasferire, tutte le informazioni (indirizzo di destinazione, qualità di servizio richiesta, opzioni, ecc.) che consentano la loro consegna a destinazione;
- ogni DLSDU è trattata indipendentemente dalle altre, senza tener conto di un eventuale appartenenza di due o più DLSDU alla stessa transazione; ciò può

comportare la consegna a destinazione delle DLSDU in ordine diverso da quello di emissione.

La modalità orientata alla connessione è quella utilizzata nel caso di applicazioni che richiedono interazioni di lunga durata tra entità in configurazione stabile, come nel caso di trasferimento di lunghi "file" di dati. Tale modalità consente di ridurre l'informazione aggiuntiva (*overhead*) in ogni DLSDU, dato che elimina la necessità di inserire in ogni DLSDU gli interi indirizzi delle entità di rete poste in corrispondenza, sostituendoli con l'identificatore della connessione.

La modalità senza connessione è invece più adatta ad applicazioni in cui il volume dei dati da scambiare è limitato e le interazioni tra le entità di rete sono sporadiche. In tal caso la maggiore quantità di informazione aggiuntiva contenuta nelle unità di dati è compensata dalla assenza delle fasi di instaurazione e di abbattimento della connessione.

Nelle LAN, il DL-servizio è normalmente di tipo senza connessione. Tale scelta è una conseguenza della particolare struttura di tali reti (mezzo condiviso) e dell'ambiente applicativo per cui sono state concepite (elaborazione distribuita). La gestione delle connessioni avrebbe infatti molto appesantito le procedure dello strato di collegamento, rendendolo inadatto all'utilizzazione in un ambiente ad accesso multiplo.

## **III.2 Funzionalità del servizio**

In questo paragrafo saranno illustrate le funzionalità del DL-servizio distinguendo tra quelle tipiche della modalità con connessione (§ III.2.1) e quelle relative alla modalità senza connessione (§ III.2.2).

### *III.2.1 Funzionalità del DL-servizio con modalità orientata alla connessione.*

Il DL-servizio di tipo orientato alla connessione fornisce ad un DLS-utente le seguenti funzionalità:

- i mezzi per instaurare una DL-connessione con un altro DLS-utente;
- i mezzi per negoziare la voluta qualità di servizio sulla DL-connessione da instaurare;
- i mezzi per trasferire DLSDU di opportuna lunghezza (non necessariamente costante) su una DL-connessione; il trasferimento delle DLSDU è completamente trasparente, nel senso che è responsabilità del DL-servizio assicurare che il loro contenuto sia consegnato inalterato a destinazione;

- i mezzi attraverso i quali un DLS-utente può regolare il ritmo di emissione delle DLSDU da parte del DLS-utente con cui è in corrispondenza (funzione di *controllo di flusso*);
- i mezzi per riportare, in caso di malfunzionamenti, lo stato di una DL-conessione ad un valore definito e per ri-sincronizzare l'attività dei DLS-utenti (funzione di *re-inizializzazione*);
- i mezzi attraverso i quali un DLS-utente o il DLS-fornitore possono abbattere, in qualsiasi momento, una DL-conessione.

Il funzionamento di una DL-conessione e le interazioni tra i DLS-utenti possono essere rappresentati dal modello funzionale raffigurato in Fig. III.2. Ogni verso di emissione delle DLSDU è modellato da una coda che rappresenta la funzione di controllo di flusso operata, per ogni verso, dal DLS-utente ricevente. Ogni coda può contenere *elementi* di diverso tipo, ognuno dei quali modella un'interazione tra un DLS-utente e il DLS-fornitore.

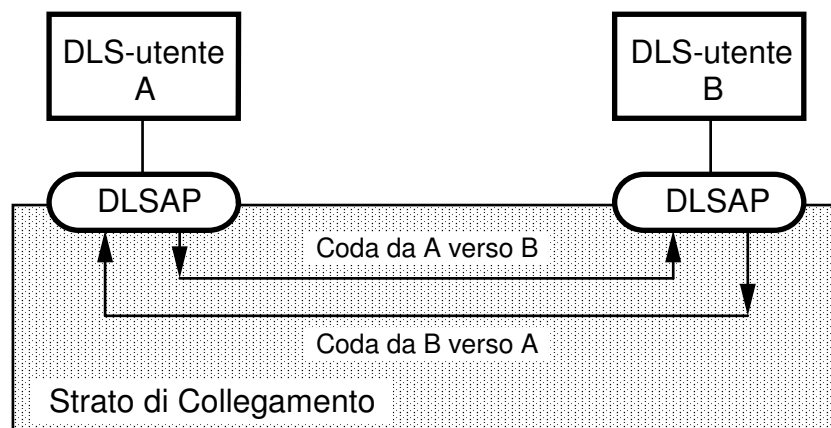


Fig. III.2 - Modello astratto di una DL-conessione.

I tipi di elementi sono:

- a) *elementi di connessione*, che rappresentano le primitive relative alla fase di instaurazione di una connessione;
- b) *elementi di dati*, che rappresentano le primitive relative al trasferimento di dati;
- c) *elementi di re-inizializzazione*, che rappresentano le primitive relative alla funzione di re-inizializzazione;
- d) *elementi di disconnessione*, che rappresentano le primitive relative alla fase di abbattimento di una connessione.

Gli elementi sono posti in una coda dal DLS-utente emittente sotto il controllo dello strato di collegamento. Gli elementi di re-inizializzazione e di disconnessione possono essere inseriti in una coda anche direttamente dal

DLS-fornitore.

La possibilità, da parte del DLS-fornitore, di inserire elementi di re-inizializzazione e di disconnessione nelle code che modellano una DL-connessione tra DLS-utenti riflette uno dei principi fondamentali dell'architettura stratificata, cioè quello di *notifica* allo strato superiore di eventuali malfunzionamenti che potrebbero comportare la non completa affidabilità di una connessione.

Se, ad esempio, il tasso di errore in linea è superiore a quello massimo consentito dalle capacità del protocollo di strato di collegamento, ovvero quest'ultimo non riesce, con i propri meccanismi di controllo d'errore, a garantire la concordata qualità di servizio, il DLS-fornitore deve interrompere la fornitura del DL-servizio attraverso l'abbattimento forzato della DL-connessione.

Una coda è considerata vuota nel momento in cui viene posto in essa un elemento di connessione. Gli elementi sono normalmente rimossi nello stesso ordine in cui sono immessi in coda. La rimozione avviene sotto il controllo del DLS-utente ricevente.

Nella fase di instaurazione, il DLS-utente che richiede la connessione inserisce nella coda un elemento di connessione (*procedura di instaurazione*). Il DLS-utente non può inserire in coda nessun altro elemento diverso da un elemento di disconnessione, se non dopo la ricezione di un elemento di connessione indicante la conferma dell'instaurazione della connessione.

L'inserimento in coda di un elemento di disconnessione costituisce l'inizio della *procedura di abbattimento* della DL-connessione. Tale procedura, che può comportare la cancellazione di altri elementi nella coda, ha termine con lo svuotamento delle code e la loro disassociazione dalla DL-connessione. L'inserimento di un elemento di disconnessione può anche rappresentare il rifiuto di un tentativo di instaurazione o l'impossibilità di completare l'instaurazione di una DL-connessione.

Durante la fase di trasferimento di dati, che consiste nell'inserire elementi di dati in coda, può essere attuata una *procedura di re-inizializzazione*, allo scopo di ri-inizializzare la connessione. Il DLS-utente che richiede la procedura di re-inizializzazione inserisce in una coda un elemento di re-inizializzazione, dopodiché il DLS-fornitore inserirà nell'altra coda un corrispondente elemento di re-inizializzazione. Se invece, è direttamente il DLS-fornitore a richiedere la re-inizializzazione, allora esso stesso inserirà in entrambe le code un elemento di re-inizializzazione.

È previsto che gli elementi emessi in una direzione possano, in alcuni casi, cambiare di ordine o essere rimossi. Dal punto di vista del modello ciò corrisponde a definire la modalità di gestione della coda, in cui possono aver luogo:

- a) *cambi di ordine* di due elementi adiacenti di tipo diverso, se la definizione dei due tipi lo consente; elementi dello stesso tipo non possono mutare di ordine;
- b) *cancellazione* di un elemento, se e solo se il successivo elemento è definito nei suoi confronti come *distruttivo*; elementi di disconnessione sono distruttivi nei riguardi di ogni altro elemento; gli elementi di re-inizializzazione sono distruttivi nei riguardi di tutti gli elementi con l'eccezione degli elementi di connessione, di disconnessione e di altri elementi di re-inizializzazione.

La *qualità di servizio* (Quality Of Service o *QOS*) di una DL-connessione è descritta da un insieme di *parametri* che sono specificati da ogni DLS-utente per descrivere le proprie necessità. Questi parametri costituiscono, per lo strato di collegamento, la base per la selezione del particolare protocollo da adottare.

I parametri di QOS possono essere divisi in due categorie, in base al modo in cui essi possono essere determinati.

La prima categoria comprende quei parametri che possono essere negoziati, tra DLS-utente e DLS-fornitore, all'atto dell'instaurazione di una DL-connessione e che possono quindi variare da connessione a connessione. Quando una DL-connessione è stata instaurata, i valori dei parametri non possono essere rinegoziati, anche se non vi è alcuna garanzia che nel frattempo tali valori possano essere effettivamente mantenuti dallo strato di collegamento. Tali parametri sono:

- la *portata*, definita come il rapporto tra il numero medio di cifre binarie trasferite unidirezionalmente in una interazione tra due DLS-utenti e il tempo necessario all'esecuzione di tale interazione; il valore di tale parametro è definito indipendentemente per ogni verso di trasferimento sulla DL-connessione;
- la *protezione*, definita come il grado di protezione e di sicurezza che lo strato di collegamento garantisce sulle informazioni trasferite su una DL-connessione; le opzioni previste sono: a) assenza di protezione; b) protezione verso ascolti non autorizzati; c) protezione contro modifiche dei dati emessi;
- la *priorità*, che specifica il grado di priorità associato ad una DL-connessione; tale parametro è utilizzato dallo strato di collegamento per

stabilire, nell'insieme di DL-conessioni da esso gestite, l'ordine in cui le DL-conessioni possono subire, in caso di necessità, una degradazione della QOS o l'abbattimento forzato.

I valori dei parametri di QOS appartenenti alla prima categoria sono negoziati tra il DLS-utente e il DLS-fornitore. Quest'ultimo potrà accettare completamente le richieste del DLS-utente, oppure, se esse non sono compatibili con lo stato attuale di funzionamento dello strato, queste dovranno essere opportunamente ridimensionate.

Se, ad esempio, un DLS-utente richiede un valore di portata troppo elevato, che non può essere soddisfatto a causa del traffico già presente sulla linea, la connessione non sarà instaurata a meno che il DLS-utente non ripieghi su un valore di portata minore, compatibile con il presente stato di congestione.

La seconda categoria comprende invece quei parametri che sono conosciuti a priori dai DLS-utenti e che non sono soggetti a negoziazione nella fase di instaurazione di una connessione. Tali parametri sono:

- il *ritardo di transito*, definito come l'intervallo di tempo intercorrente tra la emissione di una DLSDU da parte del DLS-utente emittente e la ricezione della stessa da parte del DLS-utente ricevente; il valore di tale parametro è definito indipendentemente per ogni verso di emissione sulla DL-conessione;
- il *tasso d'errore residuo*, definito come il rapporto tra il numero totale di DLSDU errate, perse o duplicate e il numero totale delle DLSDU trasferite su di una DL-conessione in un fissato intervallo di tempo;
- la *resilienza*, definita come la frequenza media secondo cui una DL-conessione è abbattuta prematuramente dal DLS-fornitore o è da questo iniziata una procedura di re-inizializzazione;

Le primitive del DL-servizio nella modalità orientata alla connessione sono riportate in Tab. III.1. L'emissione di una primitiva ad una estremità della DL-conessione ha in genere conseguenze sull'altra estremità.

<b>Servizio</b>	<b>Primitiva</b>	<b>Tipo</b>	<b>Parametri</b>
Instaurazione di una DL-connessione	DL-CONNECT	Richiesta Indicazione Risposta Conferma	Indirizzo chiamante Indirizzo chiamato Qualità di servizio
Trasferimento di dati	DL-DATA	Richiesta Indicazione	Qualità di servizio
Reset	DL-RESET	Richiesta Indicazione Risposta Conferma	Origine reset Ragioni reset
Abbattimento di una DL-connessione	DL-DISCONNECT	Richiesta Indicazione	Origine abbattimento Ragioni abbattimento

Tab. III.1 - Primitive di un DL-servizio con modalità orientata alla connessione.

Facendo riferimento ad una sola estremità della DL-connessione, tutte le possibili sequenze di primitive che hanno luogo sono definite nel diagramma di transizione di stato di Fig. III.3, in cui: a) il termine DL-DISCONNECT rappresenta o una primitiva di richiesta o una di indicazione; b) l'etichetta degli stati 5 e 6 indica l'entità, DLS-utente o DLS-fornitore, rispettivamente, che ha iniziato l'interazione locale; c) lo stato 1 ("Pronto"), che corrisponde ad assenza di DL-connessione, è quello iniziale e finale di ogni sequenza di primitive.

Riferendosi alle relazioni che legano emissione e ricezione delle primitive alle due estremità della DL-connessione, distinguiamo tra fase di instaurazione, quella di trasferimento di dati e quella di abbattimento.

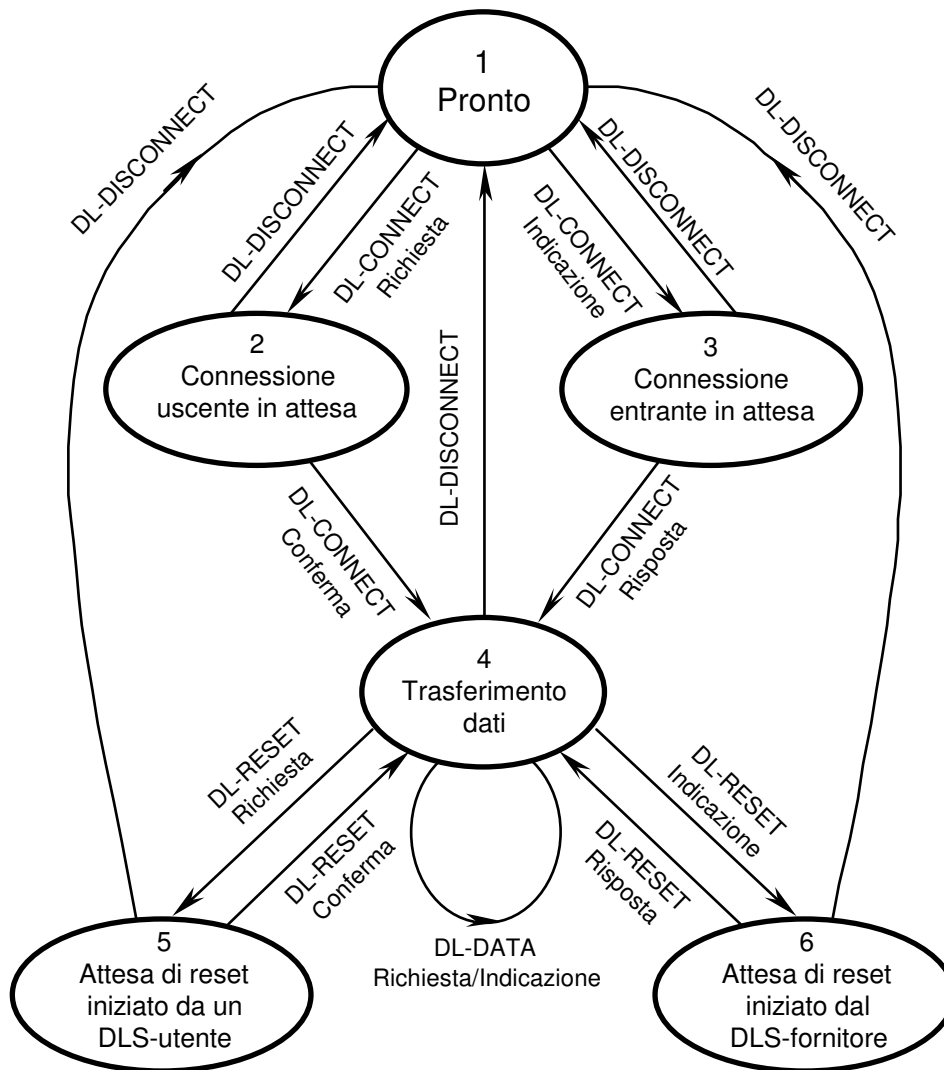


Fig. III.3 - Diagramma di transizione di stato relativo ad una estremità di DL-connesione per la fornitura di un DL-servizio orientato alla connessione.

### EVOLUZIONE DI UNA CONNESSIONE DI STRATO DI COLLEGAMENTO

Nella fase di instaurazione, le primitive DL-CONNECT sono utilizzate per instaurare una DL-connesione. Primitive di questo tipo, emesse simultaneamente in corrispondenza dei due DLSAP, sono trattate l'una indipendentemente dall'altra; possono quindi dar luogo a due, una o nessuna DL-connesione. La sequenza di primitive relative ad una instaurazione con successo di una DL-connesione è rappresentata nelle Fig. III.4a e Fig. III.4b.

La procedura di instaurazione può fallire o per l'incapacità del DLS-fornitore di instaurare una DL-connesione (Fig. III.5a) o per la volontà del DLS-utente di non accettare un'indicazione DL-CONNECT (Fig. III.5b).

Nei grafici delle Fig. III.4 e Fig. III.5, come anche in quelli dello stesso tipo che seguiranno, la presenza di una linea continua tra gli istanti di emissione di due primitive indica l'esistenza di un meccanismo protocollare di strato, tale da assicurare una definita relazione temporale tra questi istanti. Quando invece detta relazione non esiste, si sottolinea questa mancanza con il simbolo ~.

Nella fase di trasferimento di dati, le primitive del servizio di trasferimento di dati

consentono lo scambio di DLSDU in una, o in entrambe le direzioni simultaneamente, conservando la sequenzialità di dette unità. La sequenza di primitive relativa ad un trasferimento di dati è rappresentata in Fig. III.6.

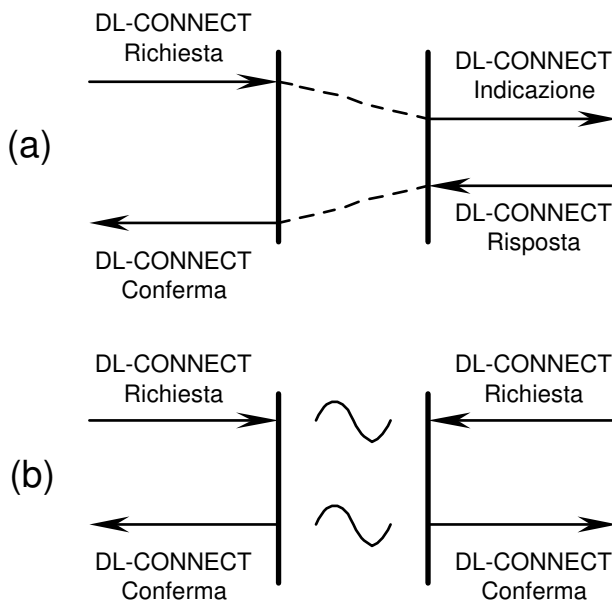


Fig. III.4 - Sequenza di primitive nel caso di instaurazione di una DL-conessione con riferimento a: (a) instaurazione iniziata da un singolo DLS-utente; (b) instaurazione iniziata simultaneamente da due DLS-utenti.

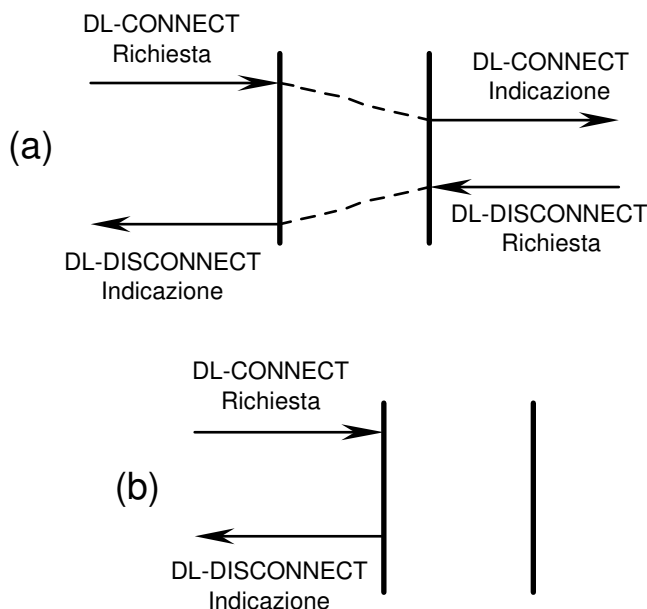


Fig. III.5 - Sequenza di primitive nel caso di mancata instaurazione di una DL-conessione con riferimento a: (a) instaurazione rifiutata di un DLS-utente; (b) instaurazione non consentita da un DLS-fornitore.

La fase di trasferimento di dati può essere interrotta da una primitiva DL-RESET o DL-DISCONNECT. Il servizio di re-inizializzazione può essere richiesto dai DLS-utenti per

risincronizzare l'uso della DL-connesione o dal DLS-fornitore per segnalare perdite di dati non recuperabili dal DL-servizio.

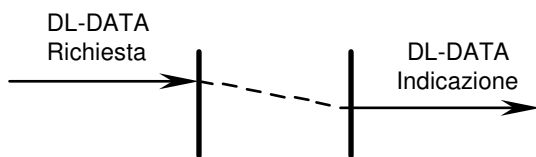


Fig. III.6 - Sequenza di primitive nel caso della fase di trasferimento di dati.

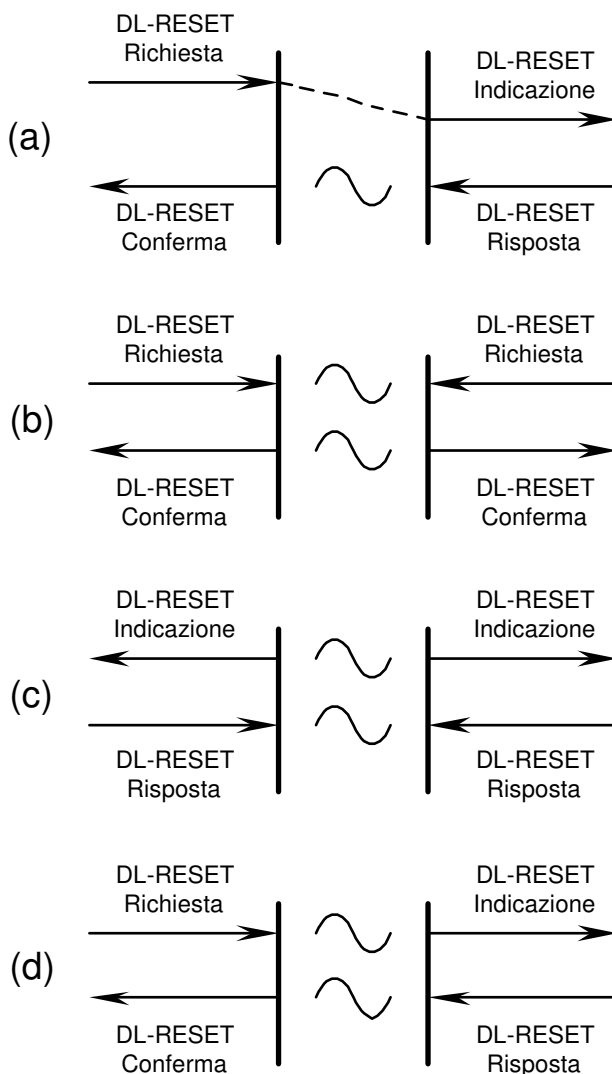


Fig. III.7 - Sequenza di primitive nel caso di attivazione di una procedura di re-inizializzazione, con riferimento a re-inizializzazione: (a) iniziata da un singolo DLS-utente; (b) iniziata simultaneamente da due DLS-utenti; (c) iniziata dal DLS-fornitore; (d) iniziata simultaneamente da un DLS-utente e dal DLS-fornitore.

Le possibili sequenze di primitive relative alla procedura di re-inizializzazione dipendono da chi origina tale procedura. Il servizio di re-inizializzazione può essere invocato da un DLS-utente (Fig. III.7a), da entrambi i DLS-utenti (Fig. III.7b), dal DLS-fornitore (Fig. III.7c) o contemporaneamente da un DLS-utente e dal DLS-fornitore (Fig. III.7d).

Nella fase di abbattimento, le primitive DL-DISCONNECT sono utilizzate per rilasciare una DL-connesione. L'abbattimento è ammesso in ogni istante indipendentemente dalla fase in cui si trova la connessione. Una richiesta di abbattimento non può essere rifiutata.

Differenti sequenze di primitive possono aver luogo, in funzione di chi origina la procedura di abbattimento di una DL-connesione, cioè dal DLS-utente (Fig. III.8a), da entrambi i DLS-utenti (Fig. III.8b), dal DLS-fornitore (Fig. III.8c) o indipendentemente da un DLS-utente e dal DLS-fornitore (Fig. III.8d).

Una richiesta di connessione presentata da un DLS-utente può non essere soddisfatta dal DLS-fornitore per vari motivi: l'impossibilità di garantire la QOS richiesta dall'utente, l'indisponibilità, per guasto o per limitazione di risorse, dell'entità di strato di collegamento chiamata oppure per malfunzionamenti dello strato fisico.

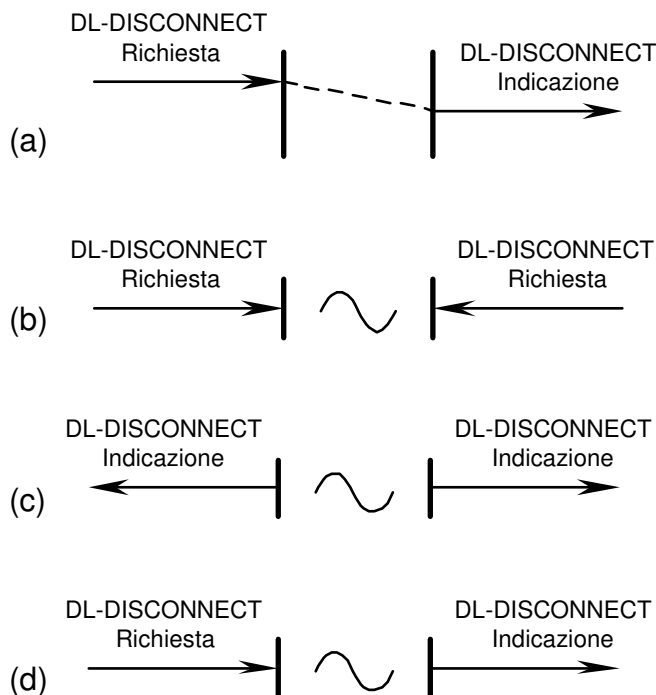


Fig. III.8 - Sequenza di primitive nel caso di abbattimento di una DL-connesione, con riferimento ad abbattimento: (a) iniziato da un DLS-utente; (b) iniziato simultaneamente da due DLS-utenti; (c) iniziato dal DL-fornitore; (d) iniziato simultaneamente da un DLS-utente e dal DLS-fornitore.

III.2.2 Funzionalità del DL-servizio con modalità senza connessione.

Il DL-servizio di tipo senza connessione fornisce i mezzi per delimitare, per identificare e per trasferire DLSDU, di lunghezza opportuna (non necessariamente costante), tra due DLS-utenti senza la necessità di instaurare una DL-connesione. Il DLS-utente, che emette una DLSDU nell'ambito di un DL-servizio senza connessione, ha la possibilità di selezionare alcuni parametri di QOS.

I parametri di QOS che un DLS-utente può selezionare all'atto della emissione di una DLSDU sono un sotto-insieme di quelli che possono essere

negoziati nella fase di instaurazione di una DL-conessione. Ad esempio, in un servizio senza connessione perde di significato la definizione di portata poiché ogni DLSDU è trattata dal DL-servizio indipendentemente dalle altre DLSDU.

Il funzionamento di un DL-servizio senza connessione e le interazioni tra i suoi DLS-utenti può essere rappresentato dal modello funzionale raffigurato in Fig. III.9. Tale modello astratto consiste in un'associazione logica permanente tra due DLSAP. Su tale associazione possono essere inseriti elementi di un solo tipo, e cioè gli elementi di dati.

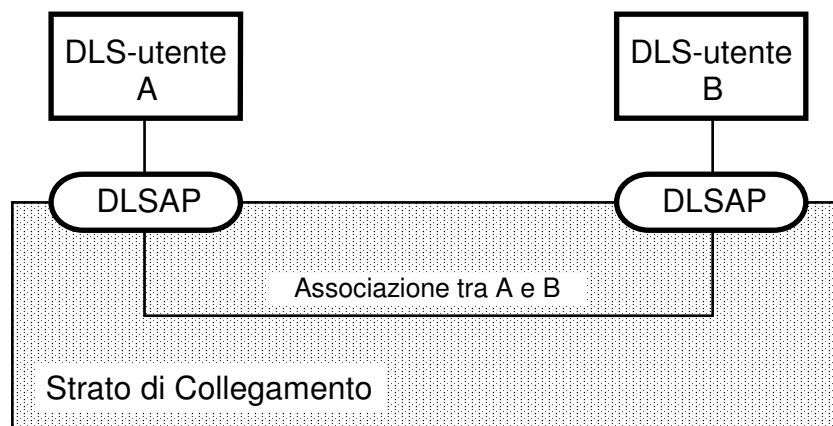


Fig. III.9 - Modello astratto di un DL-servizio nella modalità senza connessione.

In termini generali, un DL-servizio senza connessione non è in grado di assicurare ad un DLS-utente l'assenza di perdite, di duplicazioni o di modifiche dell'ordine di consegna degli elementi di dati. Tuttavia, in casi particolari, una o più delle precedenti prestazioni possono essere ugualmente soddisfatte. In tali casi, un DLS-utente ha la possibilità di selezionare un protocollo di strato di rete che tenga conto di queste particolarità di servizio.

Le LAN, caratterizzate, come più volte ricordato, da un servizio di strato di collegamento senza connessione, hanno la caratteristica di mantenere in ogni caso la sequenza di consegna dei dati rispetto a quella di emissione. Tali reti sono infatti costituite da un unico canale trasmissivo che collega tutte le stazioni connesse alla rete. In tali condizioni risulta impossibile che l'ordine di emissione delle DLSDU sia alterato. Questa particolarità può essere sfruttata per semplificare il protocollo dello strato superiore, eliminando la funzione di riordinamento.

Una caratteristica di base della modalità senza connessione risiede nel fatto che, mancando un'associazione tra DLS-utenti simile a quella offerta dal concetto di connessione, le caratteristiche del servizio di trasferimento delle informazioni non sono fissate su base connessione.

Un DLS-utente può, all'atto dell'invio di una DLSDU, richiedere una determinata QOS; tuttavia tale richiesta è basata su una conoscenza a priori delle caratteristiche e delle potenzialità del DL-servizio.

I parametri che un DLS-utente può controllare sono il *ritardo di transito*, il *tasso d'errore residuo*, il *grado di protezione delle informazioni* e la loro *priorità*. La definizione di tali parametri è analoga a quella data nel caso di servizio orientato alla connessione.

Le uniche primitive del DL-servizio nella modalità senza connessione sono la richiesta e la indicazione DL-UNITDATA. La sequenza di emissione ammessa è quella descritta in Fig. III.10, mentre in Fig. III.11 è mostrato il diagramma di transizione di stato di un DLSAP nella modalità senza connessione.

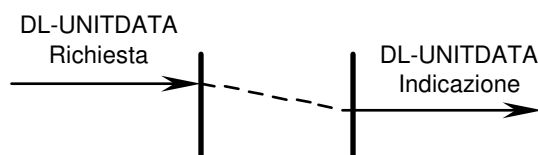


Fig. III.10 - Sequenza di primitive nel trasferimento di dati con riferimento al caso di un DL-servizio con modalità senza connessione.

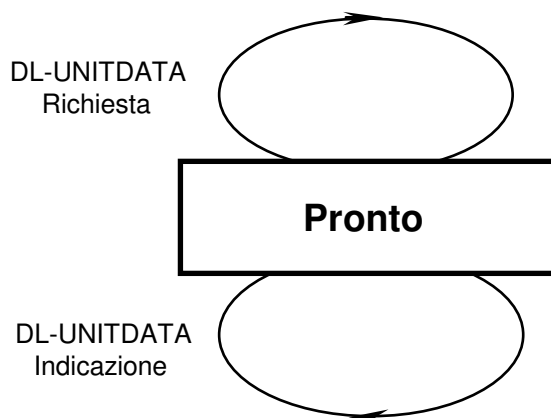


Fig. III.11 - Diagramma di transizione di stato relativo a un DL-SAP per la fornitura di un DL-servizio nella modalità senza connessione.

Una primitiva di richiesta DL-UNITDATA è usata per richiedere il trasferimento di una DLSDU. Ogni DLSDU deve essere autosufficiente ed è trattata dal DL-servizio in modo completamente indipendente da qualsiasi altra DLSDU.

L'autosufficienza deve intendersi nel senso che tutte le informazioni necessarie al corretto trasferimento della DLSDU verso la destinazione devono essere fornite dal DLS-utente al DLS-fornitore insieme ai dati veri e propri.

Il significato di questo requisito di autosufficienza delle DLSDU è evidente se si confronta il formato delle primitive DL-DATA, utilizzate nel modo orientato alla connessione, con quello delle primitive DL-UNITDATA.

Nelle DL-DATA sono contenuti solo i dati da trasferire, mentre le DL-UNITDATA contengono, oltre ai dati, anche gli indirizzi di sorgente e di destinazione unitamente all'insieme dei parametri di QOS.

È evidente che, nel primo caso, le informazioni di indirizzo e di QOS non sono necessarie. Infatti, la creazione di una connessione consente allo strato di collegamento, da un lato, di assumere preventivamente tali informazioni e, dall'altro, di associare logicamente le DLSDU provenienti dal DLS-utente alla connessione stessa.

### III.3 Protocolli di strato di collegamento

Come illustrato nei paragrafi precedenti, lo strato di collegamento ha il compito specifico di garantire allo strato di rete un trasferimento affidabile delle DLSDU, sfruttando il servizio offerto dallo strato fisico.

Ciò è assicurato mediante la cooperazione tra le DL-entità poste agli estremi della connessione di strato fisico. Tale cooperazione è regolata da un opportuno protocollo di strato di collegamento (*DL-protocollo*). Le unità di dati utilizzate da un DL-protocollo (*DLPDU*) sono indicate con terminologia varia in contesti diversi; il termine più diffuso è quello di *trama*, che, pertanto, sarà usato nel seguito di questa sezione.

Limiteremo la nostra attenzione al caso di protocolli che offrono un servizio orientato alla connessione. Tuttavia le principali funzioni svolte ed i criteri fondamentali di progetto possono essere facilmente estesi al caso di protocolli per servizi senza connessione.

Le funzioni elementari che un DL-protocollo deve svolgere sono:

- *la delimitazione e l'identificazione delle trame*; le DL-entità devono essere in grado di riconoscere, nella sequenza continua di cifre binarie consegnata dallo strato fisico, l'inizio e la fine delle trame tra loro scambiate, al fine di operare su di esse le successive funzioni di controllo;
- *la rivelazione degli errori di trasmissione*; lo strato fisico non garantisce l'assenza di errori nel trasferimento delle sequenze di cifre binarie; è compito dello strato di collegamento rivelare il verificarsi di tali eventi al fine di proteggere lo strato superiore da errate interpretazioni dell'informazione;
- *il recupero del corretto trasferimento delle trame in caso di errori*; una trama può essere corrotta, perduta, duplicata o può essere alterato il suo ordine di

- sequenza, rispetto ad altre trame, a causa di errori durante il suo trasferimento tra le DL-entità; al verificarsi di tali eventi, il DL-protocollo deve, se richiesto, poter recuperare il corretto funzionamento della DL-connessione;
- *il controllo di flusso*; nel caso in cui il ritmo di arrivo delle trame sia superiore a quello di elaborazione della DL-entità ricevente, occorre evitare la saturazione delle risorse della DL-entità e, quindi, una perdita di informazioni; la funzione è attuata costringendo la DL-entità emittente a ridurre il suo ritmo di emissione di trame;
  - *la gestione della DL-connessione*; un DL-protocollo deve comprendere i mezzi per instaurare e per abbattere una DL-connessione, nonché, eventualmente, per ri-inizializzarla.

A tutte le precedenti funzioni corrispondono, all'interno di un DL-protocollo, opportune procedure dedicate alla loro attuazione. I protocolli di strato di collegamento si differenziano, quindi, in base alle modalità con cui le funzioni sopra elencate sono eseguite.

Le procedure per l'attuazione di ognuna delle funzioni di un DL-protocollo possono essere, come sarà illustrato ampiamente nel seguito, molto diverse tra loro, pur essendo tutte orientate al soddisfacimento dello stesso obiettivo. Ovviamente, procedure diverse sono caratterizzate da prestazioni diverse e, in relazione alle condizioni al contorno, sono più o meno adatte ad essere utilizzate.

Ad esempio, alcuni protocolli prevedono, nel caso di trama ricevuta errata, l'invio all'entità corrispondente di una richiesta di ri-emissione di tutte le trame a partire da quella errata; in altri protocolli, invece, la richiesta di ri-emissione riguarda solo la trama ricevuta errata (*ri-emissione selettiva*).

È evidente che la ri-emissione selettiva migliora le prestazioni sia in termini di efficienza che di ritardo. Tuttavia richiede, rispetto alla prima modalità, una maggiore complessità funzionale della DL-entità. La scelta della modalità è quindi strettamente dipendente dalle condizioni in cui il protocollo dovrà operare.

Un importante criterio di classificazione dei DL-protocolli fa riferimento alla definizione della minima entità informativa che il protocollo è in grado di trattare. In accordo a tale criterio si distinguono: a) *protocolli orientati al carattere*; b) *protocolli orientati al bit*.

Nei protocolli orientati al carattere, la trama è costituita da una sequenza di caratteri appartenenti ad uno specifico alfabeto. La lunghezza di una trama è quindi un multiplo della lunghezza di un carattere. In questo caso, il funzionamento di una DL-entità è basato sul riconoscimento di ogni singolo

carattere della trama.

Nell'ambito dell'alfabeto utilizzato dal protocollo, è definito un insieme di caratteri speciali, detti *caratteri di controllo*, a cui sono demandate le funzioni di controllo dello scambio di informazioni. I dieci caratteri di controllo generalmente usati per questo scopo sono riportati in Tab. III.2, unitamente al loro significato.

<b>Sigla</b>	<b>Carattere di Controllo</b>	<b>Significato</b>
SOH	Start of heading	Identifica l'inizio della sequenza di caratteri che costituisce l'intestazione del messaggio
STX	Start of text	Identifica l'inizio della sequenza di caratteri che costituisce la parte informativa del messaggio
ETX	End of text	Identifica la fine della sequenza di caratteri che costituisce la parte informativa del messaggio
EOT	End of transmission	Segnala la fine della trasmissione da parte della stazione emittente
ENQ	Enquire	Sollecita l'invio di una risposta da parte di un'altra stazione
ACK	Acknowledge	Segnala la ricezione senza errori di un blocco informativo
NAK	Negative acknowledge	Segnala la ricezione con errori di un blocco informativo
SYN	Synchronous idle	Serve per stabilire e mantenere la sincronizzazione di carattere tra stazione emittente e ricevente
ETB	End transmission block	Indica la fine di un blocco dati
DLE	Data link escape	Cambia il significato del carattere immediatamente seguente, se questo è un altro carattere di controllo

Tab. III.2 - Elenco dei caratteri di controllo nei protocolli orientati al carattere.

Per chiarire meglio l'utilizzazione dei caratteri di controllo in un protocollo

orientato al carattere, supponiamo che una DL-entità debba inviare un testo ad un'entità corrispondente. In questo caso l'evoluzione della comunicazione è raffigurata in Fig. III.12. Prima dell'invio del testo, la DL-entità si assicura che l'entità corrispondente sia attiva con l'emissione del carattere ENQ. Avuta risposta affermativa, tramite la ricezione di un carattere ACK, l'invio del testo inizia con il carattere STX e termina con il carattere ETX. Ricevuto il riscontro di ricezione, la comunicazione si chiude con l'invio del carattere EOT.

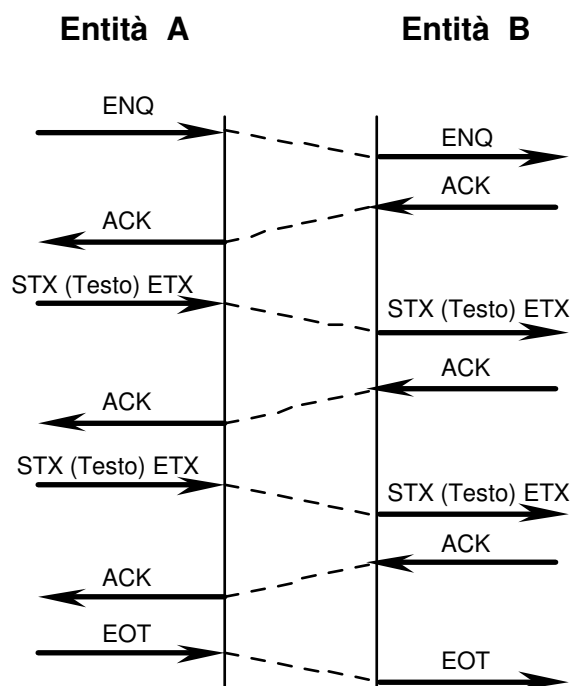


Fig. III.12 - Esempio di colloquio con un protocollo orientato al carattere.

Il limite principale dei protocolli orientati al carattere consiste nella stretta dipendenza dall'alfabeto adottato e nella loro scarsa flessibilità a supportare sofisticate procedure di controllo. Ciò li rende inadatti ad applicazioni che richiedono elevate prestazioni di efficienza e di affidabilità.

I protocolli orientati al bit sono stati sviluppati per superare i limiti che i protocolli orientati al carattere hanno progressivamente manifestato all'aumentare delle prestazioni ad essi richieste.

La caratteristica fondamentale di questi protocolli risiede nel fatto che la minima entità informativa gestita dalle DL-entità è la singola cifra binaria. Ciò consente di svincolare la lunghezza della singola trama dalla lunghezza del carattere. Ogni trama è composta da un certo numero di gruppi di cifre binarie, detti *campi*, la cui lunghezza può essere qualunque. Le funzioni di controllo della comunicazione non sono attuate da caratteri speciali, bensì dalle particolari configurazioni assunte dai campi di una trama.

La struttura generale di una trama in un protocollo orientato al bit è raffigurata in Fig. III.13. Nell'intestazione sono contenute le informazioni riguardanti il significato della trama stessa e quelle riguardanti il controllo dello stato della DL-conessione. In questo caso, quindi, esistono sia *trame di controllo*, che trasportano solo informazioni di gestione della connessione, sia *trame informative*, che invece convogliano anche informazioni d'utente.



Fig. III.13 - Struttura generale di una trama in un protocollo orientato al bit.

Nel seguito saranno esaminate separatamente le singole funzioni elementari che un protocollo di collegamento deve svolgere, individuando le alternative possibili per la loro attuazione con riferimento alle due diverse classi di protocolli. Successivamente, sarà descritto, a titolo d'esempio, il protocollo X.25 di livello 2, che è il protocollo di strato di collegamento oggi utilizzato per l'accesso alle reti per dati a commutazione di pacchetto.

### III.3.1 Delimitazione delle trame.

La funzione di delimitazione ha lo scopo di mettere in condizione le DL-entità di individuare, nella sequenza di cifre binarie consegnata dallo strato fisico, l'inizio e la fine delle singole trame e quindi di operare su di esse i successivi controlli previsti dal protocollo.

Le modalità principali con cui tale funzione può essere attuata sono: a) il conteggio dei caratteri o delle cifre binarie; b) l'inserimento di caratteri di inizio e di fine trama; c) l'inserimento di campi di inizio e di fine trama; d) la violazione del codice di linea utilizzato nello strato fisico.

È evidente che la modalità b) è possibile solo nell'ambito dei protocolli orientati al carattere, mentre la c) è tipica dei protocolli orientati al bit. La soluzione a) può essere adottata in entrambi i casi. L'ultima alternativa è invece una soluzione ibrida che sfrutta alcune particolarità del servizio di strato fisico.

Il metodo di conteggio dei caratteri o delle cifre binarie si basa sull'inserimento, nell'intestazione della trama, di un carattere (o di un campo) che indichi il numero di caratteri (o di cifre binarie) di cui è composta la trama. In fase di ricezione, la DL-entità è in grado di stabilire la fine della trama semplicemente conteggiando il numero di caratteri (o di cifre binarie) ricevuti.

È evidente che il maggior problema insito in tale metodo consiste nel fatto

che un errore trasmissivo può comportare un'errata determinazione della lunghezza della trama e quindi l'impossibilità di riconoscere il corretto inizio della trama successiva. Per questa ragione tale metodo è raramente usato.

In Fig. III.14 è mostrato un esempio di delimitazione di trama sulla base del metodo di conteggio dei caratteri.

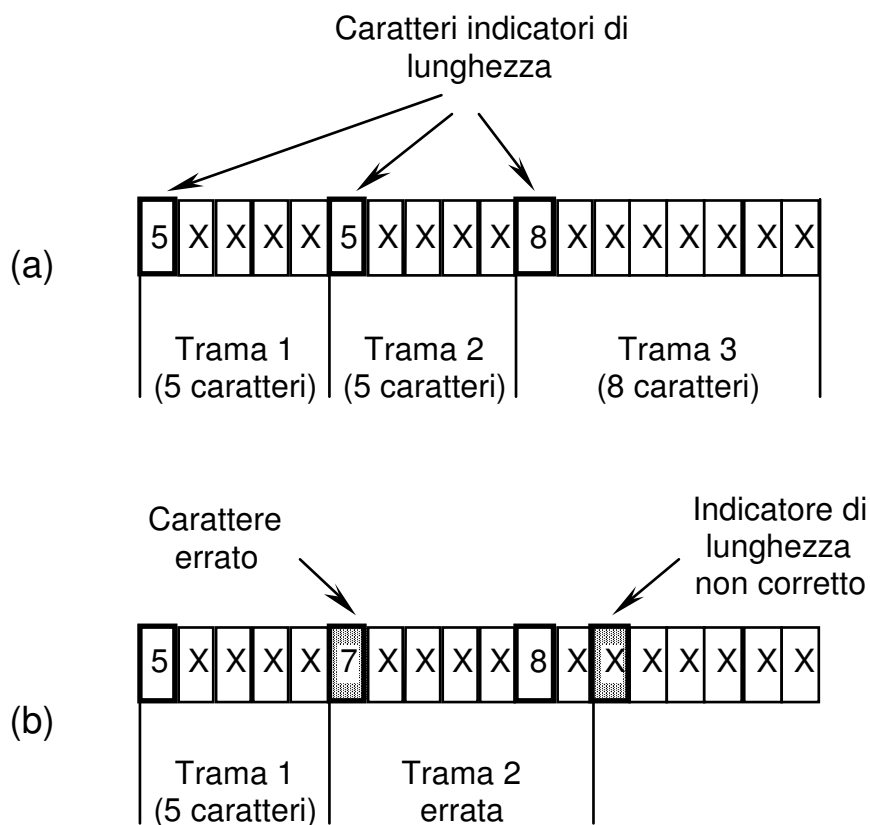


Fig. III.14 - Delimitazione della trama con il metodo del conteggio dei caratteri: (a) procedimento di ricostruzione delle trame; (b) ricostruzione delle trame in caso di errore su un carattere indicatore di lunghezza.

Nell'esempio mostrato, il primo carattere di ogni trama indica il numero di caratteri di cui è composta la trama stessa. In Fig. III.14a è mostrato il procedimento di ricostruzione delle trame a partire da una stringa di caratteri corretta, mentre in Fig. III.14b è mostrato il risultato che si avrebbe nel caso di errore nel primo carattere della seconda trama.

Il secondo metodo utilizza, per delimitare le trame, particolari caratteri dell'alfabeto utilizzato per codificare l'informazione. Ad esempio, nel caso di utilizzazione dell'alfabeto ASCII, le sequenze di caratteri utilizzate per identificare l'inizio e la fine di una trama sono rispettivamente DLE STX e DLE ETX.

Su un principio del tutto analogo si basa il terzo metodo. In questo caso, il

carattere o la sequenza di caratteri di delimitazione è sostituita da una particolare configurazione di cifre binarie, denominata *flag*. Ad esempio, nel protocollo X.25 di livello 2, la configurazione del campo di flag è 01111110, e cioè una parola di 8 cifre binarie costituita da uno "0" seguito da sei "1" e conclusa da uno "0".

Il problema che sorge nell'utilizzazione di questi due metodi è quello che la configurazione di caratteri o di cifre binarie, utilizzata per segnalare l'inizio e la fine di una trama può essere contenuta all'interno della trama stessa nei dati d'utente. È evidente come tale eventualità interferisca nella funzione di ricostruzione della trama e possa comportare errori nell'evoluzione della connessione.

Per evitare questo fenomeno si ricorre all'operazione di *riempimento* (*stuffing*). Tale operazione consiste nell'inserire, al momento della emissione, opportuni caratteri (*riempimento di carattere*) o cifre binarie (*riempimento di bit*), in modo che non sia possibile la simulazione, all'interno della trama, della sequenza di caratteri o di cifre binarie che delimita la trama stessa.

In un protocollo orientato al carattere, se la sequenza di caratteri di inizio e di fine trama è quella precedentemente indicata, la DL-entità sorgente inserisce un secondo carattere DLE ogniqualvolta rivela, all'interno della trama, un carattere DLE. È quindi evidente che, in tal modo, le sequenze di inizio e di fine trama DLE STX e DLE ETX sono uniche.

Nel caso di protocollo orientato al bit e di configurazione di flag uguale a quella citata, la medesima funzione viene eseguita aggiungendo al momento della trasmissione un bit "0" dopo ogni sequenza di cinque "1", tranne ovviamente nel caso di trasmissione del flag stesso.

Al momento della ricezione occorre ovviamente riconoscere i caratteri (cifre binarie) addizionali ed eliminarli (operazione di *svuotamento*), dato che essi non contengono informazione significativa nei riguardi del DLS-utente. Con tale accorgimento si riesce ad assicurare la completa *trasparenza* delle informazioni rispetto alla funzione di delimitazione.

In Fig. III.15 è mostrato un esempio di delimitazione di trama nel caso di protocollo orientato al carattere. In Fig. III.15a è mostrata la sequenza originaria, mentre la Fig. III.15b mostra la sequenza di caratteri dopo l'operazione di riempimento. L'entità ricevente, dopo aver rivelato i punti di inizio e di fine trama, provvede a ricostruire la sequenza originaria di caratteri (Fig. III.15c) scartando tutti i caratteri DLE che seguono un carattere DLE.

Un esempio di applicazione della funzione di riempimento nel caso di protocollo orientato al bit è mostrata in Fig. III.16. La DL-entità sorgente

inserisce nella sequenza di cifre binarie originaria (Fig. III.16a) due bit "0" dando luogo alla sequenza di cifre binarie effettivamente trasmessa (Fig. III.16b). La DL-entità ricevente, nell'esaminare la sequenza di cifre binarie ricevuta, individua l'inizio o la fine trama se dopo cinque "1" viene ricevuto un ulteriore "1". Se invece l'ulteriore cifra binaria ricevuta è uno "0", questa è scartata.

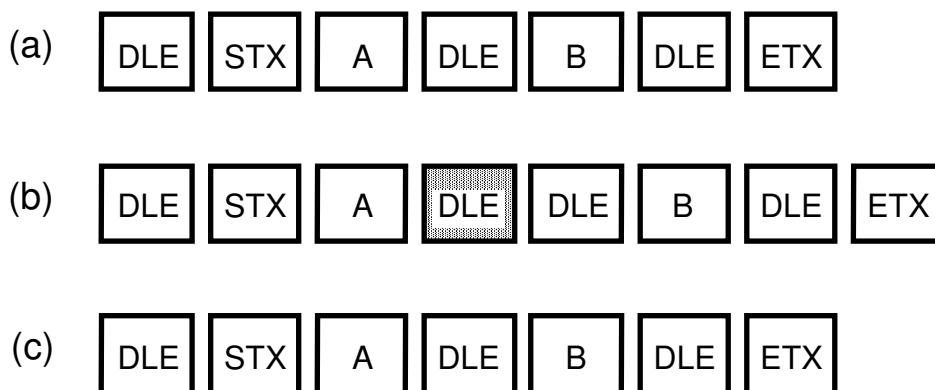


Fig. III.15 - Operazione di riempimento di carattere: (a) sequenza originaria di caratteri; (b) sequenza dopo l'operazione di riempimento; (c) sequenza dopo l'operazione di svuotamento.

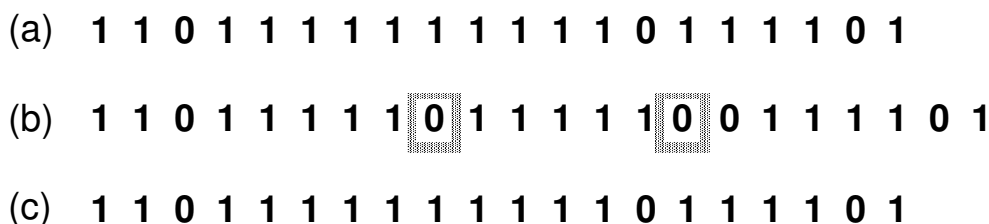


Fig. III.16 - Operazione di riempimento di bit: (a) sequenza originaria di cifre binarie; (b) sequenza dopo l'operazione di riempimento; (c) sequenza dopo l'operazione di svuotamento.

L'ultima alternativa, la violazione di codice, è applicabile solo se la codifica di linea utilizzata dallo strato fisico è di tipo particolare e contiene ridondanze che possono essere sfruttate per fini diversi.

Ad esempio, nel caso del codice di Manchester (cfr. par. II.3), un bit "1" è codificato con una transizione alto-basso, la transizione basso-alto identifica il bit "0", mentre alle configurazioni alto-alto e basso-basso non è associato alcun significato. Queste ultime possono quindi essere utilizzate per la funzione di delimitazione.

Il vantaggio della tecnica a violazione di codice è quello di non richiedere l'operazione di riempimento. Tuttavia il suo uso è limitato a casi particolari, a causa dell'onere connesso all'utilizzazione di un adatto codice di linea.

Nelle LAN, la delimitazione delle trame mediante la tecnica a violazione di

codice è la soluzione oggi più largamente adottata. Questa soluzione è stata preferita alle altre poiché l'uso del codice di Manchester è necessario per il trasferimento, tra le stazioni di rete, dell'informazione di temporizzazione. La funzione di delimitazione è quindi ottenuta senza un aggravio in termini di complessità di sistema.

### III.3.2 Rivelazione degli errori.

La funzione di rivelazione degli errori ha il compito di segnalare alla DL-entità ricevente se il valore di alcune cifre binarie o gruppi di cifre binarie di una trama sia stato alterato da cause accidentali durante il transito sul canale trasmissivo.

L'origine più comune di errore nei canali trasmissivi su mezzi in rame è il rumore impulsivo. Questo disturbo si presenta normalmente sotto forma di impulsi che hanno durata media dell'ordine dei 10 ms. In queste condizioni il numero di cifre binarie che possono essere alterate dipende dal ritmo binario utilizzato in trasmissione. Nel caso, ad esempio, di una trasmissione vocale, il valore del ritmo binario in trasmissione è di 64 kbit/s e quindi un impulso di rumore altera mediamente 640 bit.

L'approccio seguito per realizzare la funzione di rivelazione di errore è quello di aggiungere ad ogni carattere, o ad ogni trama, cifre binarie addizionali, in modo che, dall'esame di queste ultime, la DL-entità ricevente possa individuare, entro certi margini di probabilità di successo, se la trama è corretta o meno. Il recupero dell'errore è quindi demandato alle procedure di ri-emissione di cui tratteremo nel paragrafo seguente.

Un ulteriore funzione che è possibile attuare, strutturando opportunamente le cifre binarie addizionali da aggiungere ai singoli caratteri, è quella di *correzione degli errori*. In questo caso, una DL-entità non solo è in grado di rivelare, entro certi margini di probabilità di successo, gli errori trasmissivi, ma ha anche la possibilità di individuare quali cifre binarie o gruppi di cifre binarie siano stati alterati e, quindi, di apportare le necessarie correzioni, senza la necessità di richiedere la ri-emissione della trama.

È evidente che lo svolgimento di questa funzione richiede un numero di cifre binarie addizionali molto maggiore di quello necessario per la semplice rivelazione e riduce quindi l'efficienza del trasferimento; è pertanto utilizzata solo in casi particolari. Nel seguito ci occuperemo solo delle possibili alternative per l'attuazione della funzione di rivelazione.

Dalla teoria dei codici è noto che, per rivelare  $d$  errori su una stringa di cifre binarie, è necessario aggiungere ad essa almeno  $d$  cifre binarie addizionali,

mentre per rivelarli e correggerli sono necessari almeno  $2d$  cifre binarie. Da questi dati emerge come l'aggiunta di cifre binarie protettive nel caso di codici correttori d'errore è il doppio di quella necessaria per la semplice rivelazione.

I metodi più usati per la rivelazione d'errore sono tre: a) il *controllo di parità*; b) il *controllo di parità a blocchi*; c) l'impiego di *codici rivelatori d'errore polinomiali*.

Nel primo caso, viene aggiunto, ad ogni carattere in emissione, una singola cifra binaria, detta *bit di parità*. Il valore di questo bit è determinato sulla base del numero di cifre binarie "1" che compongono il carattere da emettere.

Si parla di *parità pari* se il bit aggiunto viene posto ad "1" nel caso in cui il numero di cifre binarie "1" da trasferire (compreso il bit di parità) è pari. Se, viceversa, il bit di parità è posto ad "1" se il numero di "1" del carattere è dispari, si parla di *parità dispari*.

Supponiamo di avere un carattere di lunghezza uguale a 7 cifre binarie la cui configurazione è 0101001. Poiché il numero di "1" contenuti nel carattere è uguale a tre, il bit di parità da aggiungere a tale configurazione è uguale a "1" se è utilizzata la parità pari, mentre è uguale a "0" se è invece utilizzata la parità dispari. Le configurazioni delle cifre binarie trasmesse sono quindi uguali a 01010011 nel primo caso e a 01010010 nel secondo.

In ricezione, è rivelato un errore se non esiste corrispondenza tra il numero di "1" ricevuti ed il valore del bit di parità.

Tale meccanismo di rivelazione è soddisfacente quando il numero di cifre binarie da trasferire è limitato e il tasso di errore è basso. È evidente però che, nel caso di errori di tipo a raffica, in cui il numero di cifre binarie errate consecutive è maggiore di uno, la protezione offerta da tale meccanismo è molto bassa. È sufficiente infatti un numero pari di inversioni affinché l'errore non venga rivelato.

Il verificarsi di due inversioni nelle cifre binarie di un carattere rende non efficace l'impiego del bit di parità. Riprendendo il caso esaminato in precedenza e supponendo l'utilizzazione della parità pari, se la configurazione ricevuta è 10010011, ovvero se si sono verificate le inversioni delle prime due cifre binarie, il controllo di parità darebbe esito positivo, poiché il numero complessivo di "1" è rimasto pari.

Il secondo metodo è una diretta generalizzazione del precedente ed è ottenuto controllando le parità, non solo su ogni singolo carattere, ma anche su insiemi di questi, detti *blocchi*.

Come nel metodo precedente, ad ogni carattere viene assegnato un bit di parità, ma in questo caso, ulteriori bit di parità sono generati per controllare la

parità dell'insieme delle cifre binarie occupanti la stessa posizione nei caratteri appartenenti ad un blocco. Tali bit sono trasmessi in coda ai caratteri costituenti il blocco.

In Fig. III.17a è mostrato un esempio di utilizzazione del controllo di parità a blocchi, supponendo che l'alfabeto sia composto da caratteri di 7 cifre binarie, che la dimensione di un blocco consista di 8 caratteri e che, inoltre, sia utilizzata la parità pari. Si noti che la struttura del blocco è quella di una matrice in cui il numero di "1" su ogni riga e su ogni colonna è sempre pari.

Notiamo che l'ultimo bit in basso a destra può essere interpretato anche come il bit di parità della stringa di cifre binarie complessiva, generata da tutti i caratteri del blocco. Tale bit ha il significato di parità pari se le modalità utilizzate sulle righe e sulle colonne sono uguali (parità entrambe pari o entrambe dispari); altrimenti indica una parità dispari.

Il vantaggio del controllo di parità a blocchi è quello di poter rivelare raffiche di errori di lunghezza uguale o inferiore ad  $n$ , se  $n$  è il numero di caratteri costituenti un blocco. La probabilità di non rivelare raffiche di errori di lunghezza maggiore di  $n$  è invece diversa da zero.

In Fig. III.17b è mostrato un esempio di errore trasmissivo sul blocco raffigurato in Fig. III.17a, che non può essere rivelato dal controllo di parità a blocchi utilizzato. Nel caso mostrato le inversioni di cifre binarie avvengono in numero pari e contemporaneamente sia sulle righe che sulle colonne corrispondenti. Entrambi i controlli di parità risultano quindi inefficaci.

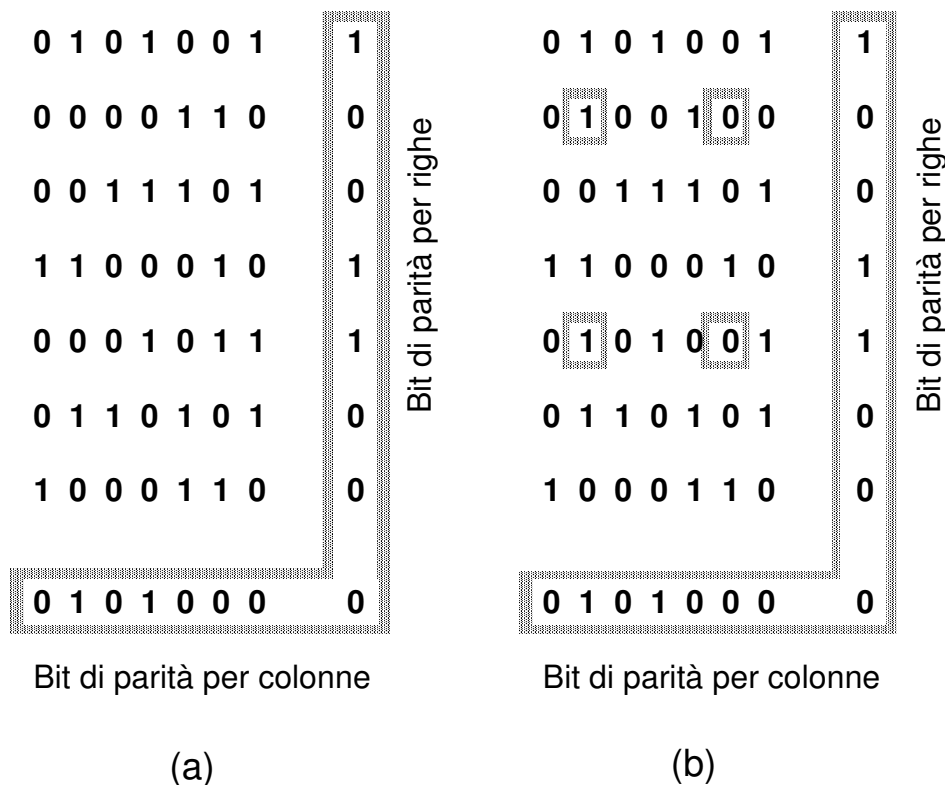


Fig. III.17 - Controllo di parità a blocchi: (a) blocco di carattere originario e configurazione dei bit di parità; (b) errore trasmissivo non rivelato dal controllo di parità.

L'ultimo metodo è quello basato sull'impiego dei codici polinomiali. Le singole cifre binarie di una stringa da emettere sono trattate come coefficienti (di valore "0" e "1") di un polinomio  $P(x)$ . Le cifre binarie di una trama di lunghezza uguale a  $k$  cifre binarie sono quindi considerate come i coefficienti di un polinomio completo di grado  $k-1$ . In particolare, l' $i$ -ma bit della trama è il coefficiente del termine  $x^{i-1}$  nel polinomio  $P(x)$ .

Quando è utilizzato il metodo dei codici polinomiali, le DL-entità emittente e ricevente utilizzano un polinomio comune  $G(x)$ , detto *polinomio generatore*.

Questo polinomio è usato dall'entità emittente come divisore del polinomio rappresentato dalle cifre binarie della trama da trasmettere. Il resto della divisione  $P(x)/G(x)$ , sotto forma di cifre binarie, è inserito in un apposito campo, detto *sequenza di verifica di trama* (Frame Check Sequence o *FCS*), che è posto in coda alla trama.

La DL-entità ricevente esegue, con lo stesso polinomio  $G(x)$ , la stessa operazione di divisione e confronta il resto ottenuto localmente con quello contenuto nella trama. Se i due resti sono uguali, la trama è considerata corretta, altrimenti uno o più errori si sono verificati nel corso della trasmissione e la

trama non è accettata.

Un codice polinomiale ad  $r$  bit è in grado di rivelare qualsiasi errore o gruppo di errori di lunghezza inferiore o uguale ad  $r$ .

Il polinomio generatore  $G(x)$  utilizzato nel protocollo X.25 di livello 2 ha l'espressione  $x^{16}+x^{12}+x^5+x^0$  che, in forma binaria, assume la configurazione 1000010000001000001, di lunghezza uguale a 17 cifre binarie. Per utilizzare tale polinomio è necessario che la lunghezza della trama sia superiore o uguale a 18 cifre binarie. Il resto della divisione tra il polinomio rappresentato dalla trama ed il polinomio generatore è un polinomio di grado minore o uguale a 15. Quindi, il campo FCS deve essere di lunghezza uguale a 16 cifre binarie.

### III.3.3 *Recupero in caso di errori.*

La funzione di recupero in caso di errori ha lo scopo di riportare alla normalità il flusso delle trame scambiate tra le DL-entità nel caso di errori, di duplicazioni, di perdite di trame o di alterazioni nel loro ordine originale.

La procedura più usata per assicurare un trasferimento affidabile delle trame è basata sull'invio, da parte della DL-entità ricevente verso quella emittente, di particolari trame di controllo che recano i *riscontri*, positivi o negativi, delle trame ricevute.

Un riscontro positivo indica la corretta ricezione della trama, mentre un riscontro negativo segnala che l'entità ricevente ha riscontrato errori nella ricezione della trama; è allora necessaria la sua ri-emissione.

Un esempio di utilizzazione dei riscontri positivi o negativi è mostrato in Fig. III.18. Ad ogni ricezione di una trama corretta, l'entità ricevente invia una trama di riscontro (ACK). La ricezione di un ACK viene interpretata, dall'entità emittente, come un'autorizzazione a proseguire l'emissione delle trame. In caso di ricezione di una trama errata, viene invece emessa una trama di riscontro negativo (NACK). Tale trama impone all'entità emittente la ri-emissione della trama.

Il precedente meccanismo non è sufficiente, da solo, a garantire il recupero da qualsiasi situazione creatasi in seguito a eventi di errore. Ad esempio, se un'entità emette una trama, rimanendo in attesa di un riscontro da parte dell'entità corrispondente e se tale trama viene persa, ad esempio per malfunzionamenti delle apparecchiature, si creerebbe una situazione di stallo (*deadlock*), in cui entrambe le entità rimangono indefinitamente in attesa di un evento proveniente dall'altra entità.

Per evitare questi *stati di stallo*, si introduce nel protocollo un opportuno *temporizzatore* (*timer*). Quando un'entità emette una trama, fa

contemporaneamente partire un temporizzatore. Questo è regolato in modo da arrestarsi dopo un intervallo di tempo sufficiente affinché, in condizioni normali di funzionamento, la trama arrivi a destinazione, sia elaborata e ne sia ricevuto il relativo riscontro. Se il temporizzatore si arresta prima dell'arrivo del riscontro della trama, questa viene ri-emessa.

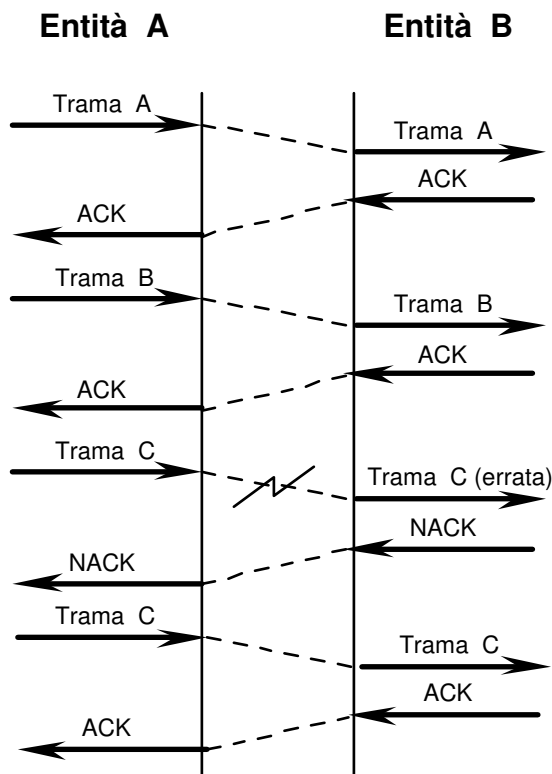


Fig. III.18 - Esempio di utilizzazione di riscontri positivi e negativi.

In Fig. III.19a è mostrato il caso di ri-emissione di una trama avvenuta per l'azzeramento del temporizzatore. Nell'esempio, la trama è stata perduta a causa di malfunzionamenti delle apparecchiature e quindi l'entità ricevente non ha avuto la possibilità di inviare alcun tipo di riscontro. La ri-emissione risolve in questo caso il problema poiché permette, entro i margini di probabilità di successo consentiti dalla funzione di rivelazione di errore, la ricezione corretta della trama.

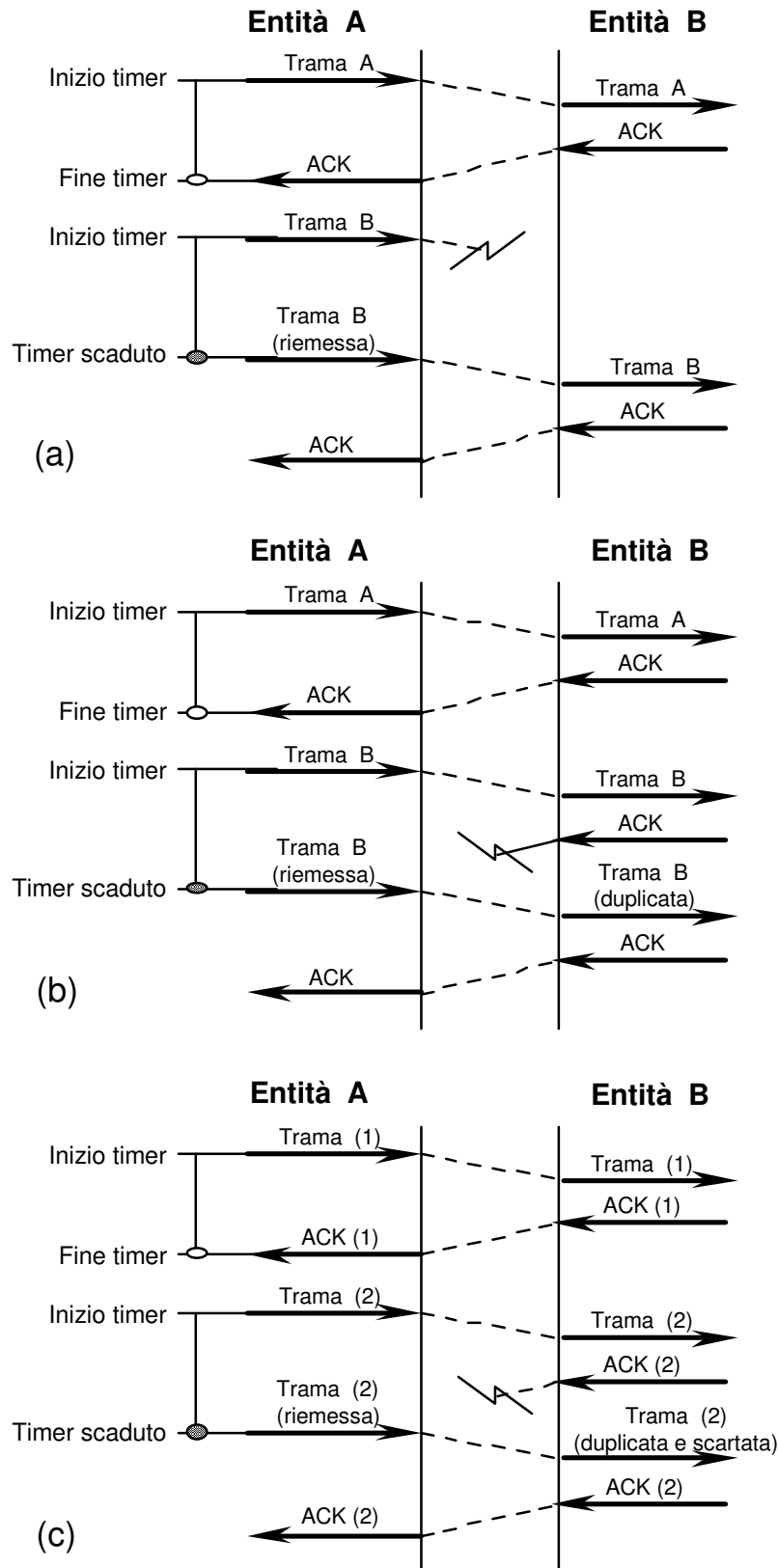


Fig. III.19 - Esempi di uso dei temporizzatori e della numerazione delle trame: (a) recupero di errore su una trama; (b) duplicazione di una trama in seguito a perdita di un riscontro; (c) uso del numero di sequenza in emissione.

In Fig. III.19b è illustrato un caso di ri-emissione di trama, analogo a quello sopra trattato, in cui però questa volta la trama perduta non è quella informativa, bensì la trama di riscontro ACK. In questo caso la ri-emissione provoca la ricezione di una doppia copia della trama e quindi un ulteriore errore nel trasferimento delle informazioni.

Come mostrato in precedenza, la ri-emissione di una trama comporta il rischio di duplicazione della stessa in ricezione e quindi la possibilità che la DL-entità ricevente consegni allo strato di rete un'informazione non corretta.

La soluzione di questo problema è la seguente: nelle trame viene inserito un campo che indica il numero d'ordine sequenziale (*numero di sequenza*) in cui è avvenuta l'emissione di ciascuna trama. In questo modo l'entità ricevente può distinguere la trama originaria da quelle duplicate.

L'inconveniente di duplicazione di una trama illustrato in precedenza può essere superato numerando in sequenza le trame emesse (Fig. III.19c). In questo caso, quindi, l'entità ricevente rivela la ricezione di due trame con lo stesso numero di sequenza (il numero 3 nell'esempio) e quindi è in grado di scartare la trama duplicata.

Riassumendo, un DL-protocollo ha a disposizione, per garantire l'affidabilità del trasferimento delle trame, i seguenti meccanismi: 1) l'impiego di riscontri (positivi o negativi); 2) l'uso di richieste di ri-emissione; 3) l'uso di temporizzatori; 4) l'introduzione dei numeri di sequenza delle trame.

In relazione all'utilizzazione dei precedenti meccanismi, possiamo distinguere tre modalità principali per le procedure di recupero degli errori: a) *modalità a riscontro positivo con ri-emissione*; b) *modalità a finestra scorrevole con ri-emissione non selettiva*; c) *modalità a finestra scorrevole con ri-emissione selettiva*.

Nella modalità a riscontro positivo con ri-emissione, una DL-entità può emettere una sola trama alla volta. Per emettere la successiva deve attendere il riscontro positivo dalla DL-entità ricevente. In questo caso, è sufficiente che il campo indicante il numero di sequenza della trama sia composto da una sola cifra binaria (*bit di sequenza*), equivalente ad una numerazione modulo 2.

In caso di ri-emissioni, provocate dall'azzeramento del temporizzatore, il bit di sequenza della trama riemessa ha lo stesso valore di quello della trama originale. La DL-entità ricevente accetta, quindi, solo trame che rechino bit di sequenza dal valore alternato.

Un caso di utilizzazione della modalità a riscontro positivo con ri-emissione è mostrato in Fig. III.20. In questo caso la numerazione delle trame è

indicata da una sola cifra binaria, per cui la ricezione di trame consecutive recanti lo stesso valore del bit di sequenza è indice di duplicazione. Dalla figura si nota come tale meccanismo sia sufficiente a far superare eventi, sia singoli che multipli, in cui vengano perse sia trame informative, sia trame di riscontro.

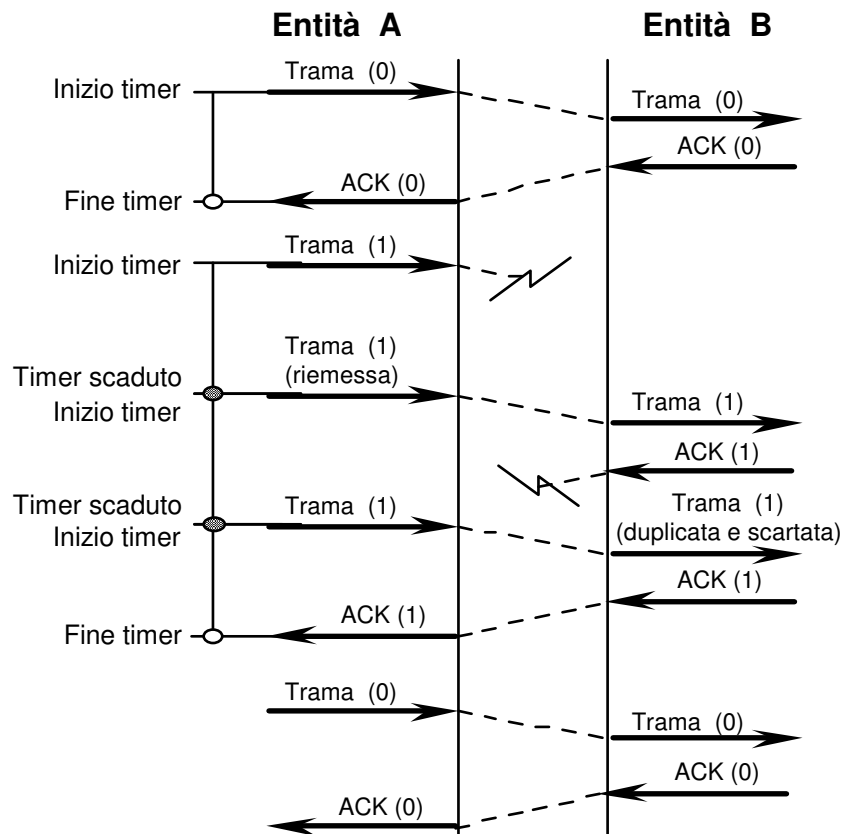


Fig. III.20 - Esempio di utilizzazione della modalità a riscontro positivo con ri-emissione.

Nelle modalità a finestra scorrevole con ri-emissione sia non selettiva (b) che selettiva (c), l'entità emittente ha la possibilità di inviare, senza attendere il riscontro, un numero  $W_T$  di trame, con  $W_T > 1$ . Il valore di  $W_T$  costituisce la larghezza della cosiddetta *finestra di trasmissione*. I limiti inferiore e superiore di tale finestra sono aggiornati progressivamente in base ai riscontri ricevuti e corrispondono ai numeri di sequenza minimo e massimo delle trame che possono essere emesse.

Detto  $L_{inf}$  il valore del limite inferiore, il valore  $L_{sup}$  dell'estremo superiore è uguale a  $L_{inf} + W_T - 1$ .

In Fig. III.21 è riportato un esempio di aggiornamento della finestra di trasmissione in base all'evoluzione della comunicazione tra due DL-entità.

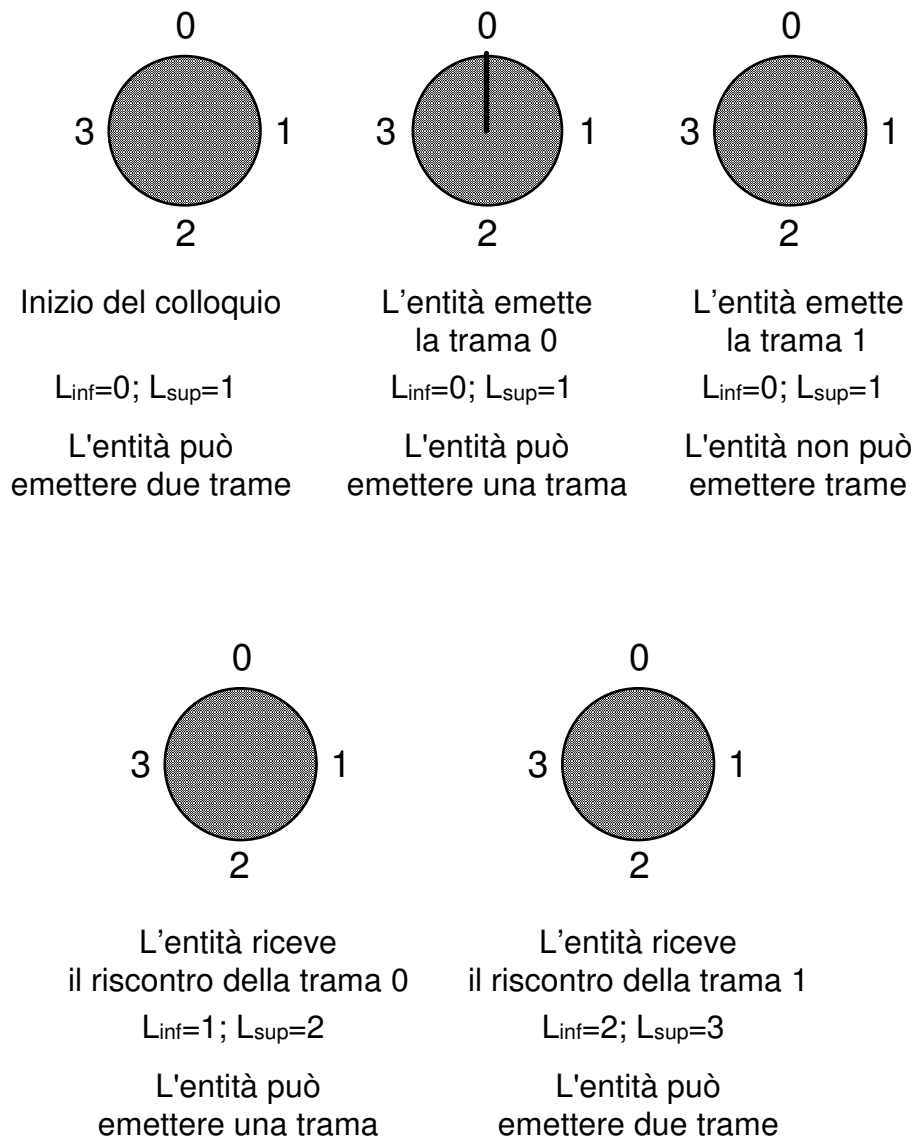


Fig. III.21 - Esempio di aggiornamento della finestra di trasmissione.

Nel caso mostrato, il numero di sequenza delle trame è codificato modulo 4, la finestra ha larghezza  $W_T$  uguale a 2 e, all'inizio della connessione, i limiti inferiore e superiore sono uguali rispettivamente a 0 e 1. All'inizio della fase di trasferimento dei dati, l'entità sorgente emette due trame consecutivamente, esaurendo la finestra. Non appena iniziano ad arrivare le trame ACK, che in questo caso recano il numero di sequenza della trama da riscontrare, il limite inferiore della finestra è progressivamente aggiornato. Questo indica sempre il numero di sequenza della prima trama che non è stata ancora riscontrata o che deve essere ancora emessa.

Si noti che questo meccanismo consente il riscontro di più trame contemporaneamente se alla trama ACK si dà il significato di riscontro di tutte

le trame che hanno un numero di sequenza inferiore o uguale a quello indicato dalla trama ACK stessa.

Nel caso di finestra di trasmissione di larghezza uguale a  $W_T$ , il campo recante il numero di sequenza di una trama deve avere una lunghezza uguale almeno a  $\lceil \log_2(W_T+1) \rceil$  bit, avendo indicato con  $\lceil x \rceil$  il più piccolo intero superiore a  $x$ .

Nel caso di finestra di trasmissione di larghezza uguale 3, il campo di numerazione delle trame deve essere di lunghezza uguale almeno a 2 cifre binarie, equivalenti ad una numerazione modulo 4. Mentre, se si volesse una finestra di trasmissione di 4 trame, sarebbero necessario prevedere un campo di numerazione di lunghezza uguale almeno a 3 cifre binarie (numerazione modulo 8).

Infatti, se si utilizza una numerazione modulo 4, non è possibile emettere più di tre trame senza riscontro. Altrimenti, si potrebbe verificare una ambiguità nel riscontro delle stesse. Supponiamo infatti che in un certo istante il limite inferiore della finestra si uguale a 0. Se la finestra avesse larghezza uguale a 4 trame, l'entità emittente potrebbe emettere le trame numerate da 0 a 3. Se a seguito di questa emissione, arrivasse una trama ACK recante il numero di riscontro 0, non sarebbe possibile stabilire se questo riscontro si riferisca ancora alla trama 3 emessa in una fase precedente e già riscontrata, oppure alla trama 3 attualmente in attesa di riscontro.

Una finestra analoga può essere definita anche per il lato ricezione (*finestra di ricezione*).

In questo caso, i limiti inferiore e superiore della finestra indicano rispettivamente i valori minimo e massimo dei numeri di sequenza delle trame che possono essere accettate in ricezione a prescindere dall'ordine di arrivo.

Se, quindi,  $W_R$  indica la larghezza della finestra di ricezione, il numero massimo di trame che possono essere accettate, anche in presenza alterazioni nella sequenza di ricezione, è uguale a  $W_R - 1$ . L'aggiornamento dei limiti della finestra di ricezione è progressivamente operato all'arrivo delle trame corrispondenti al limite inferiore della finestra.

Detto  $L_{inf}$  il valore del limite inferiore, il valore  $L_{sup}$  del limite superiore è uguale a  $L_{inf} + W_R - 1$ .

Un esempio di funzionamento della finestra di ricezione è mostrato in Fig. III.22, in cui sono state scelte una numerazione modulo 4 delle trame e una larghezza della finestra uguale a 3 trame. Come si può notare, fenomeni di fuori sequenza sono tollerati, mentre l'aggiornamento dei limiti della finestra è

effettuato solo quando la sequenza delle trame ricevute è corretta. Ad esempio, l'aggiornamento del limite inferiore con valore 1 avviene solo all'arrivo della trama 1, anche se prima di essa sono giunte le trame 2 e 3.

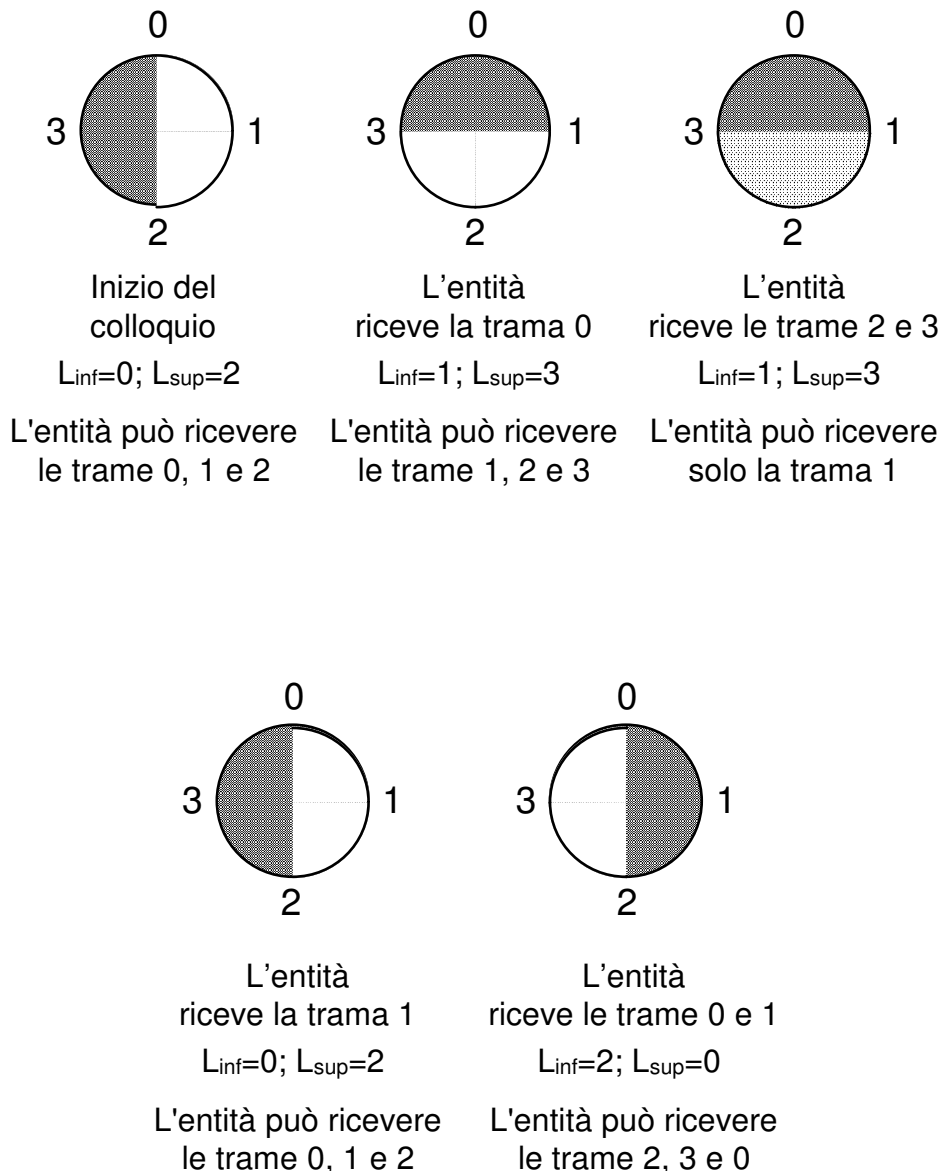


Fig. III.22 - Esempio di aggiornamento della finestra di ricezione.

La differenza tra le due modalità b) e c) consiste nella diversa gestione delle ri-emissioni.

Nella modalità con ri-emissione non selettiva, quando si rivela un'anomalia nella sequenza delle trame ricevute, viene richiesta la ri-emissione della trama mancante e di tutte le successive, a prescindere dal fatto che queste ultime siano state, o saranno, ricevute correttamente. Tale modalità corrisponde ad assumere una finestra di ricezione di larghezza unitaria.

Nel caso di ri-emissione selettiva, invece, la richiesta di ri-emissione riguarda esclusivamente la trama mancante, mentre le successive, se corrette, sono accettate ugualmente. In questo caso la finestra di ricezione ha larghezza maggiore di uno.

In Fig. III.23a è riportata l'evoluzione del protocollo se è utilizzata la modalità a ri-emissione non selettiva. In questo caso, non appena si verifica una situazione di fuori sequenza (trame 4 e 5 arrivate prima della 3), la DL-entità richiede, tramite l'invio di una opportuna trama, la ri-emissione sia della trama 3 che, implicitamente delle trame seguenti. Le trame 4 e 5, anche se arrivate correttamente, sono scartate. Ciò equivale a dire che, in ogni istante, solo la trama con numero di sequenza uguale a quello corrispondente all'estremo inferiore della finestra di ricezione può essere accettata. La larghezza della finestra di ricezione è in questo caso uguale ad 1.

In Fig. III.23b è mostrato il comportamento della modalità a ri-emissione selettiva nello stesso caso di Fig. III.23a. Come è facile rendersi conto, tale modalità evita la ri-emissione delle trame 4 e 5 garantendo una maggiore efficienza della connessione.

La scelta della larghezza della finestra di trasmissione dipende essenzialmente da una scelta tra minore complessità delle DL-entità e migliori prestazioni del protocollo. In alcuni casi, tuttavia, tale scelta è obbligata dalle condizioni operative.

Supponiamo ad esempio una DL-connessione, realizzata su una tratta via satellite, caratterizzata da un ritmo binario di 100 kbit/s e da un ritardo complessivo di propagazione di 300 ms. Se la finestra di trasmissione avesse larghezza uguale a 1 e la trama avesse lunghezza uguale a 1000 cifre binarie, l'emissione di una trama occuperebbe la stazione per 10 ms, a cui seguirebbe un intervallo di tempo minimo, uguale a 610 ms, prima dell'arrivo del riscontro e quindi della possibilità di inviare un'ulteriore trama. Tale tempo è dovuto al doppio tempo di propagazione della tratta via satellite e dal tempo di trasmissione del riscontro. In questo caso l'efficienza di utilizzazione della tratta sarebbe molto bassa, non superiore a circa l'1.5%.

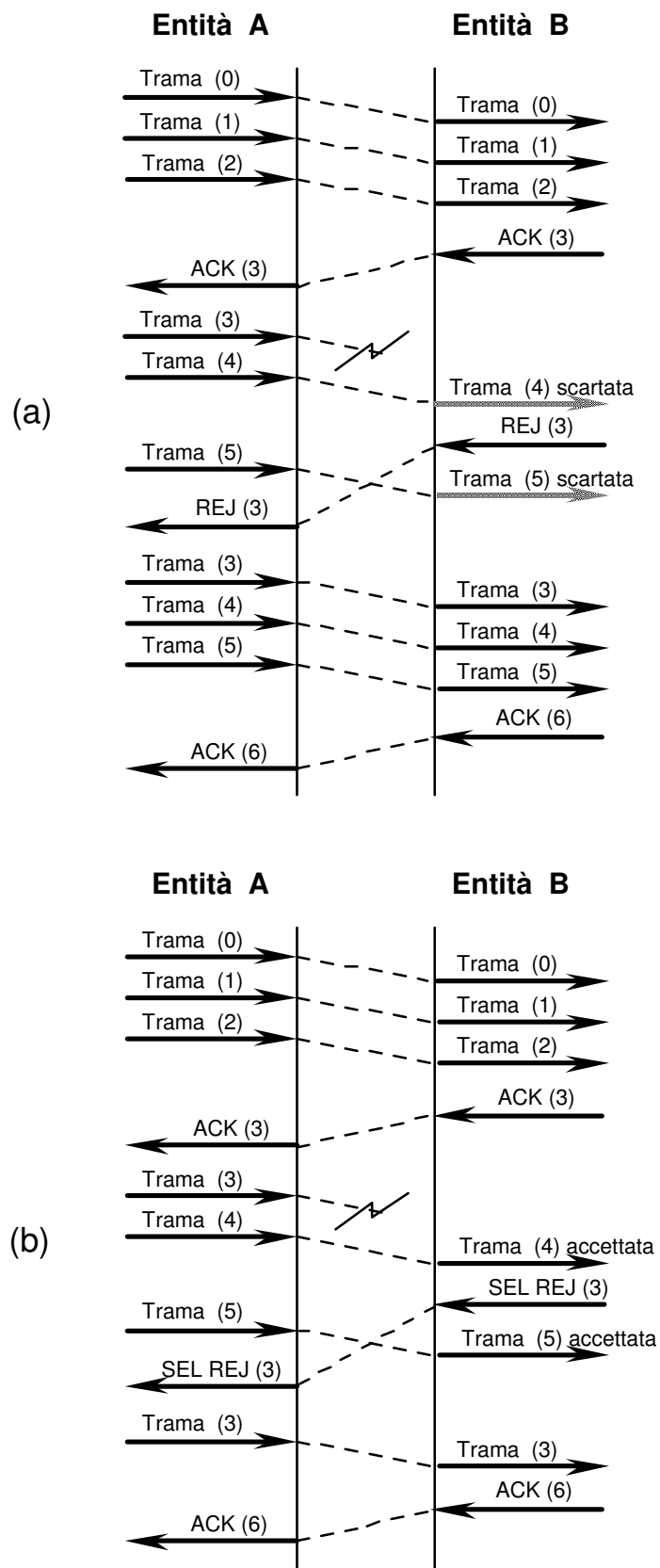


Fig. III.23 - Esempio di utilizzazione della modalità con ri-emissione per il recupero di errore: (a) ri-emissione non selettiva; (b) ri-emissione selettiva.

L'uso di una finestra di trasmissione più larga garantirebbe una maggiore efficienza. Infatti, durante l'attesa del riscontro alla prima trama, una entità potrebbe proseguire nella emissione delle trame successive. Nel caso esaminato, il valore ottimale della larghezza della finestra è uguale a 62. Con tale valore di larghezza di finestra ed in condizioni ottimali di funzionamento, l'entità emittente ha la possibilità di emettere continuativamente trame, senza tempi morti dovuti all'attesa di riscontri.

#### III.3.4 *Controllo di flusso.*

La funzione di controllo di flusso ha il compito di assicurare che, nel caso in cui il ritmo d'arrivo delle trame sia superiore alla capacità di elaborazione della DL-entità ricevente, non si verifichino perdite di informazione a causa della saturazione delle sue risorse di memorizzazione.

La soluzione classica a questo problema è quella di realizzare un meccanismo di controreazione tale che, in caso di necessità, l'entità ricevente possa imporre a quella emittente di rallentare l'emissione delle trame.

Sono possibili vari schemi di controllo di flusso. La maggior parte di questi utilizza le già discusse regole protocollari, definite per consentire il recupero in caso di errori.

Abbiamo già visto che, in accordo a queste regole, un'entità non può emettere trame se prima non ne ha ricevuto il relativo permesso, implicito o esplicito, da parte dell'entità ricevente. Tale principio può essere quindi utilizzato anche per regolare il flusso delle trame.

Ad esempio, il meccanismo a finestra di trasmissione, discusso nel paragrafo precedente, rappresenta anche un metodo di controllo di flusso. Infatti, l'entità ricevente ha la possibilità di regolare l'emissione dei riscontri in modo da ottenere il voluto comportamento da parte dell'entità emittente.

Ulteriori dettagli a riguardo saranno presentati nel par. IV.6, con riferimento a un controllo di flusso operato in una connessione di rete.

#### III.3.5 *Gestione del collegamento.*

La funzione di gestione del collegamento ha lo scopo di fornire alle DL-entità i mezzi per instaurare e per abbattere una DL-connessione ed, eventualmente, per ri-inizializzarla in caso di gravi malfunzionamenti.

Nell'ambito dell'evoluzione di una DL-connessione, una DL-entità può svolgere due tipi di funzioni: da un lato le funzioni di *primario* e di *secondario*, dall'altro quelle di *sorgente* e di *collettore*. Queste ultime hanno un ruolo ovvio; anzi si può naturalmente accettare che esse debbano essere ambedue presenti in

una DL-entità, che operi nell'ambito di una connessione su cui l'informazione è trasferita in modo bidirezionale.

Non altrettanto si può affermare per le funzioni di primario e di secondario. La prima di queste è quella che ha la responsabilità del trasferimento di informazione verso una DL-entità che svolge la funzione di secondario. Quest'ultima deve quindi ricevere *comandi* dalla funzione di primario e deve fornire *risposte* in relazione alle azioni richieste dai comandi.

Per effetto delle diverse possibili combinazioni tra queste coppie di funzioni, un trasferimento di informazione da una sorgente a un collettore può avvenire, in tre fasi, secondo due modalità principali, e cioè *a seguito di comandi* ovvero *a seguito di risposte*.

La prima modalità si verifica quando, all'estremità A della connessione, la funzione di primario si associa a quella di sorgente, mentre all'estremità B la funzione di secondario è associata a quella di collettore. In questo caso (Fig. III.24a) il trasferimento di informazione si attua nel modo seguente: l'estremità A invia alla B (prima fase) un *invito a ricevere* (selecting) e fa quindi seguire (seconda fase) l'emissione delle sue unità di dati; l'interazione si chiude con l'emissione (terza fase), da parte dell'estremità B, di un riscontro di ricezione. Il trasferimento avviene quindi a seguito di un comando.

La seconda modalità prevede invece che all'estremità A la funzione di primario sia associata a quella di collettore e che all'estremità B si abbia un'associazione tra le funzioni di secondario e di sorgente. In queste condizioni (Fig. III.24b), la prima fase del trasferimento, detta anche *fase di interrogazione*, si attua con l'emissione, da parte di A, di un *invito ad emettere* (polling) rivolto all'estremità B; questa, se ha informazioni da trasferire ad A, emette le sue unità di dati (seconda fase); l'interazione si chiude con un riscontro di ricezione da parte di A (terza fase). Il trasferimento si realizza quindi a seguito di una risposta.

Queste due modalità di trasferimento possono essere impiegate singolarmente o in forma combinata, per ognuno dei due versi della connessione. Ciò dipende dalla struttura delle apparecchiature che sono interconnesse, chiamate nel seguito *stazioni*, e dal tipo di procedura di trasferimento impiegata.

Circa la struttura delle stazioni, esistono soluzioni in cui si ha un'*associazione permanente* delle funzioni primaria o secondaria da un lato e quelle di sorgente e di collettore dall'altro, ferma restando la presenza in ogni stazione, di ambedue queste ultime. A tale riguardo esistono due alternative significative.

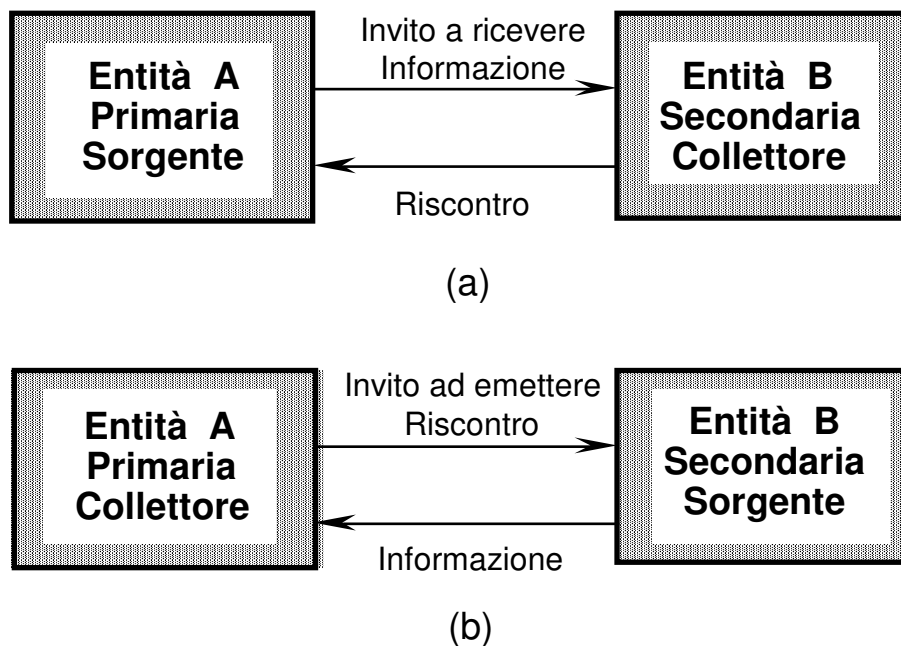


Fig. III.24 - Controllo del trasferimento di informazione: (a) con invito a ricevere; (b) con invito a emettere.

Nella prima (alternativa 1), una stazione include solo la funzione primaria ovvero quella secondaria e ognuna di queste è associata, permanentemente e separatamente, con la funzione di sorgente o con quella di collettore; si parla allora di *stazione primaria* o di *stazione secondaria* a seconda dei casi.

Nella seconda (alternativa 2), in ogni stazione la funzione di sorgente è associata permanentemente alla funzione primaria, mentre quella di collettore è accoppiata con quella secondaria.

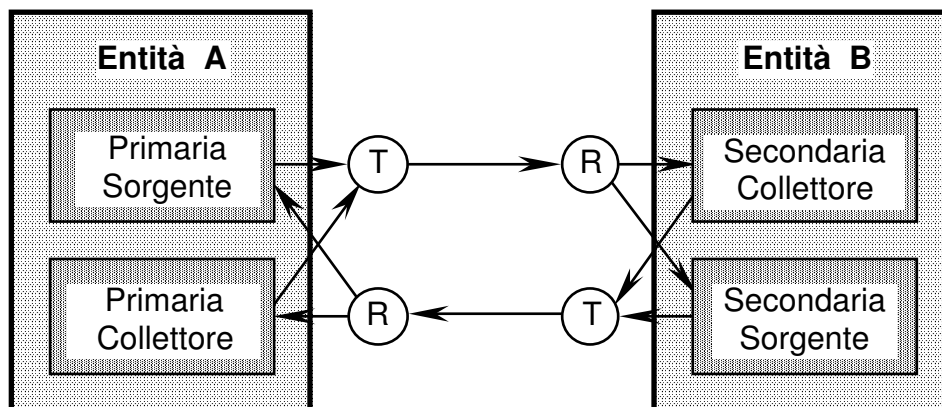
Esiste anche una possibilità ulteriore rispetto alle alternative 1 e 2. Questa riguarda le cosiddette *stazioni combinate*, in cui esistono sia la funzione primaria che quella secondaria (come nell'alternativa 2), ma ognuna di queste è dissociata dalle funzioni di sorgente e di collettore. Ciò consente alle funzioni primaria e secondaria di condividere le altre due funzioni.

Riguardo poi al tipo di procedura, si distinguono *procedure bilanciate* e *procedure sbilanciate*. Nelle seconde il trasferimento dei dati in un verso della connessione avviene solo a seguito di comandi, secondo lo schema di Fig. III.24a, ovvero solo a seguito di risposte, come avviene invece nel caso illustrato in Fig. III.24b.

Per concludere, esaminiamo tre soluzioni di DL-connessione che possono emergere dall'applicazione di questi concetti generali.

La prima soluzione è illustrata in Fig. III.25, ove, alle due estremità A e B della DL-connessione sono presenti stazioni non combinate, che sono

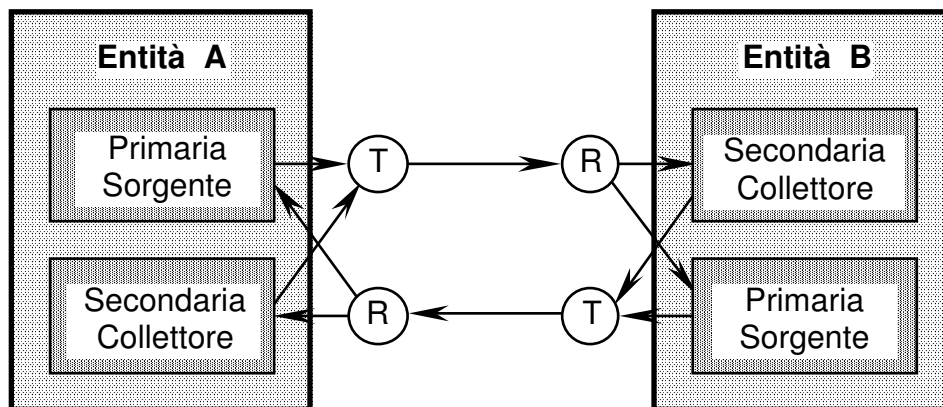
strutturate secondo l'alternativa 1; ma, mentre all'estremità A è presente una stazione con sola funzione primaria (stazione primaria), all'estremità B si ha la sola funzione secondaria (stazione secondaria). Si applica allora, per un verso della connessione, lo schema di Fig. III.24a e, per l'altro, quello di Fig. III.24b. Conseguentemente il trasferimento in un verso (da A B) avviene a seguito di comandi e quello nell'altro verso (da B ad A) solo a seguito di risposte. La procedura di trasferimento è quindi di tipo sbilanciato. Inoltre la configurazione di connessione è *asimmetrica*, data la disomogeneità funzionale delle stazioni connesse; essa consente sia topologie punto-punto che punto-multipunto.



E : Emittitore; R : Ricevitore

Fig. III.25 - Configurazione asimmetrica con procedura sbilanciata.

Una seconda soluzione è quella presentata in Fig. III.26, ove alle due estremità A e B sono presenti stazioni non combinate che sono strutturate secondo l'alternativa 2. Pertanto, per ambedue i versi della connessione, si applica lo schema di Fig. III.24a e il trasferimento avviene solo a seguito di comandi. La procedura di trasferimento è quindi di tipo sbilanciato, mentre la configurazione di connessione è *simmetrica*, dato che alle due estremità sono presenti stazioni funzionalmente omogenee. Tale configurazione è quindi impiegabile, a differenza della precedente, solo in topologie punto-punto ovvero in topologie punto-multipunto in cui siano contemporaneamente attive solo due stazioni.



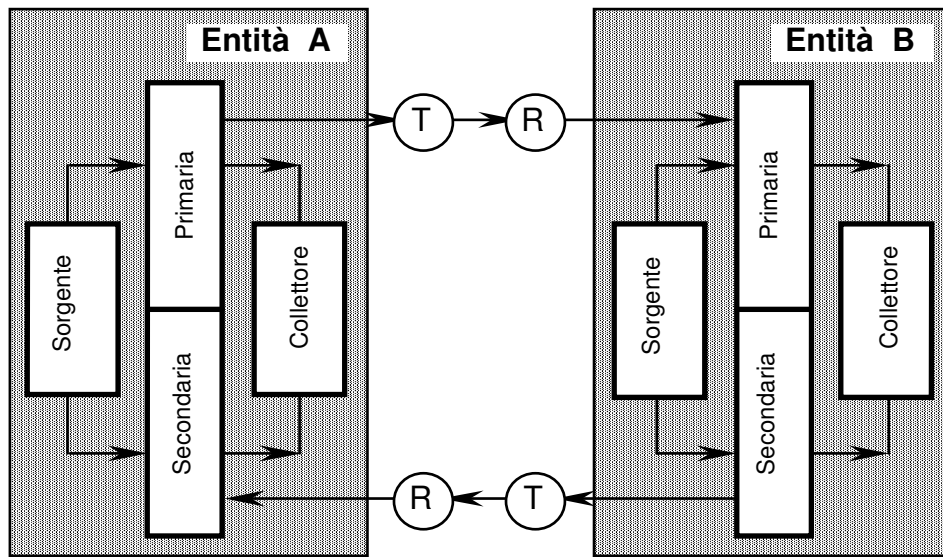
E: Emittitore; R : Ricevitore

Fig. III.26 - Configurazione simmetrica con procedura sbilanciata.

Infine in Fig. III.27 è presentata la terza soluzione, e cioè il caso di una connessione tra due stazioni combinate, che è utilizzabile solo in topologie punto-punto. Qui il trasferimento di informazione, per ogni verso, avviene indifferentemente a seguito di comandi e di risposte. La relativa procedura è quindi di tipo bilanciato.

Alle soluzioni ora illustrate corrispondono tre possibili modalità di comunicazione:

- il *modo di risposta normale* (NRM - Normal Response Mode), che impiega la configurazione di collegamento mostrata in Fig. III.25; in questo caso la stazione primaria governa lo scambio delle trame mediante l'instaurazione e l'abbattimento della connessione ed il controllo del flusso di trame da e verso le stazioni secondarie; queste ultime possono iniziare a emettere solo dopo aver ricevuto dalla prima una esplicita autorizzazione;
- il *modo di risposta asincrono* (ARM - Asynchronous Response Mode), che impiega la configurazione di collegamento illustrata in Fig. III.26; in questo caso ognuna delle due stazioni può instaurare ed abbattere la connessione; successivamente ognuna di esse può emettere trame senza attendere l'autorizzazione dall'altra;
- il *modo bilanciato* (ABM - Asynchronous Balanced Mode), che impiega la configurazione di collegamento illustrata in Fig. III.27; in questo modo ogni stazione combinata può instaurare ed abbattere la connessione; può anche emettere e ricevere trame nell'ambito di sue iniziative, non legate al comportamento dell'altra stazione.



E : Emittitore; R : Ricevitore

Fig. III.27 - Procedura bilanciata.

Un'applicazione delle modalità di colloquio NRM è costituita dalle reti a topologia stellare governate da protocolli d'accesso del tipo ad interrogazione e risposta (*polling*). In tal caso, la stazione primaria, posta al centro stella (normalmente un computer di medie-grandi proporzioni) interroga ciclicamente le stazioni secondarie (i terminali) per richiedere l'eventuale invio di dati.

Nelle reti a lunga distanza a commutazione di pacchetto, il colloquio tra DL-entità poste all'interno dei nodi di commutazione avviene normalmente in accordo alla modalità ABM. Poiché esiste in questo caso la completa simmetria nella connessione, entrambe le entità hanno la possibilità di eseguire qualsiasi funzione protocollare in qualsiasi istante.

È evidente che in collegamenti punto-punto i modi ARM e ABM sono più efficienti, poiché non richiedono procedure di interrogazione.

Si possono infine distinguere tre *modi di utilizzazione* di un collegamento; a) *unidirezionale*, in cui l'informazione è trasferita solo da una stazione A ad una stazione B e non viceversa; b) *bidirezionale alternato*, in cui l'informazione è trasferita sia da A a B, che da B ad A, ma non contemporaneamente; c) *bidirezionale simultaneo*, in cui le informazioni sono trasferite nelle due direzioni anche contemporaneamente. La corrispondenza tra modalità di trasmissione dei dati a livello fisico e modi di utilizzazione del collegamento è descritta in Tab. III.3.

Modalità di trasmissione dati	Modalità di utilizzazione del collegamento		
	Unidirezionale	Bidirezionale alternata	Bidirezionale simultanea
Semplice	X		
Semiduplice	X	X	
Duplice	X	X	X

Tab. III.3 - Corrispondenza tra modalità di trasmissione dei dati e modalità di utilizzazione del collegamento.

### III.4 II LAP-B

Il protocollo *X.25 di livello 2* (Link Access Protocol - Balanced, LAP-B) è il protocollo di strato di collegamento standardizzato dal CCITT per l'utilizzazione nelle reti a pacchetto per il colloquio tra DTE e DCE. Le procedure dell'*X.25 di livello 2* sono derivate da quelle utilizzate nel protocollo *HDLC* (High-level Data Link Control) standardizzato dall'ISO.

In questo paragrafo saranno descritte le funzionalità del protocollo *X.25 di livello 2*, cercando di individuare, per ogni funzione, quale modalità esecutiva tra quelle illustrate nei paragrafi precedenti è utilizzata. Per i dettagli procedurali si rimanda il lettore alla consultazione della specifica raccomandazione del CCITT.

L'*X.25 di livello 2* è un protocollo orientato al bit, che usa il formato di trama indicato in Fig. III.28. Esso si compone dei seguenti campi:

- *delimitatore* (*F* - Flag): sequenza di 8 cifre binarie nella configurazione (01111110), con funzione di sincronizzazione e di delimitazione iniziale e finale della trama;
- *indirizzo* (*A* - Address): sequenza di 8 cifre binarie, che discrimina un comando da una risposta; delle 256 configurazioni possibili di questo campo ne vengono utilizzate solo due; questa ridondanza è da mettere in relazione alle esigenze del protocollo HDLC (da cui l'*X.25* deriva), dove è previsto l'indirizzamento di una molteplicità di stazioni secondarie nelle configurazioni punto-multipunto;
- *controllo* (*C* - Control): sequenza di 8 cifre binarie (nella versione di base), che specifica il tipo di trame: *informativa* (*I* - Information), *di supervisione* (*S* - Supervisory) o *non numerata* (*U* - Unnumbered);
- *informazione* (*I* - Information): sequenza di un numero variabile di cifre

binarie che costituisce l'informazione di utente (pacchetto), così come viene offerta e ricevuta dallo strato di rete; in questo campo viene inserito un solo pacchetto;

- *sequenza di controllo della trama (FCS - Frame Check Sequence):* sequenza di 16 cifre binarie, utilizzata per la funzione di rivelazione di errore.

Delimitatore	Indirizzo	Controllo	Informazione	Sequenza di controllo della trama	Delimitatore
F 01111110	A 8 bit	C 8 bit	I N bit	FCS 16 bit	F 01111110

Fig. III.28 - Formato della trama del protocollo X.25 di livello 2.

La funzione di delimitazione delle trame è realizzata tramite i campi delimitatori; è quindi necessaria la funzione di riempimento di bit (cfr. par.. III.3) per assicurare la trasparenza del contenuto informativo della trama.

La funzione di rivelazione d'errore è attuata mediante l'applicazione di codici polinomiali attraverso il campo FCS. Tale campo, come discusso in par.. III.3, contiene, in forma binaria, il resto  $R(x)$  della divisione, tra il polinomio  $P(x)$  associato alla stringa delle cifre binarie costituenti i campi A, C ed I ed un polinomio generatore  $G(x)$  comune alle entità emittente e ricevente. Il polinomio generatore utilizzato ha l'espressione  $x^{16}+x^{12}+x^5+x^0$ . Una trama è ritenuta corretta se il resto  $R'(x)$  della divisione, tra il polinomio individuato dal valore ricevuto delle cifre binarie della trama ed il polinomio generatore, è uguale al polinomio  $R(x)$  inserito nel campo FCS. Se ciò non accade la trama viene scartata.

La modalità utilizzata per l'attuazione delle funzioni di recupero in caso di errori e di controllo di flusso è quella a finestra scorrevole con ri-emissione non selettiva.

La larghezza della finestra di trasmissione può essere al massimo uguale a 7 nella versione di base. Mentre la finestra di ricezione è sempre di larghezza unitaria. La richiesta di ri-emissione, effettuata tramite un apposita trama di supervisione, avviene quindi quando l'entità ricevente rileva una situazione di fuori sequenza nelle trame ricevute corrette. Tale richiesta di ri-emissione interessa, implicitamente, tutte le trame il cui numero di sequenza è maggiore di quello indicato nella richiesta di ri-emissione.

Per quanto riguarda la gestione del collegamento, sono possibili le modalità ARM e ABM. L'attivazione di una tra le due modalità è operata

mediante l'invio di una specifica trama non numerata.

La codifica del campo di controllo per le trame informative, di supervisione e non numerate è riportata in Tab. III.4.

Tipo di trama	Comandi	Risposte	Codifica							
			1	2	3	4	5	6	7	8
Informativa	I		0	N(s)			P	N(r)		
Supervisione	RR	RR	1	0	0	0	P/F	N(r)		
	RNR	RNR	1	0	1	0	P/F	N(r)		
	REJ	REJ	1	0	0	1	P/F	N(r)		
Non numerata	SARM	DM	1	1	1	1	P/F	0	0	0
	SABM		1	1	1	1	P	1	0	0
	DISC		1	1	0	0	P	0	1	0
		UA	1	1	0	0	F	1	1	0
		FRMR	1	1	1	0	F	1	1	0

Tab. III.4 - Tipi di trame definite nel protocollo X.25 di livello 2.

Una trama informativa contiene, nel campo C, i due numeri di sequenza  $N(s)$  e  $N(r)$ , che indicano, rispettivamente, il numero (modulo 8) di sequenza in trasmissione ed il numero (modulo 8) di sequenza in ricezione.  $N(r)$  ha il significato di numero di sequenza della trama che la stazione emittente si aspetta di ricevere.

È il caso di sottolineare che, in questo protocollo è possibile inviare i riscontri delle trame ricevute, oltre che con le apposite trame di supervisione, anche attraverso le trame informative inviate nella direzione opposta della connessione. Questo è quindi lo scopo del numero  $N(r)$  contenuto in esse.

Sono definiti tre tipi di trame di supervisione: 1) *pronto a ricevere* (RR - Receive Ready); 2) *rigetto* (REJ - Reject); 3) *non pronto a ricevere* (RNR - Receive Not Ready).

Le trame RR e REJ hanno funzioni, rispettivamente, di riscontro positivo e di riscontro negativo delle trame I, mentre la trama RNR indica la temporanea indisponibilità a ricevere da parte della stazione che la emette.

Come nelle trame informative, il numero  $N(r)$ , contenuto nelle trame di supervisione indica il numero di sequenza della trama che l'entità si aspetta di ricevere e quindi, implicitamente, ha il significato di riscontro della ricezione di tutte le trame fino a quella numerata con  $N(r)-1$ .

Le trame non numerate SARM (Set ARM) e SABM (Set ABM) svolgono la

funzione di instaurazione della connessione tra DTE e DCE, rispettivamente secondo la modalità ARM o ABM. La trama *DISC* (Disconnect) è invece utilizzata per abbattere la connessione.

Le risposte *UA* (Unnumbered Acknowledgement) e *DM* (Disconnect Mode) indicano, rispettivamente, l'accettazione e il rifiuto della connessione, mentre la trama *CMDR/FRMR* (Command Reject/Frame Reject) segnala la non accettazione di una trama. Le trame *DM* e *FRMR* sono anche utilizzate per ottenere una re-inizializzazione della connessione (*funzione di re-inizializzazione*).

Il bit *P/F* (Poll/Final) consente di associare logicamente una trama di risposta (in cui il bit è denominato *final*) alla trama di comando ad essa relativa (in cui il bit è denominato *poll*).

Un diagramma di stato semplificato delle fasi di instaurazione e di abbattimento di una DL-connessione, riferita al modo bilanciato del protocollo X.25 di livello 2, è riportato in Fig. III.29.

#### *PROCEDURE DI INSTAURAZIONE E DI ABBATTIMENTO DI UNA DL-CONNESSIONE*

Dall'esame della Fig. III.29 emerge che la sequenza di trame che sono scambiate all'atto dell'instaurazione di una DL-connessione, con modalità ABM, è la seguente. L'entità, che inizia la procedura emette una trama *SABM* e si pone nello stato di richiesta di connessione. Alla ricezione di una trama *UA*, la connessione entra direttamente nella fase di trasferimento dei dati.

La Fig. III.29 mette anche in evidenza che le trame *FRMR* e *DM* sono utilizzate per realizzare anche la funzione di re-inizializzazione della connessione in caso di gravi situazioni di errore, ad esempio nel caso di ricezione di una trama non permessa in un certo stato o di fallimento delle procedure di ri-emissione.

In questi casi, l'emissione di una delle due trame costringe a reinizializzare la connessione attraverso la ri-emissione di una trama *SABM*.

La Fig. III.30 mostra invece alcuni esempi della procedura di trasferimento di dati utilizzata nel protocollo X.25 di livello 2.

In tale figura, si fa riferimento alla sequenza temporale di trame in emissione e in ricezione ad una interfaccia A-B, nell'ipotesi che il collegamento sia già stato instaurato e che il ritardo di trattamento e di propagazione sia nullo.

Nel simbolismo adottato in questa figura, nell'interno di ogni riquadro rappresentante una trama, sono indicati il tipo della trama (*I*, *RR*, *REJ*), i numeri di sequenza  $N(s)$  e  $N(r)$  per le trame informative, nonché il numero  $N(r)$  per le sole trame di supervisione.

Il caso di Fig. III.30a si riferisce a un trasferimento senza errori di trasmissione e mette in evidenza l'uso di trame *RR* per inviare riscontri positivi in assenza di trame *I* da emettere.

La Fig. III.30b si riferisce invece al caso in cui, nel trasferimento da A a B, una trama *I* (la trama "I,4,1") è colpita da errore. A seguito della ricezione della trama "I,5,3", la stazione B rivela la situazione di fuori sequenza, scarta le trame ricevute fuori sequenza (le trame "I,5,3" e "I,6,3"), richiede la ri-emissione con una trama *REJ* (la trama "REJ, 4") ed esce dalla

condizione di fuori sequenza alla ricezione della prima trama I richiesta (la trama "I,4,3").

Infine, la Fig. III.30c riguarda il caso di errori che colpiscono trame di tipo S. In particolare viene esaminato il caso in cui la trama "RR,3", inviata da B ad A, viene rivelata errata in A e quindi scartata. L'esempio mette in evidenza che, come criterio generale di funzionamento della procedura di trasferimento, la stazione emittente, dopo l'invio di ciascuna trama I, attiva un temporizzatore  $T_I$  entro cui si aspetta la ricezione del riscontro di tale trama. Quando tale intervallo è concluso, la trama I è nuovamente emessa. Nell'esempio la trama "I,2,0" viene emessa due volte dalla stazione A, determinando pertanto una situazione di *duplicazione*, che viene recuperata con l'invio, da parte della stazione B, della trama "REJ, 3".

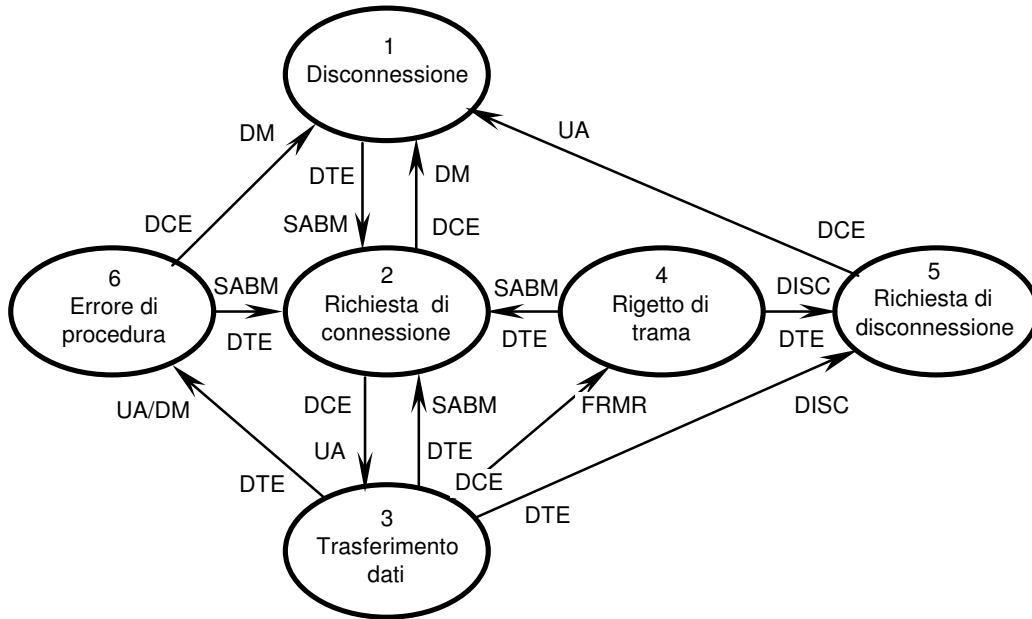


Fig. III.29 - Diagramma semplificato di transizione di stato delle procedure di instaurazione e di abbattimento di una DL-connezione mediante il protocollo X.25 di livello 2.

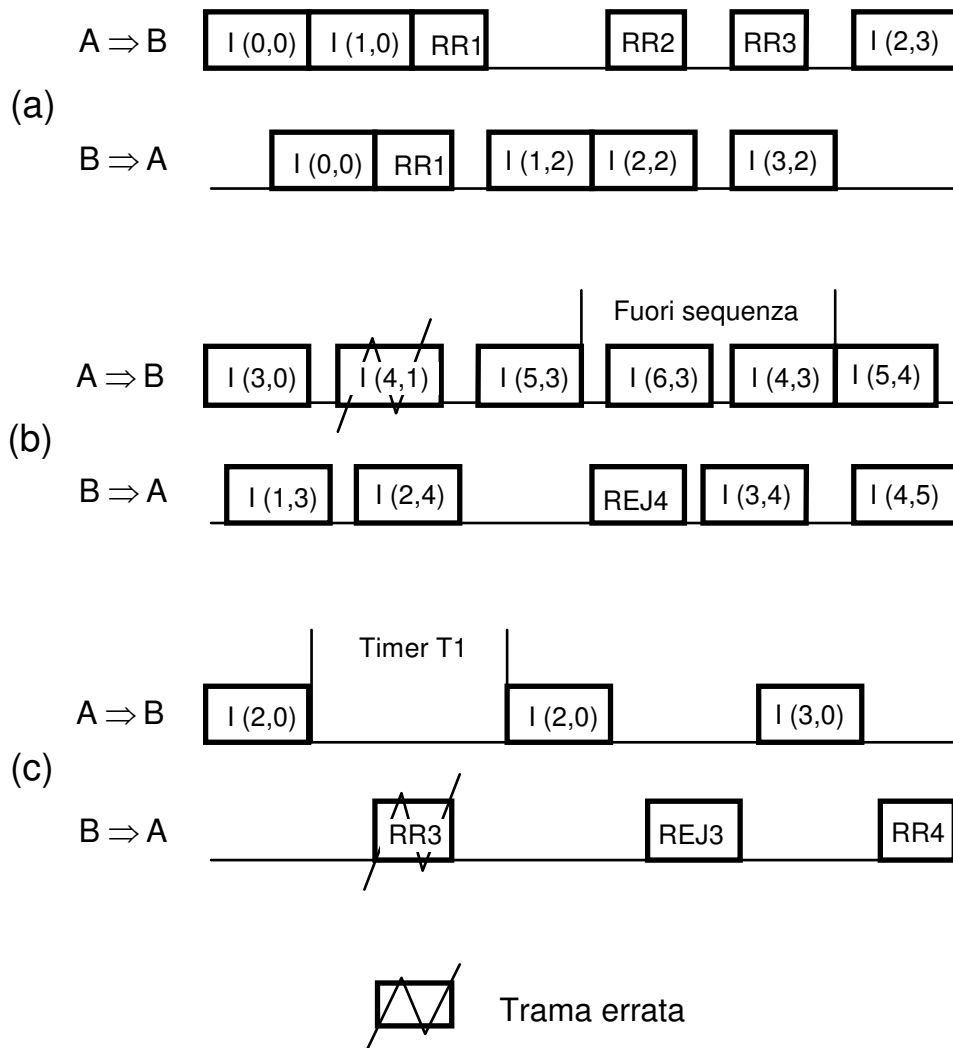


Fig. III.30 - Esempi di applicazione della procedura di recupero degli errori nel protocollo X.25 di livello 2, con riferimento a: (a) trasferimento corretto delle trame; (b) errore su una trama informativa; (c) errore su un riscontro.



## IV STRATO DI RETE

La trattazione in questo capitolo, dedicato allo strato di rete, è suddivisa in più parti. Dapprima si affrontano le caratteristiche del servizio di strato di rete (par. IV.1). Quindi si introducono alcune funzioni che sono tipicamente svolte in questo strato (par. IV.2, IV.3 e IV.4). Il par. IV.5 è dedicato alle funzionalità del servizio di strato di rete. Infine nel par. IV.6 si introduce un protocollo di strato di rete avente significatività in contesti specifici (protocollo X.25 di livello 3); un ulteriore protocollo di strato di rete (il protocollo IP) è descritto nel par. **Errore. L'origine riferimento non è stata trovata.**

### IV.1 Caratteristiche del servizio

Secondo il modello OSI, il servizio dello strato di rete (*servizio di rete o N-servizio*), fornisce, fondamentalmente, il trasferimento trasparente di dati tra entità di trasporto (*utenti dell'N-servizio o NS-utenti*). Esso rende invisibili a questi NS-utenti le modalità secondo cui le risorse di supporto sono utilizzate per realizzare il trasferimento.

In particolare l'N-servizio assicura:

- l'*indipendenza dai sottostanti mezzi di comunicazione*, anche quando la sorgente e la destinazione del trasferimento sono in reti diverse, ognuna delle quali offre servizi dissimili per caratteristiche e per prestazioni;
- il *trasferimento da estremo e estremo*, e cioè tra NS-utenti residenti in sistemi terminali;
- la *trasparenza dell'informazione trasferita*, per ciò che riguarda sia l'informazione d'utente che quella di controllo; cioè la struttura (formato e codice) e il contenuto dei dati trattati sono determinati esclusivamente dagli strati superiori e non sono quindi in alcun modo vincolati dall'N-servizio;
- la *scelta della qualità di servizio* (Quality Of Service o *QOS*) da parte degli NS-utenti, che ne possono concordare le caratteristiche mediante opportuni parametri;
- l'*indirizzamento degli NS-utenti*, che consente a questi ultimi di riferirsi l'un l'altro in modo univoco.

Si tratta, come è evidente, di caratteristiche peculiari dello strato di rete, che ben differenziano questo strato rispetto ai due strati sottostanti. L'N-strato è infatti il primo (salendo nella gerarchia OSI) in cui si pone il problema di un *trasferimento da estremo a estremo*, con possibili attraversamenti di reti diverse, quali reti in area locale (LAN) e in area geografica (WAN).

Queste reti sono normalmente di tecnologie profondamente disomogenee (anche in relazione ai tempi e agli ambienti in cui si sono sviluppate) e presentano, in generale, prestazioni quanto mai diversificate. È quindi spesso assai arduo conseguire condizioni di *inter-funzionamento*, che assicurino la QOS desiderata per specifiche applicazioni.

Le linee guida per risolvere questo problema sono: (i) l'N-servizio dovrebbe essere indipendente dalla tecnologia della rete o delle reti che ne consentono la fornitura; (ii) lo strato di trasporto dovrebbe essere opportunamente "schermato" nei confronti del numero, del tipo e della topologia delle reti da utilizzare; (iii) gli indirizzi, che identificano gli utenti del servizio (e cioè gli *indirizzi di rete*) e che vengono resi disponibili allo strato di trasporto, dovrebbero corrispondere a un *piano di numerazione* uniforme, anche nell'ambito di LAN e WAN.

Sempre in un ambito OSI, sono previsti due tipi di N-servizio, quello *orientato alla connessione* e quello *senza connessione*. Le differenze concettuali tra questi due approcci sono state già sottolineate nel par. II.1 per il PH-servizio e nel par. III.1 per il DL-servizio. Per questi due casi la distinzione tra i due approcci non ha significato (come nel caso dello strato fisico) o riveste una importanza marginale (come nel caso dello strato di collegamento).

Invece, nella definizione dell'N-servizio, il problema della scelta tra le due alternative acquista una rilevanza particolare, dato che è proprio l'N-servizio a porre in corrispondenza *diretta* gli utenti interessati, secondo le loro esigenze applicative, a un trasferimento di informazione con il rispetto di opportuni requisiti prestazionali.

I vantaggi e gli svantaggi dei due approcci sono stati in passato oggetto di attento esame e di aspre dispute negli ambienti delle telecomunicazioni e dell'informatica, e, più in particolare, in quelli dei gestori delle reti e degli utenti di sistemi di elaborazione: i primi a favore di un servizio orientato alla connessione e i secondi a sostegno del servizio senza connessione.

La discussione ha messo in evidenza che le due alternative di servizio hanno caratteristiche tra loro complementari e che è auspicabile una via intermedia. Ad esempio, nel caso in cui si adotti la soluzione orientata alla connessione, nell'ambito di questa dovrebbe essere reso disponibile un servizio che emuli le caratteristiche più attraenti della soluzione senza connessione.

## IV.2 Controllo del traffico

Una fenomenologia tipica che si riscontra in una rete di telecomunicazione

è legata agli effetti del traffico sulle prestazioni della rete e delle sue parti componenti. Ciò è legato al largo impiego di risorse condivise e ai conseguenti fenomeni di congestione, che si possono manifestare sia nelle strategie di pre-assegnazione individuale o collettiva (*congestione di pre-assegnazione*), sia nel corso dell'evoluzione di attività di utilizzazione, che agiscono nell'ambito di strategie con assegnazione a domanda o con pre-assegnazione collettiva (*congestione di utilizzazione*).

Nei fenomeni di congestione di pre-assegnazione o di utilizzazione si può osservare che, al di sopra di un carico-limite (*soglia di sovraccarico*), le prestazioni dell'insieme peggiorano bruscamente. Ad esempio, se le contese di pre-assegnazione sono risolte con modalità a perdita, all'aumentare della domanda di pre-assegnazione al di sopra di questa soglia, si verifica un brusco aumento della probabilità di rifiuto. Se si fa invece riferimento a contese di utilizzazione gestite con modalità ad attesa, al crescere del carico offerto, il sistema risponde con un rapido incremento del valor medio dei tempi di attesa e con una diminuzione, altrettanto brusca, della sua portata.

Per evitare condizioni di questo tipo, che possono condurre rapidamente al completo collasso di intere sezioni di rete, occorre prevedere opportune forme di controllo del traffico, che consentano di filtrare le domande in modo che: 1) l'insieme operi con una intensità di traffico offerto che sia sempre inferiore alla soglia di sovraccarico; 2) le risorse componenti siano utilizzate in modo efficiente. Notiamo che questa seconda condizione è in parziale contrasto con la prima, dato che impone di lavorare in prossimità della soglia.

Come è facile intuire, il problema ora delineato è normalmente di soluzione molto complessa. Ci limiteremo quindi a illustrare qui, a titolo d'esempio, solo alcune linee-guida oggi seguite nell'ambito delle reti per dati a pacchetto.

In questo caso sono previste due forme di controllo di traffico, e cioè quelle di tipo *preventivo* e quelle di tipo *reattivo*. Nel controllo di tipo preventivo, che è applicabile solo a reti con trans-servizio orientato alla connessione, è previsto un meccanismo, detto di *controllo di accettazione di chiamata*. In base a questo una chiamata virtuale viene accettata solo quando sono disponibili sufficienti risorse da pre-assegnare virtualmente in modo che siano soddisfatte due condizioni: a) la nuova chiamata possa usufruire di una adeguata qualità di servizio; b) le chiamate già instaurate e tuttora in corso continuino ad evolvere con la qualità di servizio concordata al momento della loro accettazione.

La gestione delle situazioni di contesa è effettuata normalmente con

modalità a perdita e l'accettazione di chiamata è operata, almeno nel caso qui considerato delle reti per dati a pacchetto, con assegnazione a domanda media.

Nei controlli di tipo reattivo, quali sono applicati almeno nelle reti per dati a pacchetto, sono utilizzati, durante la fase di trasferimento dell'informazione, meccanismi *a retro-pressione*, tra i quali si possono citare le tecniche di *controllo di flusso*. Tale controllo, che è normalmente gestito con modalità ad attesa, viene effettuato attraverso il rilascio, a cura della parte ricevente e nei confronti di quella emittente, di opportune autorizzazioni relative al numero massimo di unità informative che possono essere emesse e/o ricevute. Ciò consente di ridurre o di eliminare gli effetti delle singole cause di congestione di utilizzazione, all'atto in cui si presentano condizioni di carico elevato in certe parti della rete.

Un esempio di controllo reattivo è offerto dalla tecnica di *controllo di flusso a finestra*, che sarà descritta con riferimento al protocollo X.25. Tale tecnica si realizza, per ciascun verso di trasferimento, attraverso l'adozione di due dispositivi, l'uno posto nella estremità emittente (*finestra di trasmissione*), l'altro in quella ricevente (*finestra di ricezione*).

Con riferimento, ad esempio, al flusso di UI scambiate fra due estremità *A* e *B* e al verso di trasferimento da *A* a *B*, la finestra di trasmissione (collocata in *A*) specifica qual è il numero massimo di UI che *A* può emettere in funzione della disponibilità di risorse in *B*. La finestra di ricezione (collocata in *B*) precisa invece se il numero massimo di UI ricevute da *B* è effettivamente quello autorizzato ad essere emesso da parte di *A*.

La disponibilità al trasferimento di UI da *A* verso *B* viene aggiornata esclusivamente da *B*: ciò avviene agendo direttamente sulla finestra in ricezione e inviando ad *A* opportune informazioni per il controllo della finestra in trasmissione.

E tuttavia da notare che questa tecnica, come del resto qualunque altra a retro-pressione, può essere impiegata quando il ritmo binario di trasferimento non è troppo elevato in relazione al valore che assume il ritardo di propagazione sulla connessione fisica tra *A* e *B*. Infatti un eventuale ritardo nel rivelare la condizione di congestione e nel prendere le opportune azioni correttive può rendere inefficace l'applicazione di meccanismi del tipo ora descritto.

### **IV.3 Indirizzamento**

L'indirizzo di un utente identifica in modo univoco quest'ultimo nell'ambito di una rete di telecomunicazione. La modalità con cui sono definiti

ed assegnati gli indirizzi (*piano di indirizzamento*) ha significative conseguenze sulla *funzione di instradamento*, e cioè la funzione il cui scopo è guidare l'informazione di utente verso la destinazione voluta (cfr. par. IV.4). Si possono distinguere tre modalità di indirizzamento:

- 1- nella prima l'indirizzo è stabilmente legato ad un luogo fisico e il piano di indirizzamento è definito in modo tale che l'indirizzo, oltre a identificare un utente, dia anche informazioni su dove lo stesso si trovi;
- 2- nella seconda l'indirizzo di per se stesso non consente di dedurre in modo immediato la localizzazione dell'utente, anche se esiste una corrispondenza stabile tra l'utente e il luogo ove esso si trova;
- 3- nella terza non esiste una corrispondenza stabile tra indirizzo e luogo fisico in cui l'utente si trova in un dato momento.

Nel caso di cui al precedente punto 2 l'algoritmo di instradamento deve prima stabilire *dove* l'utente si trova e quindi scegliere una strada per raggiungerlo. Nel terzo caso bisogna prima stabilire *dove* l'utente si trova *in un dato momento* e quindi scegliere una strada per raggiungerlo.

Per esemplificare prendiamo in considerazione diversi casi, in ordine crescente di complessità, essendo il caso più semplice quello in cui un indirizzo ci da' direttamente informazioni stabili sulla localizzazione della destinazione voluta:

- nella rete telefonica fissa gli indirizzi (=numeri telefonici) sono organizzati in modo gerarchico per cui un indirizzo, oltre ad identificare uno specifico utente, fornisce precise informazioni sulla localizzazione dello stesso; ad esempio il numero telefonico "39 06 44 58 XXXX" è legato ad uno specifico luogo (e apparato telefonico): 39 è l'identificativo dell'Italia, 06 il prefisso per la città di Roma, 44 può specificare una particolare centrale di commutazione, 58 un centralino privato e XXXX un interno del centralino; quando all'algoritmo di instradamento giunge una richiesta di connessione verso tale numero è immediato sapere dove si trova la destinazione;
  - nell'attuale Internet basata sul protocollo IP senza prestazioni di mobilità, l'associazione tra indirizzo e luogo fisico in cui l'utente si trova può essere considerata relativamente stabile ma l'indirizzo fornisce limitate informazioni su dove il relativo utente si trova (il piano di indirizzamento non è organizzato in modo gerarchicamente legato alla localizzazione territoriale degli utenti: non esistono prefissi internazionali, interurbani etc.). Ciò significa che parte integrante dell'algoritmo di instradamento è anche stabilire la localizzazione dell'utente; si noti però che una volta stabilita l'associazione indirizzo-luogo questa può essere assunta stabile, a meno di cambiamenti che possono però avvenire solo in seguito ad operazioni di carattere gestionale e non quindi sotto il controllo di un generico utente;
- si noti che un caso simile si presenta anche nella rete telefonica fissa quando si usufruisce del servizio "numero verde"; in tal caso l'indirizzo (del tipo 167 X...X) non rientra nel piano gerarchico di indirizzamento, non si usano i consueti prefissi inter-urbani e non è possibile, esaminando il numero, stabilire la localizzazione del relativo utente di destinazione; in tal caso devono quindi essere invocate opportune procedure (nella fattispecie ottenute mediante la cosiddetta Rete Intelligente) per ottenere la corrispondenza

- indirizzo->localizzazione prima di poter iniziare a scegliere una strada opportuna;
- nelle reti radio-mobile cellulari un indirizzo è in corrispondenza con un dato utente ma la posizione di quest'ultimo varia nel tempo; l'algoritmo di instradamento deve quindi stabilirne la localizzazione in modo dinamico, prima di poter scegliere una strada attraverso la quale fargli pervenire una data informazione; infatti, un numero telefonico di un terminale cellulare non fornisce alcuna indicazione su dove il relativo utente si trovi in un certo momento.

L'evoluzione delle reti di telecomunicazione ha reso inoltre la questione appena discussa ancora più complessa. Nei classici paradigmi di comunicazione vi è infatti la tendenza ad identificare un terminale con l'utente (o gli utenti) che ne fa (fanno) uso e l'indirizzo si riferisce al terminale ma anche, per estensione, all'utente. Il concetto di mobilità è stato però esteso dalla mobilità di terminale e cioè dalla possibilità di accedere ad un servizio di telecomunicazione con un terminale in movimento (vedi punto precedente) alla mobilità personale. Nel primo caso l'indirizzo è legato ad un terminale fisico; quest'ultimo può spostarsi da un luogo ad un altro ma permane un'associazione tra terminale ed utente, ovvero l'indirizzo è legato ad uno specifico terminale. Nel secondo caso (mobilità personale) l'indirizzo è legato all'utente che può quindi registrarsi su diversi terminali, ovvero usare diversi terminali mantenendo la propria identità.

Il sistema GSM è un esempio di tale nuovo concetto di mobilità: un utente è identificato da un indirizzo contenuto in una carta: inserendo quest'ultima in un terminale si personalizza il terminale. Le chiamate indirizzate verso quell'utente saranno instradate sul terminale su cui l'utente stesso si è registrato e la tariffazione delle chiamate effettuate da un terminale sarà addebitata all'utente registrato su quel terminale. Tale possibilità è in progetto di essere estesa anche alle reti non mobili: un utente potrà registrarsi su qualunque terminale, sia fisso che mobile, e la rete, per instradare l'informazione verso un dato utente, dovrà determinare su quale terminale un dato utente è registrato, poi dove quel terminale si trova in quel dato momento (se si tratta di un terminale mobile) e quindi scegliere una strada che porti a quel terminale.

#### **IV.4 Instradamento**

L'instradamento è una funzione di natura logica avente lo scopo di guidare l'informazione di utente verso la destinazione voluta. Decidere un criterio di instradamento significa stabilire come scegliere un percorso attraverso la rete logica in modo tale che l'informazione emessa da un utente giunga ad un altro utente *rispettando* opportuni requisiti prestazionali. Più in dettaglio si può dire che l'instradamento è una funzione decisionale che si attua con lo svolgimento di un opportuno algoritmo: per ogni coppia origine-destinazione occorre individuare un percorso, soddisfacendo determinati criteri prestazionali sia dal punto di vista degli utenti, (qualità del servizio da questi percepita: ritardo di trasferimento, perdita di informazione, eventuale ritardo di instaurazione, etc.) sia da quello degli operatori di rete (efficienza di utilizzazione delle risorse: traffico totale che la rete può trasportare, resistenza ai guasti, economicità e semplicità di implementazione e di gestione, etc.).

La questione della localizzazione di un indirizzo, trattata nel paragrafo precedente, può essere isolata come una prima parte della funzione di instradamento (alcuni la considerano una funzione preventiva all'instradamento vero e proprio). Quale che sia il piano di indirizzamento usato esisterà dunque una prima fase (più o meno complessa) in cui dall'indirizzo si risalirà alla localizzazione dell'utente destinatario nell'ambito di una data rete. A questo punto la funzione di instradamento può scegliere il percorso migliore verso la località precedentemente individuata.

#### IV.4.1 Generalità sugli algoritmi di instradamento

Con queste premesse, e assumendo risolto il problema della localizzazione dell'utente destinatario (problema che è risolto in modo specifico all'interno di determinate reti di telecomunicazione), ci si concentrerà nel seguito nel problema generale di individuare un percorso da una data origine ad una data destinazione. Si noti che, in una certa situazione, potrebbe esistere più di un percorso che soddisfa i requisiti voluti; in tal caso, tra i percorsi possibili, si può scegliere quello che ottimizza determinati parametri.

La funzione di instradamento è complessa per tre ordini di motivi:

- richiede un coordinamento tra diversi sistemi ed i sistemi coinvolti possono essere molto numerosi;
- deve poter far fronte ad eventuali guasti e malfunzionamenti sia dei nodi che dei rami di rete, eventualmente re-instradando il traffico per cui sono state già prese decisioni di instradamento e tenendo conto delle mutate condizioni per il traffico da instradare in seguito;
- deve, ove possibile e conveniente, tenere in conto lo stato di occupazione delle risorse di rete in modo dinamico e/o adattativo; ciò al fine di scegliere i percorsi evitando porzioni di rete che risultino in congestione, così da ottimizzare sia l'efficienza di utilizzo delle risorse stesse sia le prestazioni percepite dagli utenti.

La funzione di instradamento comporta anche una funzione ancillare e cioè la distribuzione a tutti i nodi di rete delle informazioni necessarie per svolgere tale funzione. In altri termini bisogna che ogni nodo di rete sia reso edotto delle informazioni di cui ha necessità per poter prendere decisioni di instradamento; tali informazioni possono essere:

- dati non preventivamente elaborati e riguardanti, ad esempio, l'occupazione delle risorse, la topologia di rete, i guasti e le riparazioni di elementi di rete; in tal caso ogni nodo elabora localmente le decisioni di instradamento, sulla base dei dati ricevuti. In tal caso si parla di algoritmi di instradamento

*distribuiti*;

- informazioni già elaborate da altri sistemi di rete che forniscono quindi direttamente istruzioni sulle scelte da prendere. In tal caso si parla di algoritmi di instradamento *centralizzati*.

Tale funzione può essere considerata un problema aggiuntivo e a se stante ma che nondimeno risulta di considerevole importanza: la rete deve trasportare non solo le informazioni di utente ma anche altre informazioni necessarie ai nodi di rete per svolgere la funzione di instradamento.

Le prestazioni di rete sono fortemente influenzate dalla funzione di instradamento sia in termini della *quantità di servizio* (portata, di precipuo interesse dell'operatore di rete) sia in termini della *qualità di servizio* (ritardo di trasferimento ed eventuale perdita di informazione, parametri direttamente percepiti dagli utenti). La funzione di instradamento interagisce con la funzione di controllo di congestione nel determinare tali misure prestazionali mediante un meccanismo a contro reazione mostrato in Fig. IV.1.

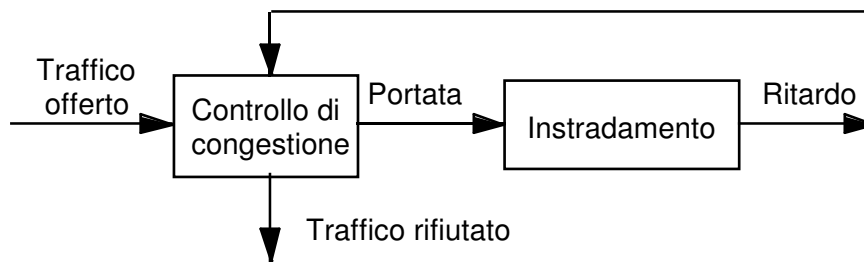


Fig. IV.1 - Interazione tra instradamento e controllo di flusso

Quando il traffico offerto alla rete è considerato eccessivo, il controllo di congestione ne rifiuterà una parte e le prestazioni viste dal traffico accettato saranno influenzate dalla funzione di instradamento. D'altra parte il controllo di congestione decide se accettare o meno il traffico sulla base delle prestazioni che gli utenti percepiranno (ad es. il ritardo di trasferimento) per cui, se la funzione di instradamento opera in modo adeguato, essa riuscirà a mantenere basso il ritardo di trasferimento e il controllo di congestione accetterà una maggiore quantità di traffico. In altri termini lo scopo della funzione di instradamento è anche quello di fare operare il controllo di congestione lungo una curva qualità-portata (ad esempio ritardo-portata) che sia la più favorevole possibile (Fig. IV.2); ovvero aumentare la portata, a parità di ritardo di trasferimento subito dagli utenti in condizioni di alto traffico, e diminuire il ritardo, in condizioni di basso traffico.

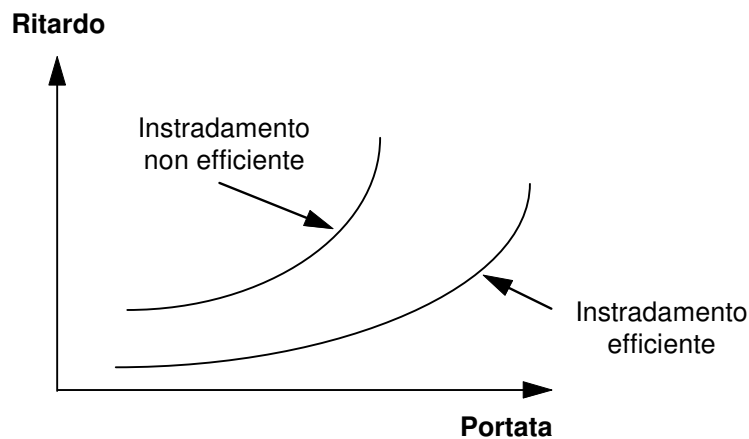


Fig. IV.2 - Interazione tra instradamento e controllo di flusso

Entrando ora in maggiore dettaglio, si può definire l'instradamento come quella funzione decisionale, svolta in ogni nodo di rete, ed avente lo scopo di stabilire il ramo di uscita verso cui deve essere inoltrata una unità informativa (UI) che perviene da un dato ramo di ingresso. (xxx compito decisionale e attuativo ?)

Per approfondire questa definizione, è opportuno distinguere i due casi di servizio di trasferimento con connessione e senza connessione. Nel secondo caso (che si verifica, ad esempio, in un modo di trasferimento a pacchetto con servizio a datagramma) ogni UI viene considerata come entità a se stante e viene instradata nella rete in maniera indipendente dalle altre UI che hanno uguali origine e destinazione. Ciò equivale a dire che la funzione di instradamento viene operata per ogni singola UI che attraversa il nodo.

Nel caso invece di servizio di trasferimento con connessione (quale si ha in un modo di trasferimento a circuito o in uno a pacchetto con servizio a chiamata virtuale) la funzione di instradamento viene attivata solo durante la fase di instaurazione della connessione e definisce, nodo per nodo, il *percorso di rete* che devono seguire *tutte* le UI trasferite nell'ambito della comunicazione stessa.

In ambedue i casi, una procedura di instradamento consiste nel fissare le regole che rendono possibile il soddisfacimento delle *richieste di traffico*, facendo uso delle *risorse di rete* in modo da garantire opportune *prestazioni di trasferimento*. L'efficacia di una procedura di instradamento è quindi legata alla sua capacità di fare fronte a situazioni mutevoli nel tempo sia a causa delle variazioni delle richieste provenienti dagli utenti della rete, sia a causa della vulnerabilità (ad es. per effetto di guasti o di altre cause di fuori-servizio) delle risorse di cui la rete dispone.

Dal punto di vista di questa caratteristica, si possono allora distinguere

procedure di instradamento *statiche*, ovvero *dinamiche*, ovvero infine *adattative*.

Le procedure di instradamento statiche sono basate su regole che, in generale, assumono come invarianti nel tempo sia la topologia della rete (e cioè le sue risorse), sia le caratteristiche delle richieste di traffico di accesso. In base a queste informazioni, in ogni nodo della rete vengono predisposte opportune tabelle (*tabelle di instradamento fisse*), che consentono di operare, di volta in volta, la decisione di instradamento. In queste tabelle è possibile prevedere, oltre a un instradamento-base, anche instradamenti alternativi da scegliere nel caso in cui si verificano modifiche dello stato della rete tali da rendere non conveniente o impossibile la scelta dell'instradamento-base.

Le procedure di instradamento dinamiche utilizzano invece tabelle che variano nel tempo, mentre quelle adattative dipendono da certe stime sullo stato della rete nell'istante in cui viene presa la decisione di instradamento. Notiamo che una procedura di instradamento dinamica non è necessariamente adattativa, dato che la variabilità nel tempo delle tabelle di instradamento può essere di tipo programmato (con variazioni a livello giornaliero, settimanale, mensile, ecc.) senza l'effettuazione di stime sullo stato della rete. Invece un instradamento adattativo è normalmente dinamico, a meno che lo stato della rete rimanga immutato nel tempo.

Nelle procedure di instradamento adattative è possibile distinguere vari gradi di adattività a seconda della rapidità con cui le relative regole vengono aggiornate per fare fronte alle mutate situazioni in cui si opera. In generale, è conveniente limitare questa rapidità di aggiornamento, in quanto la raccolta dei dati sulle modifiche dello stato della rete implica un traffico addizionale che può limitare gli altri flussi informativi di utente e/o di segnalazione attraverso la rete. Anche per le procedure di instradamento adattative sono impiegate tabelle, una per ogni nodo della rete, che, a differenza di quelle fisse nell'instradamento statico, possono modificarsi nel tempo.

Come è evidente, le procedure di instradamento adattative sono più evolute rispetto a quelle statiche, in quanto consentono di conseguire migliori prestazioni globali. Sono però anche più complesse e quindi più costose e ciò ne ha finora limitato l'impiego alle reti di concezione più recente, quali, ad esempio, le reti che operano con modo di trasferimento orientato al pacchetto (nel seguito indicate per brevità come reti a pacchetto). Nelle reti di impostazione sistemistica più tradizionale, quali ad esempio le reti telefoniche, si sono fino ad oggi impiegate procedure di instradamento statiche. E tuttavia iniziato, anche sulla base dell'aumentata capacità elaborativa presente negli

autocommutatori di tecnica più recente, l'impiego di procedure di instradamento dinamiche.

Le procedure di instradamento da adottare in reti a pacchetto dovrebbero idealmente possedere i seguenti attributi:

- *semplicità computazionale*; cioè l'algoritmo di instradamento dovrebbe impegnare in misura minima la capacità di elaborazione dei nodi della rete;
- *robustezza*; cioè dovrebbe essere garantita la possibilità dell'algoritmo ad adattarsi ai cambiamenti dei livelli di flusso di traffico attraverso la rete e a trovare percorsi di rete alternativi quando i nodi e/o i rami si guastano o ritornano in servizio dopo la riparazione;
- *stabilità*; cioè l'algoritmo dovrebbe essere in grado di convergere verso una soluzione accettabile senza oscillazioni eccessive quando si adatta ai cambiamenti di traffico e di topologia;
- *equità*; cioè le prestazioni di qualità di servizio assicurate dall'algoritmo dovrebbero essere fornite senza discriminazioni tra gli utenti, entro i possibili vincoli di assegnate priorità;
- *ottimalità*; cioè l'algoritmo dovrebbe individuare il percorso di rete più conveniente per minimizzare il valore medio del ritardo di trasferimento di un pacchetto e per massimizzare la portata della rete; particolare attenzione deve essere rivolta anche alla continuità di servizio e ai connessi problemi di affidabilità del trasferimento.

Per rispondere a questo insieme di esigenze, le procedure di instradamento che sono state proposte e, in alcuni casi, realizzate per reti a pacchetto sono normalmente di tipo adattativo. In questo ambito si distinguono procedure con controllo *centralizzato* da quelle con controllo *distribuito*.

Le procedure di instradamento con controllo centralizzato prevedono l'esistenza di un *centro di controllo di rete* (CCR), al quale confluiscono le informazioni riguardanti sia il livello del traffico di ingresso alla rete, sia l'integrità e il grado di occupazione dei rami e dei nodi. Il CCR elabora queste informazioni secondo l'algoritmo di instradamento che è stato prescelto, aggiorna di conseguenza le tabelle relative ai vari nodi della rete e provvede a trasferirne i contenuti a questi ultimi.

Nel caso di soluzione centralizzata, si può conseguire più agevolmente la condizione di ottimalità della procedura di instradamento, sollevando al tempo stesso i nodi di rete dai compiti di elaborazione del relativo algoritmo. Questo può allora essere di maggiore complessità computazionale. Lo svantaggio principale di un controllo centralizzato risiede nella vulnerabilità del sistema, dato che le conseguenze di un guasto del CCR possono essere disastrose ai fini

della continuità del servizio: la rete può infatti continuare a funzionare basandosi sull'ultimo aggiornamento fornito dal CCR prima del suo guasto, ma questa informazione diventa rapidamente priva di utilità, con la conseguenza che le prestazioni del trasferimento diventano inaccettabili.

Nel caso di procedure di instradamento adattative con controllo distribuito, ogni nodo provvede a raccogliere le informazioni per lui accessibili sullo stato della rete e provvede a elaborarle per aggiornare le proprie tabelle; conseguentemente ogni decisione di aggiornamento viene presa localmente. Per queste procedure, si possono distinguere i casi in cui ogni nodo: a) opera senza interazioni con gli altri nodi, limitandosi a utilizzare le informazioni sullo stato della rete che possono essere raccolte localmente (*algoritmi isolati*); b) aggiorna la propria tabella di instradamento sulla base di dati scambiati con gli altri nodi e, più in particolare, con quelli ad esso adiacenti (*algoritmi cooperativi*).

In ambedue i casi di controllo centralizzato o distribuito (con algoritmo cooperativo) sorge, come già detto, la necessità di limitare il traffico di aggiornamento. Una soluzione è rappresentata da procedure quasi-statiche, che prevedono un aggiornamento dello stato della rete a intervalli regolari abbastanza lunghi (*aggiornamento sincro*) o addirittura solo in occasioni in cui si verificano variazioni di un certo interesse (*aggiornamento asincro*). Nel primo caso un intervallo di aggiornamento tipico può essere dell'ordine di una decina di secondi.

Gli algoritmi di instradamento attualmente in uso o semplicemente proposti sono molto numerosi e sono caratterizzati da diversi livelli di sofisticazione e di efficienza. La ragione di una tale varietà è dovuta sia a ragioni storiche che alle diverse esigenze di specifiche reti. Nel seguito della sezione si esamineranno in modo qualitativo alcune classi di procedure di instradamento adottate nelle reti per dati a pacchetto.

#### IV.4.2 Procedure basate sulla ricerca del percorso più breve

Molti algoritmi di instradamento usati in situazioni reali sono basati sulla nozione di *percorso più breve* tra due nodi (shortest path routing). La rete logica viene dapprima rappresentata mediante un grafo orientato. Ad ogni ramo è assegnata una lunghezza (o peso). La scelta di questi pesi dipende dalla metrica che si intende utilizzare per valutare la distanza tra due generici nodi, definita come somma dei pesi dei rami che occorre percorrere per passare da un nodo all'altro. L'opzione più semplice consiste nel definire pari all'unità il peso di ogni ramo; in tal modo il percorso più breve è semplicemente il percorso che attraversa il numero minimo di rami (percorso a numero di salti minimo). Il

peso di un ramo può essere però determinato anche con metriche più complesse quali, ad esempio, la distanza fisica tra le sue estremità, la sua capacità trasmissiva, il ritardo di trasferimento di un pacchetto-tipo attraverso il nodo a monte e il ramo considerato, la lunghezza della coda a monte, e così via.

Si vede quindi che, in funzione della metrica utilizzata, un simile algoritmo può essere sia statico che dinamico che adattativo. Inoltre questa procedura è stata realizzata con soluzioni a controllo sia centralizzato che distribuito. Una volta definita la metrica esistono diversi algoritmi che consentono di identificare il percorso tra due nodi tale che la distanza tra questi (misurata secondo la metrica prescelta) sia minima. A titolo di chiarificazione illustriamo un algoritmo di questo tipo.

L'algoritmo di instradamento del percorso più breve è sviluppabile in una molteplicità di passi, con complessità crescente all'aumentare del numero di nodi della rete. Per chiarirne lo sviluppo facciamo riferimento alla Fig. IV.3, dove il punto di partenza dell'algoritmo è definito in (a). In questo grafo, ai vari rami componenti sono stati attribuiti i relativi pesi. L'applicazione dell'algoritmo è riferita a una coppia di nodi di origine e di destinazione, che, nel caso di Fig. IV.3, si identificano con  $A$  e  $D$ . I passi di sviluppo dell'algoritmo, che nel caso specifico di Fig. IV.3, sono cinque, sono illustrati ordinatamente in (b), (c), (d), (e) e (f).

Si parte dal nodo di origine ( $A$  nel caso specifico) e si individua, tra i nodi adiacenti, quello a distanza minima. Questo è assunto come *nodo di lavoro*. Il risultato del primo passo (Fig. IV.3b) individua il nodo  $B$  come primo nodo di lavoro, in quanto la sua distanza dal nodo di origine è inferiore a quella del nodo  $G$ . I nodi  $B$  e  $G$  sono etichettati (tra parentesi) con la loro distanza dal nodo di origine e, accanto a tale distanza, si indica anche il nodo  $A$  da cui si è partiti per valutarne la distanza. Dato che, in questo primo passo, non è stato esplorato alcun percorso, al di fuori di quelli corrispondenti ai rami  $AB$  e  $AG$ , gli altri nodi del grafo sono etichettati con distanza infinita.

Man mano che l'algoritmo procede e i percorsi possibili sono esplorati, le etichette di ogni nodo sono variate in relazione alla distanza di questo dal nodo di origine lungo il percorso più breve identificato nel passo precedente. L'etichetta di uno specifico nodo  $X$ , sempre comprendente la distanza di  $X$  dal nodo di origine e l'indicazione del nodo attraverso cui può essere raggiunto  $X$  con quella distanza, può quindi essere *di tentativo* o *permanente*. Inizialmente tutte le etichette sono di tentativo. Quando si accerta che l'etichetta di un nodo  $X$  è relativa al percorso più breve possibile per raggiungere  $X$  partendo dal nodo di origine, essa è resa permanente e quindi non più modificabile nei passi

successivi dell'algoritmo.

Ogni passo è caratterizzato dal proprio nodo di lavoro (indicato con una freccia in Fig. IV.3), nel senso che vengono presi in considerazione tutti i nodi ad esso adiacenti e per questi vengono modificate, se possibile, le relative etichette. A questo punto si esamina l'intero grafo per cercare il nodo con etichetta di tentativo contenente il minimo valore di distanza dal nodo di origine. Il nodo così individuato assume etichetta permanente e diventa il nodo di lavoro del passo successivo. E' immediato verificare l'applicazione di queste regole nei contenuti di Fig. IV.3, da cui emerge che il percorso più breve tra i nodi A e D è quello che attraversa i nodi B,E,F e H.

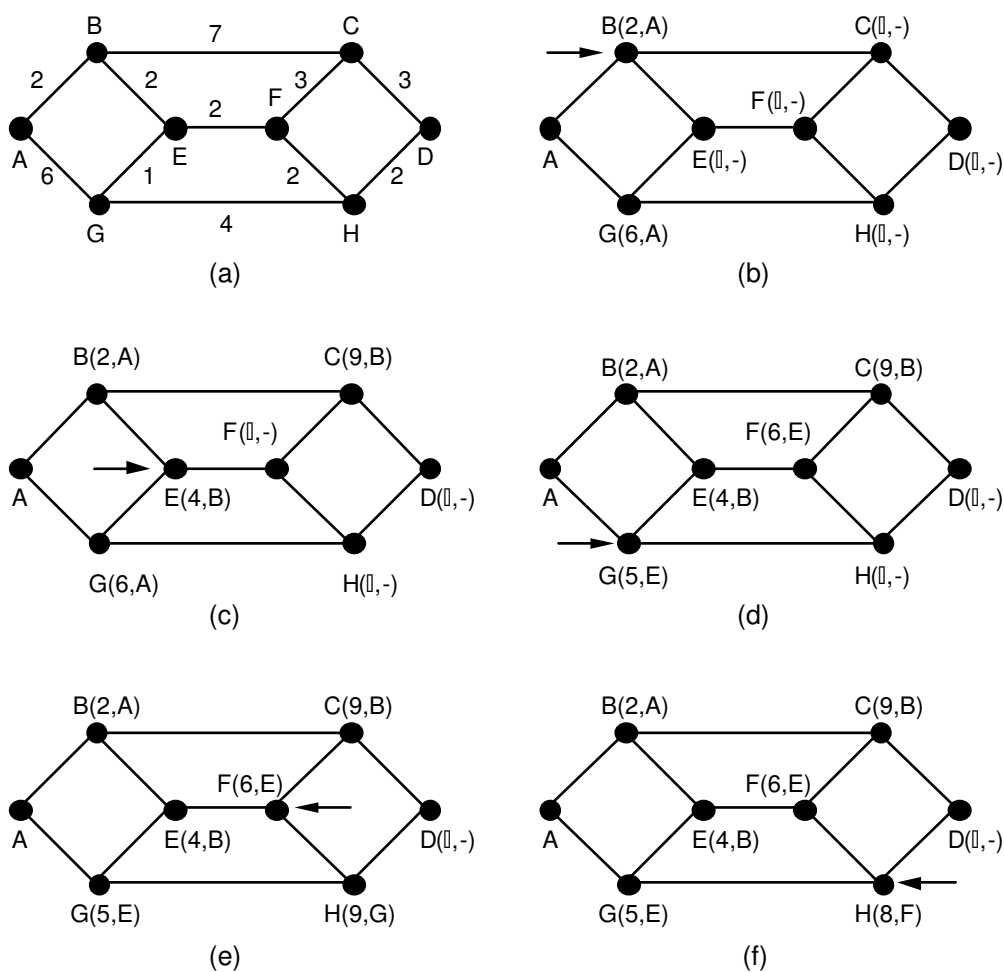


Fig. IV.3 - Applicazione dell'algoritmo di instradamento del percorso più breve; A e D sono i nodi di origine e di destinazione.

Per giustificare come l'algoritmo ora descritto consenta di risolvere il problema della ricerca del percorso più breve si consideri la Fig. IV.3c, corrispondente al secondo passo dell'algoritmo, in corrispondenza del quale il nodo E viene designato come nodo di lavoro. Occorre dimostrare che non può esistere un percorso più breve di ABE per andare da A ad E.

Supponiamo allora che esista un percorso, ad esempio *AXYZE*, che sia più breve di *ABE*. Esistono due possibilità: l'etichetta di Z è stata già resa permanente o non lo è stata. Se lo è

stata (prima possibilità) allora  $E$  è già stato preso in considerazione in un passo seguente a quello in cui l'etichetta di  $Z$  è stata resa permanente. In queste condizioni il percorso  $AXYZE$  non è sfuggito alle iterazioni dell'algoritmo.

Se invece  $Z$  ha ancora una etichetta di tentativo (seconda possibilità) si possono verificare due casi:

- l'etichetta di  $Z$  è maggiore o uguale a quella di  $E$ , nel qual caso il percorso  $AXYZE$  non può essere più breve di  $ABE$ ;
- l'etichetta di  $Z$  è minore di quella di  $E$ , nel qual caso  $Z$  e non  $E$  deve assumere etichetta permanente; conseguentemente è garantito che  $E$  sarà preso in considerazione da  $Z$  e vale quindi la stessa conclusione a cui si è giunti relativamente alla prima possibilità.

#### IV.4.3 Instradamento ottimo

L'instradamento basato sulla ricerca del percorso più breve è abbastanza diffuso, grazie anche alla sua relativa semplicità realizzativa, ma presenta due limitazioni principali:

- la portata che riesce ad offrire è potenzialmente limitata poiché tale algoritmo prevede un solo cammino per ogni coppia origine-destinazione;
- la sua capacità di adattarsi a mutate condizioni di traffico o di disponibilità di risorse di rete è ridotta a causa di fenomeni di oscillazione.

Per chiarire questi due punti ci si avvale delle seguenti considerazioni. Si consideri la rete (avente scopo esclusivamente didattico) mostrata Fig. IV.4 e composta da 5 nodi e da 5 rami. Si supponga che:

- la rete trasferisce pacchetti di lunghezza costante;
- ogni ramo possa trasportare 10 pacchetti nell'unità di tempo;
- al nodo 1 siano offerti 15 pacchetti destinati al nodo 5;
- alla rete non sia offerto altro traffico.

Nel caso in cui si adotti un algoritmo di instradamento che prevede un solo cammino per ogni coppia origine destinazione, solo 10 pacchetti possono essere trasportati: ad esempio lungo il percorso 1->2->5, che risulta essere il percorso più breve nel caso in cui si scelga pari ad uno il peso di ogni ramo; i restanti 5 pacchetti saranno rifiutati. Se invece per ogni coppia origine-destinazione l'algoritmo di instradamento potesse usare più di una solo percorso, i 15 pacchetti potrebbero essere divise, inviando ad esempio 8 pacchetti lungo il percorso 1->2->5 e i restanti 7 lungo il percorso 1->3->4->5, evitando così di rifiutare del traffico.

Si noti però che in quest'ultima alternativa non è più garantita la consegna in sequenza dei pacchetti. Quindi il prezzo da pagare per l'aumento di efficienza così ottenuto è la necessità di prevedere una operazione di ri-sequenziamento dei pacchetti giunti a destinazione.

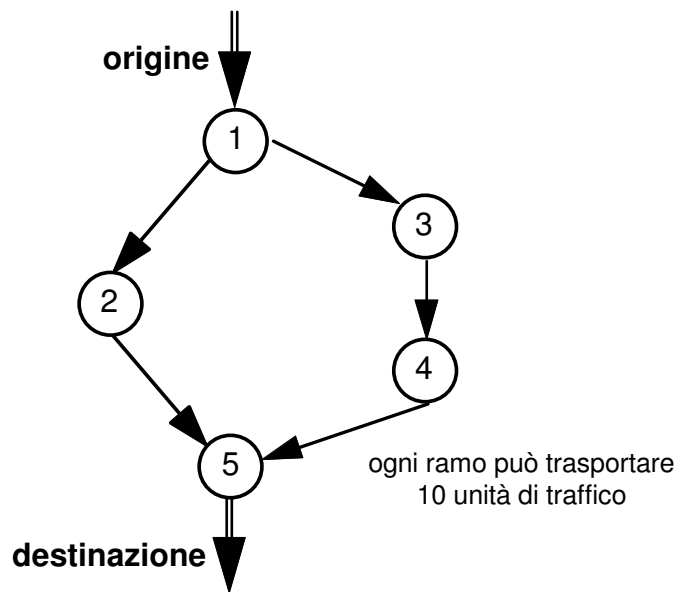


Fig. IV.4 - Limiti dell'algoritmo di instradamento del percorso più breve

Per quanto riguarda il secondo svantaggio dell'algoritmo di instradamento del percorso più breve e cioè la presenza di eventuali oscillazioni si consideri sempre la rete di Fig. IV.4. Se la metrica sui cui è basato l'algoritmo utilizza pesi legati allo stato di occupazione dei rami (come è naturalmente il caso, a meno di ridursi a considerare solo il numero e le capacità statiche dei rami attraversati), allora, assumendo che il peso di un ramo sia legato al ritardo di attraversamento, si può verificare la seguente catena di eventi:

- l'algoritmo sceglie un percorso (ad esempio 1->2->5) perché in quel momento risulta non occupato da altro traffico e quindi è caratterizzato da un basso valore del ritardo di attraversamento;
- a causa del traffico appena instradato lungo tale percorso il relativo ritardo di attraversamento aumenta, superando quello del percorso 1->3->4->5;
- di conseguenza l'algoritmo re-instrada il traffico che aveva appena diretto lungo il percorso 1->2->5, spostandolo lungo il percorso 1->3->4->5;
- a causa di questa scelta, il ritardo di attraversamento lungo il percorso 1->3->4->5 supera quello relativo al percorso 1->2->5 e l'algoritmo riporta il traffico sul precedente percorso 1->2->5.

E' immediato vedere che questo modo di operare conduce ad un andamento oscillatorio in cui il traffico è continuamente spostato da un percorso all'altro.

L'instabilità tipica di un tale fenomeno a contro-reazione può essere

ridotta: i) limitando la velocità con cui l'algoritmo può cambiare strade; ii) aggiungendo una costante positiva di elevato valore al peso di ogni singolo ramo, così da diminuire l'effetto della variazione del peso dei rami; iii) rendendo asincroni i messaggi che i nodi si scambiano al fine di comunicarsi le variazioni dei pesi dei rami, in modo tale che la modifica dello stato di un dato percorso non si propaghi in un breve intervallo di tempo in tutta la rete, ma avvenga in modo graduale. Purtroppo l'introduzione di simili accorgimenti limita anche la sensibilità dell'algoritmo alle variazioni delle condizioni di traffico o di disponibilità di risorse di rete.

L'instradamento ottimo, invece, si pone come fine l'ottimizzazione globale su tutta la rete di una data misura prestazionale, ad esempio il ritardo medio di trasferimento. In altre parole non si limita a trovare quale sia il percorso migliore, in un dato momento, per ogni singola relazione di traffico alla volta, ma opera in modo tale da cercare un stato di instradamento che sia ottimo per tutta la rete nel suo insieme. E' facile vedere infatti come cercare la strada migliore per una data relazione di traffico non equivalga a rendere migliori le prestazioni dell'intera rete. In un dato momento ci si può trovare di fronte alla seguente alternativa::

- scegliere la strada migliore per una certa coppia origine-destinazione anche se tale scelta va a peggiorare le prestazioni viste da altre relazioni di traffico (e queste ultime dovranno quindi adeguarsi alle mutate situazioni);
- scegliere per una certa coppia origine-destinazione una strada che non sia la migliore possibile ma tale che le prestazioni medie globali siano migliori rispetto all'alternativa precedente (possibilmente le migliori in assoluto).

Nel primo caso si sta applicando il principio del percorso più breve, nel secondo il principio dell'instradamento ottimo.

L'instradamento ottimo può dividere il traffico di una data relazione lungo più strade e allocare gradualmente il traffico lungo cammini alternativi; i corrispondenti algoritmi sono ovviamente più sofisticati e complessi.

#### IV.4.4 Instradamento a deflessione

Con il nome di *instradamento a deflessione* o *della patata bollente* si designa un algoritmo d'instradamento molto semplice e robusto, che può essere classificato come adattativo di tipo isolato. E' applicato nei casi in cui le capacità di memorizzazione dei nodi siano molto limitate e si basa sul principio di rilanciare un pacchetto immediatamente, senza immagazzinarlo, eventualmente indirizzandolo verso un'uscita diversa da quella preferita (nel senso che non è quella che lo porterebbe più vicino alla sua destinazione) se

quest'ultima è già impegnata.

Supponiamo che più di un pacchetto “preferisca” andare verso una certa destinazione nello stesso momento: solo uno di questi vi sarà inviato mentre gli altri saranno inoltrati verso altre destinazioni “meno gradite” da cui poi saranno rilanciati verso la destinazione finali se vi saranno risorse disponibili, altrimenti saranno ancora “deflessi” verso ulteriori destinazioni sub-ottimali. Prendendo opportune precauzioni, tra cui quella di evitare che un pacchetto sia inserito in un ciclo chiuso da cui non riesce ad uscire, si può garantire che la totalità dei pacchetti (o una grande maggioranza di essi) arrivi a destinazione. Lo svantaggio di tale schema è che non preserva la sequenza dei pacchetti e che introduce un ritardo di trasferimento che può essere notevolmente variabile.

#### IV.4.5 Instradamento diffusivo

Come accennato in precedenza, durante il funzionamento di un algoritmo di instradamento può essere necessario diffondere alcune informazioni che consentono all'algoritmo stesso di poter operare (ad es. lo stato di occupazione delle risorse di rete). In molte situazioni, tali informazioni devono essere inviate da un nodo di origine a *tutti* gli altri nodi: come esempio di nodo di origine si pensi a

- un centro di controllo di rete, nel caso di algoritmi centralizzati, che deve inviare i dati da esso elaborati a tutti i nodi di commutazione;
- un nodo di commutazione che si trovi in stato di congestione e che deve rendere nota tale sua situazione a tutti gli altri nodi (nel caso di algoritmi distribuiti).

In altri casi può essere necessario inviare in modo diffusivo anche l'informazione di utente: si pensi a servizi di telecomunicazione di tipo appunto diffusivo (come potrebbe essere il servizio televisivo).

Una metodologia molto usata per diffondere a tutti i nodi di rete una certa informazione è nota come *inondazione* (“flooding”) e opera come segue. Il nodo di origine invia l'informazione in oggetto a tutti i suoi “vicini” (cioè ai nodi a cui è direttamente connesso). Ogni vicino la rilancia verso i suoi vicini e così via finché l'informazione in tal modo arriva a tutti i nodi della rete. Al fine di limitare il numero di pacchetti trasmessi si osservano due regole aggiuntive:

- un nodo non deve rilanciare un pacchetto verso il nodo da cui lo ha ricevuto;
- un nodo deve trasmettere un pacchetto a ogni vicino una sola volta.

Ciò può essere attuato includendo in ogni pacchetto un identificativo del nodo di origine ed un numero di sequenza che è incrementato con ogni pacchetto emesso dal nodo di origine. Memorizzando il numero di sequenza più

alto ricevuto per ogni nodo di origine e non rilanciando pacchetti con numero di sequenza più basso di quello memorizzato ogni nodo può evitare di trasmettere lo stesso pacchetto più di una volta.

Questo modo di operare è illustrato nella Fig. IV.5 che illustra anche come il numero totale di pacchetti trasmessi possa variare tra  $L$  e  $2L$  dove  $L$  è il numero di rami bi-direzionali della rete.

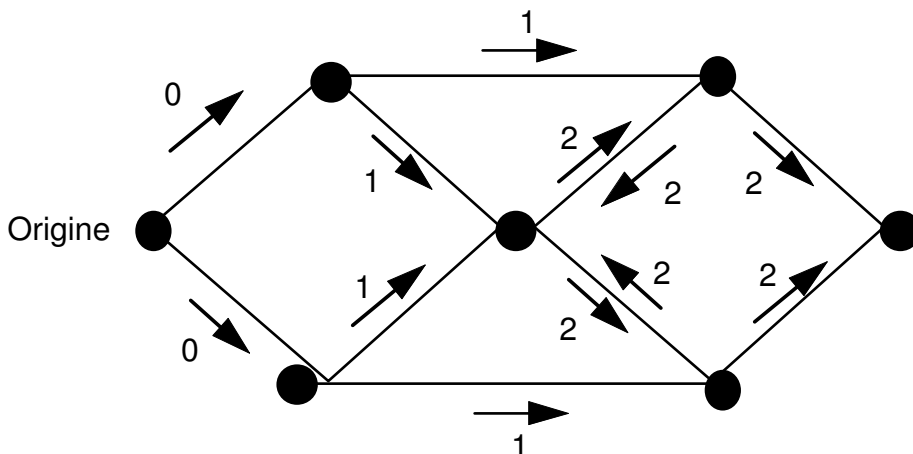


Fig. IV.5 - Esempio di algoritmo diffusivo

Un altro metodo di instradamento diffusivo è basato sulla nozione di *spanning tree* (albero pervasivo).

Secondo la usuale nomenclatura usata nei grafi geometrici, uno *spanning tree* associato a un grafo  $G$  è uno sotto-grafo di  $G$  (cioè con rami che sono un sotto-insieme di quelli di  $G$ ) che include il minor numero di rami necessario per fornire completa connettività tra tutti i nodi di  $G$  (senza formare percorsi chiusi)

Nell'inoltro a *spanning tree* ogni nodo della rete logica conserva informazione su quali dei suoi rami uscenti appartengano allo *spanning tree* che è associato al grafo della rete. I pacchetti che sono instradati in modo diffusivo (*pacchetti diffusivi*) sono distinti da uno speciale indirizzo di destinazione (formato ad es. da tutti "1"). Quando un nodo riceve un pacchetto di tipo diffusivo, ne inoltra una copia su tutti i suoi rami uscenti che appartengono allo *spanning tree*, con l'eccezione di quello che ha la stessa estremità del ramo entrante. Il nodo fornisce anche una copia del pacchetto a tutti i calcolatori che ad esso fanno capo.

A titolo di esempio, si consideri la Fig. IV.6 ove è mostrato lo *spanning tree* associato al grafo di Fig. IV.5. Grazie all'algoritmo ora descritto è possibile effettuare una comunicazione diffusiva che richiede un numero totale di pacchetti trasmessi pari a  $N-1$ , se  $N$  è il numero di nodi del grafo

corrispondente alla rete. Un tale aumento di efficienza rispetto all'algoritmo di inondazione va pagato con la necessità di memorizzare e di aggiornare in ogni nodo interessato lo spanning tree della rete, anche in seguito a variazioni topologiche.

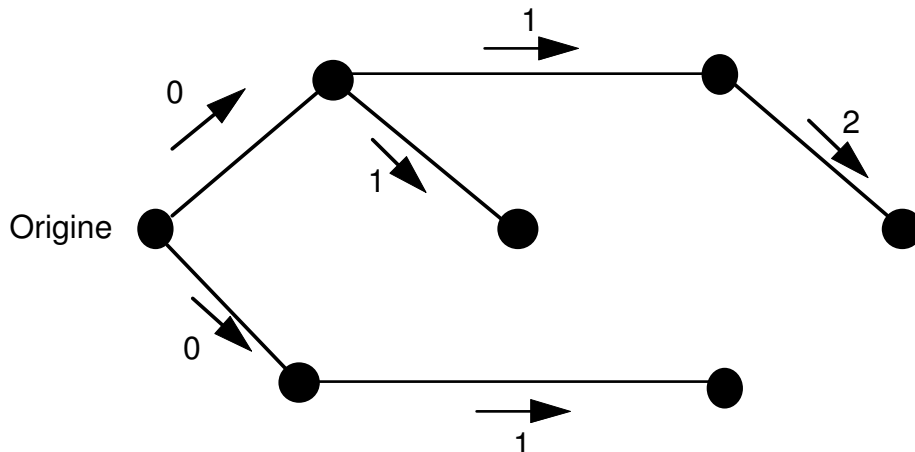


Fig. IV.6 - Esempio di spanning tree

#### IV.4.6 Instradamento in reti interconnesse

In una inter-rete (cfr § I.1.5 e Cap. VII) è necessario fare uso di dispositivi di interconnessione che fungano da interfaccia tra una sotto-rete e l'altra, "traducendo" i protocolli di una nei corrispondenti protocolli dell'altra, e che consentano ad una unità informativa proveniente da una data sotto-rete di giungere ad un'altra sotto-rete, dove si trova il sistema terminale verso cui quell'unità informativa è destinata (Fig. IV.7).

Tra le funzioni che tali dispositivi devono svolgere, vi è, in generale, la funzione di instradamento. Supponendo che l'inter-rete in questione adotti nello strato di rete una modalità di trasferimento senza connessione, un sistema di interconnessione deve svolgere, fra gli altri, i seguenti compiti:

- per ogni unità informativa che prende in consegna, deve determinare dove si trova, nell'ambito della inter-rete, la destinazione verso la quale quella unità informativa è diretta (compito di natura decisionale);
- per ciascuna unità informativa deve individuare una strada che porti quest'ultima verso la destinazione voluta (compito di natura decisionale);
- deve inoltrare fisicamente ogni unità informativa verso la strada prescelta (compito di natura attuativa).

Nel caso di modalità di trasferimento con connessione i compiti di natura decisionale sono svolti solo durante la fase di instaurazione mentre quello di natura attuativa è svolto per ogni unità informativa trattata.

## Inter-rete

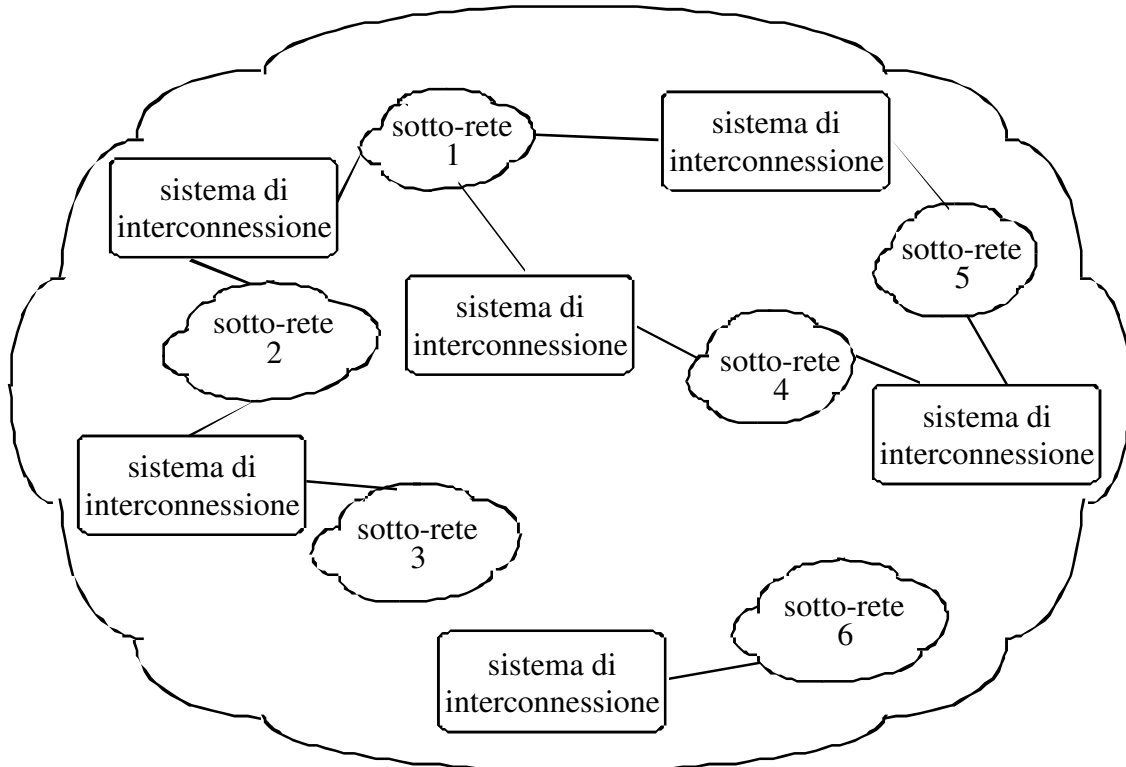


Fig. IV.7 - Esempio di inter-rete

Dal punto di vista della funzione di instradamento, una inter-rete può essere vista in due modi alternativi:

- 1) in una visione ad *un* livello, in cui i dispositivi di interconnessione hanno il ruolo di nodi aggiuntivi rispetto a quelli delle sotto-reti; ogni sistema di interconnessione è connesso a due o più sotto-reti e quindi a due o più nodi appartenenti alle sotto-reti che interconnette; l'intera inter-rete è vista come un insieme di nodi di numerosità pari alla somma di tutti i nodi delle sotto-reti componenti e di tutti i sistemi di interconnessione;
- 2) in una visione a *due* livelli; al primo di questi livelli ogni sotto-rete è vista come entità a se stante e costituita da un certo numero di nodi; tra tali nodi ve ne sono alcuni di natura particolare (i sistemi di interconnessione) che la mettono in comunicazione con il mondo esterno ad essa. Al secondo livello, ogni sotto-rete è vista come un macro-nodo e la struttura interna di ogni sotto-rete non è visibile dall'esterno. L'intera inter-rete è quindi vista, a questo secondo livello, come costituita da un insieme di nodi: alcuni di questi sono macro-nodi, rappresentanti le sotto-reti componenti, mentre i rimanenti sono i sistemi di interconnessione che interconnettono i macro-nodi. A questo secondo livello la numerosità di "nodi equivalenti" è quindi ben minore di

quella di cui al caso 1) poiché la struttura interna delle sotto-reti non è visibile.

L'instradamento basato sul primo di questi due modelli è chiamato *non gerarchico* e può appartenere a una qualunque delle tipologie discusse nei paragrafi precedenti; ad esempio può essere del tipo a ricerca del percorso più breve. In tal caso l'algoritmo di instradamento dovrà cercare il percorso più breve tra *tutti* i nodi che costituiscono l'inter-rete.

L'instradamento basato sul secondo dei modelli di cui sopra è detto *gerarchico*. Nel caso di instradamento gerarchico, si ha una procedura a due livelli. Al livello più basso vi è un algoritmo di instradamento, separato per ogni sotto-rete, il quale:

- gestisce fino a destinazione il traffico locale (interno alla sotto-rete);
- instrada il traffico diretto ad altre sotto-reti verso i dispositivi di interconnessione, limitandosi a farlo arrivare ad uno di questi ultimi e non ponendosi il problema di quale strada fargli percorrere fino al sistema terminale di destinazione finale.

Al livello più alto vi è un algoritmo di instradamento che determina la sequenza di sotto-reti e dei relativi dispositivi di interconnessione che ogni unità informativa deve attraversare se vuole, partendo da un sistema di interconnessione di una sotto-rete, arrivare ad un sistema di interconnessione appartenente ad una diversa sotto-rete. In altre parole l'algoritmo di basso livello gestisce il traffico "locale" e l'algoritmo di alto livello gestisce il traffico di "lunga distanza", guidandolo attraverso diverse sotto-reti per consegnarlo infine all'algoritmo di basso livello della sotto-rete di destinazione, il quale lo farà pervenire al sistema terminale di destinazione.

L'instradamento di tipo gerarchico è tipicamente preferibile in reti di grandi dimensioni (ad esempio è quello usato in Internet e nella rete telefonica mondiale). Per meglio chiarire quanto appena esposto si ricorre ad un esempio.

Si consideri una nazione con diverse città, ognuna delle quali ha una sua rete cittadina connessa alla rete nazionale tramite dispositivi di interconnessione. Supponendo di usare l'algoritmo del percorso più breve, se si usa un instradamento non gerarchico, non vi è distinzione tra le reti cittadine: ogni nodo deve calcolare il percorso più breve verso tutti i nodi di tutta la nazione; ogni nodo deve conoscere la topologia di tutta la rete e la dimensione delle tabelle di instradamento contenute in ogni nodo (ovvero una serie di informazioni, elaborate durante la funzione decisionale, che specificano per ogni destinazione quale strada far seguire all'informazione di utente) è pari al numero di tutti i nodi della rete nazionale.

Se invece si usa un approccio gerarchico, esiste un algoritmo di instradamento locale di basso livello per ogni città, le cui tabelle di instradamento devono contenere informazioni solo per i nodi appartenenti alla stessa città. L'algoritmo di alto livello si occupa

dell'instradamento quando una comunicazione non è limitata all'interno di una sola città e opera nei dispositivi di interconnessione. Ogni dispositivo di interconnessione deve conoscere la topologia della nazione come insieme di città, senza entrare nel merito di come è costituita ogni singola città, e ha una tabella di instradamento la cui dimensione è pari al numero delle città.

Per vedere come l'instradamento gerarchico complessivo lavori, si consideri una unità informativa generata da un sistema appartenente a una data città; l'algoritmo di instradamento locale, di basso livello, prende in consegna l'unità informativa e determina se essa è destinata alla stessa città da cui è stata generata o meno. Nel primo caso è capace di portarla a destinazione, altrimenti la consegna ad un dispositivo di interconnessione; qui è presa in consegna dall'algoritmo di alto livello che ne determina la strada attraverso diversi dispositivi di interconnessione (e quindi attraverso città) fino ad arrivare alla città di destinazione; qui l'unità informativa è di nuovo presa in consegna dall'algoritmo di basso livello che la consegnerà al sistema terminale di destinazione.

E' facile vedere come nel caso di instradamento gerarchico la dimensione delle tabelle di instradamento ed i calcoli da effettuare per determinare la strada più opportuna sono considerevolmente minori che nel caso di un algoritmo non gerarchico. D'altra parte un instradamento gerarchico risulta tipicamente meno ottimale di uno non gerarchico. Bisogna quindi scegliere un opportuno compromesso tra questi due approcci a seconda delle specifiche esigenze.

Si noti infine che l'instradamento gerarchico può essere organizzato in più di due livelli (ad esempio una rete internazionale costituita da più reti nazionali ognuna delle quali comprende diverse città).

## IV.5 Funzionalità del servizio

Limitiamo il nostro esame a un N-servizio orientato alla connessione. Questo offre le seguenti funzionalità:

- 1) i mezzi per instaurare una connessione di rete (*N-connessione*) che lo associa logicamente con un altro NS-utente, allo scopo di trasferire informazione sotto forma di *NSDU*, e cioè di unità di dati dell'*N-servizio* (*instaurazione della N-connessione*);
- 2) la definizione di un accordo, tra due NS-utenti e le entità dello strato di rete (*NS-fornitore*), relativamente a una certa QOS associata ad ogni *N-connessione* (*negoiazione della QOS*);
- 3) i mezzi per trasferire, in sequenza e in modo trasparente, le *NSDU* su una *N-connessione* (*trasferimento di dati normali*);
- 4) i mezzi con i quali l'*NS-utente* ricevente può controllare il ritmo secondo cui l'*NS-utente* emittente può inviargli le *NSDU* (*controllo di flusso*);
- 5) i mezzi per trasferire, quando necessario, *NSDU veloci*, in sequenza e separatamente; queste ultime sono di lunghezza limitata e sono soggette a un controllo di flusso diverso da quello operato sulle *NSDU normali* (*trasferimento di dati veloci*);

- 6) i mezzi in base ai quali, con l'uso di un *servizio di re-inizializzazione*, una N-connessione può essere riportata a uno stato definito, consentendo di sincronizzare le attività dei due NS-utenti tra loro connessi (*re-inizializzazione di una N-connessione*);
- 7) i mezzi che consentono a un NS-utente, in opportuni casi, di confermare al suo corrispondente (e cioè da estremo a estremo) la ricezione di una NSDU (*conferma di ricezione*);
- 8) i mezzi per abbattere senza condizionamenti una N-connessione da parte sia degli NS-utenti che nell'NS-fornitore (*abbattimento di una N-connessione*).

Mentre le due funzionalità di cui in 5) e in 7) possono essere rese disponibili dall'NS-fornitore solo su base opzionale, tutte le altre sei debbono essere fornite obbligatoriamente e sono quindi sempre disponibili.

L'N-servizio è descritto quindi in base alla qualità del servizio che viene fornito in corrispondenza degli N-SAP. Una volta che la N-connessione è instaurata, gli NS-utenti alle due estremità della connessione hanno la stessa percezione della QOS conseguita, anche se questa è realizzata attraverso più reti di diversa qualità.

La QOS di un N-servizio è a sua volta precisata in termini di opportuni *parametri* e la definizione di ognuno di questi specifica il modo secondo cui può essere misurato o calcolato il valore del parametro. Si distinguono i parametri di QOS che esprimono le *prestazioni* dell'N-servizio da quelli che descrivono le *caratteristiche* di quest'ultimo.

Con riferimento ad una N-connessione e alle sue tre fasi evolutive, alla prima categoria appartengono parametri quali:

- a) il *ritardo* e la *probabilità di insuccesso* dell'instaurazione;
- b) la *portata* (throughput), il *ritardo di transito*, il *tasso di errore residuo* (corrispondente a corruzione, duplicazione o perdita delle NSDU) e la *probabilità di insuccesso* del trasferimento;
- c) il *ritardo* e la *probabilità di insuccesso* dell'abbattimento.

Nella seconda categoria troviamo invece, ad esempio, la *priorità di una N-connessione*, con riferimento all'ottenimento o alla tenuta di quest'ultima e quindi all'ordine secondo cui N-conessioni debbono essere interrotte per recuperare risorse a favore di altre. Una ulteriore priorità riguarda l'ordine secondo cui è possibile degradare la QOS di N-conessioni, al fine di favorirne altre.

Con riferimento al caso di servizio orientato alla connessione, in Tab. IV.1 è riportato un elenco delle primitive dell'N-servizio e dei relativi parametri. La

prima e l'ultima primitiva riguardano, rispettivamente, la fase di instaurazione e quella di abbattimento, mentre le rimanenti quattro sono relative alla fase di trasferimento dei dati. Gli elementi di servizio dell'N-strato sono quindi due in più rispetto al caso del DL-strato; quelli aggiuntivi (trasferimento di dati veloci e conferma di ricezione) sono però offerti su base opzionale. Osserviamo poi che la negoziazione della QOS avviene tramite le primitive N-CONNECT, mentre il controllo di flusso è sostanzialmente gestito dalle primitive N-DATA.

Per ragioni di brevità ci limitiamo, a titolo di esempio, a illustrare le primitive N-CONNECT e N-DISCONNECT.

Servizio	Primitiva	Tipo	Parametri
Instaurazione di una connessione di rete	N-CONNECT	Richiesta Indicazione Risposta Conferma	Indirizzo chiamante Indirizzo chiamato Opzione dati veloci Qualità di servizio Dati d'utente
Trasferimento di dati normale	N-DATA	Richiesta Indicazione	Dati d'utente Richiesta di conferma
Trasferimento di dati veloci	N-EXPEDITED DATA	Richiesta Indicazione	Dati d'utente
Conferma di ricezione	N-DATA ACKNOWLEDGE	Richiesta Indicazione	----
Reset	N-RESET	Richiesta Indicazione Risposta Conferma	Iniziatore Ragioni
Abbattimento di una connessione di rete	N-DISCONNECT	Richiesta Indicazione	Dati d'utente Ragioni abbattimento

Tab. IV.1 - Primitive del servizio di rete e loro parametri.

Le primitive N-CONNECT sono utilizzate per instaurare una N-connessione nell'ambito di un servizio di tipo confermato. Tra i loro parametri vengono specificati, innanzitutto, gli indirizzi del chiamato e del chiamante, identificando, nell'ambito di un opportuno indirizzamento di strato, i punti di accesso dell'N-servizio (N-SAP), che sono la destinazione e l'origine dell'N-connessione.

Sono poi contenute due variabili booleane, usate per richiedere i due elementi di

servizio opzionali, e cioè quelli riguardanti il trasferimento di dati veloci e la conferma di ricezione. Ulteriori parametri riguardano la QOS e i dati di utente.

Circa la QOS, per ogni parametro di questa è definito un insieme di sub-parametri, che sono oggetto di negoziazione. Tra questi si distinguono:

- 1) un valore obiettivo, che è quello desiderato dall'NS-utente chiamante;
- 2) un valore di qualità minima accettabile, che è il minimo valore di QOS in grado di soddisfare l'NS-utente chiamante;
- 3) un valore disponibile, che è il valore di QOS che l'NS-fornitore è disposto a fornire;
- 4) un valore selezionato, che è il valore di QOS sul quale l'NS-utente chiamato esprime accordo.

Nel caso in cui la negoziazione non vada a buon fine, la instaurazione della connessione non ha successo.

Infine il parametro "dati di utente" permette all'NS-utente chiamato di includere propri dati (entro una certa lunghezza massima) nella richiesta di connessione. Il chiamato può prendere visione di questi dati, prima di decidere se accettare o meno la richiesta.

Le primitive N-DISCONNECT riguardano un servizio non confermato e sono emesse per l'abbattimento di una N-connessione. Questa funzione può essere effettuata: 1) da uno degli NS-utenti o da entrambi o dall'NS-fornitore, per abbattere una connessione instaurata; 2) dall'NS-utente chiamato, per rifiutare la instaurazione richiesta dal chiamante; 3) dall'NS-fornitore, per indicare la sua inabilità a instaurare la connessione richiesta.

Per le primitive N-DISCONNECT, i parametri sono: l'indicazione dell'iniziatore della procedura di abbattimento, le ragioni dell'emissione della primitiva e i dati di utente.

Circa le ragioni dell'emissione della primitiva, occorre distinguere tra cause transitorie e permanenti, mentre le cause sono legate all'iniziatore. Quando questo è l'NS-fornitore, cause di rifiuto sono, ad esempio, un indirizzamento sconosciuto, un N-SAP non raggiungibile, una QOS non disponibile. Quando l'iniziatore è invece un NS-utente, si può avere abbattimento a seguito, ad esempio, di rifiuto a una richiesta di instaurazione e di QOS non disponibile.

Infine, il parametro "dati di utente" ha lo stesso significato già precisato per le primitive N-CONNECT e, in unione con questo secondo parametro, consente uno scambio di dati contestualmente alle fasi di instaurazione e di abbattimento.

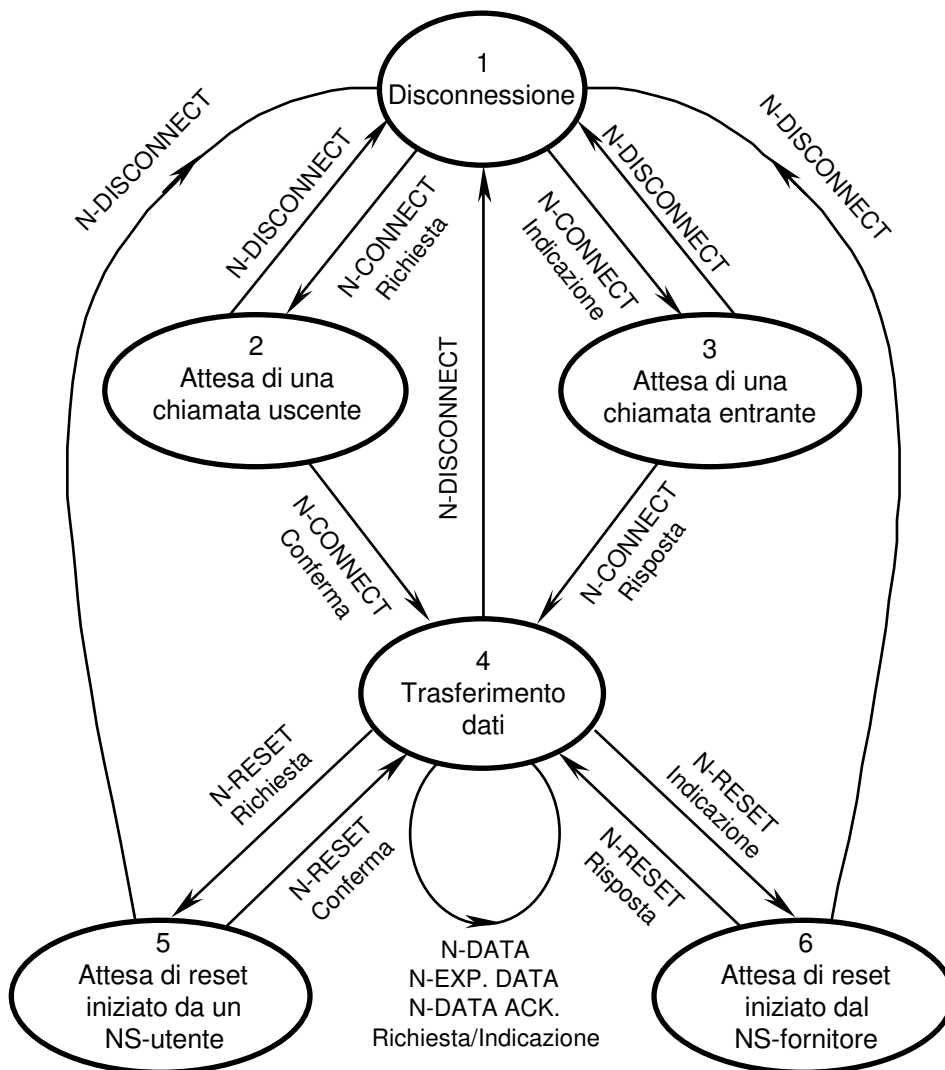
Il funzionamento di una N-connessione può essere rappresentato con un modello a coda del tipo riportato in Fig. IV.8, ove è prevista una coda per ognuna delle due direzioni di flusso informativo e ogni coda collega i due N-SAP.

Questo modello può essere interpretato con le stesse considerazioni svolte nel par. III.2, purché riferite all'interazione tra terzo e quarto strato. L'insieme degli elementi che trovano posto nelle code è tuttavia più ampio, dato il più ricco insieme di primitive dell'N-servizio rispetto a quelle del DL-servizio.

Per quanto riguarda le possibili sequenze di primitive scambiate attraverso gli N-SAP, valgono le stesse considerazioni svolte nel par. III.2 circa le primitive scambiate attraverso i DL-SAP. Alle sequenze previste per il DL-strato vanno tuttavia aggiunte, nell'N-strato, quelle derivanti dall'esistenza dei

due servizi opzionali.

In tal modo il diagramma di transizione di stato, che è mostrato in Fig. IV.8 e che descrive le possibili transizioni aventi luogo a una estremità della N-connesione, differisce da quello di Fig. III.3 solo per l'etichetta dell'arco che rappresenta la permanenza nello stato "Trasferimento dei dati".



N.B. : N-DISCONNECT indica la primitiva di richiesta o quella di indicazione.

Fig. IV.8 - Diagramma di transizione di stato relativo a una estremità di N-connesione per la fornitura di un N-servizio orientato alla connessione.

Infine la Fig. IV.9, che riassume le sequenze di primitive attraversanti i due N-SAP, reca in più, rispetto alle Fig. III.4, Fig. III.5, Fig. III.6, Fig. III.7 e Fig. III.9, le due sequenze che rappresentano la conferma di ricezione (fig. Fig.

IV.9o) e il trasferimento di dati veloci (Fig. IV.9p).

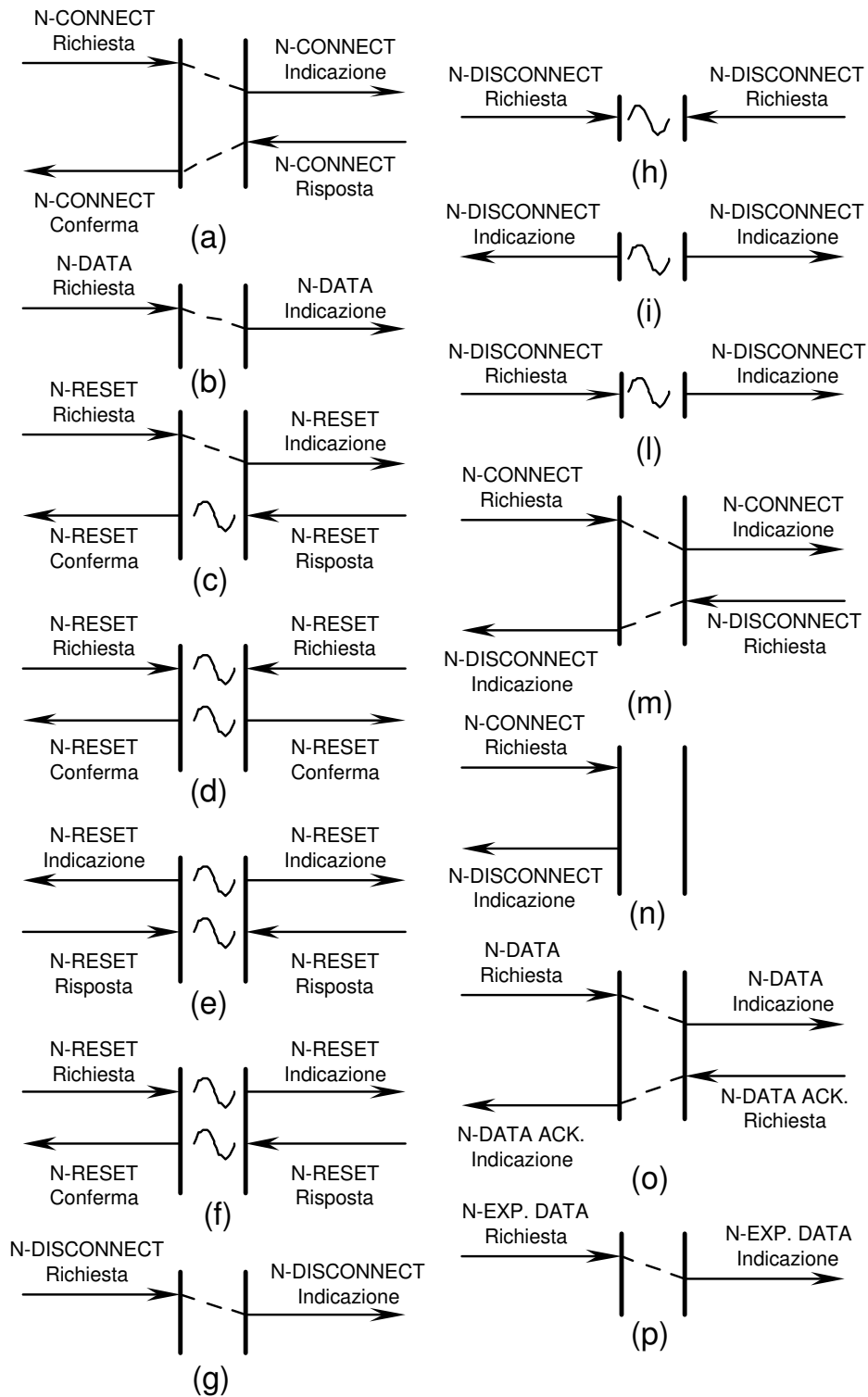


Fig. IV.9 - Possibili sequenze di primitive dell'N-servizio con riferimento a:

- (a) *instaurazione della connessione di rete;*
- (b) *trasferimento di dati;*
- (c,d,e,f) *procedure di re-inizializzazione;*

(g,h,i,l,m,n)	<i>abbattimento della connessione;</i>
(o)	<i>trasferimento di dati con conferma di ricezione;</i>
(p)	<i>trasferimento di dati veloci.</i>

## IV.6 Protocollo X.25 di livello 3

Nel seguito del capitolo verrà descritto, a titolo di esempio, un protocollo di strato di rete: *il protocollo X.25 di livello 3*, che è normalizzato dalla Racc. X.25 in unione all'omonimo protocollo di livello 2 (descritto nel par. III.4) e al protocollo X.21 (trattato nel par. II.4).

La Racc. X.25 è stata già più volte citata in precedenza come l'unico standard internazionale per l'accesso a reti pubbliche per dati a commutazione di pacchetto. Il livello 3 di questa raccomandazione riguarda lo strato di rete dell'interfaccia DTE-DCE.

Tale protocollo è stato definito in epoca precedente alla normalizzazione dell'architettura OSI. Esso, pertanto, presenta varie discordanze, sia formali che sostanziali, rispetto al modello astratto del servizio di rete, che è stato descritto nei par. IV.1 e par. IV.5.

Ci riferiamo all'interfaccia di accesso a una rete per dati a commutazione di pacchetto, operante nel modo a chiamata virtuale e consideriamo lo scambio di informazione che avviene tra DTE e DCE.

Le NPDU trasferite su questa interfaccia sono allora i *pacchetti*, ottenuti dalla segmentazione delle unità di dati dello strato di trasporto, che, in questo contesto, sono chiamate *messaggi*.

La N-conessione è denominata *circuito virtuale* e può utilizzare, nell'ambito di ogni interfaccia della rete (e quindi, in particolare, dell'interfaccia di accesso DTE-DCE), un *canale logico* scelto tra quelli che l'interfaccia rende accessibili in uno schema di *multiplazione statistica*.

In questo quadro di riferimento, la Racc. X.25 consente, per lo strato di rete, la fornitura dei servizi di *chiamata virtuale* (Virtual Call o VC) e di *circuito virtuale permanente* (Permanent Virtual Circuit o PVC).

Nel primo caso la chiamata, relativa a una comunicazione di dati, si svolge attraverso le tre fasi di instaurazione del circuito virtuale, di trasferimento dei dati su tale circuito e di abbattimento del circuito stesso. Nel secondo caso si ha invece una associazione semi-permanente tra due DTE mediante un circuito virtuale ed è quindi necessaria la sola fase di trasferimento dei dati. Pertanto il servizio PVC è un caso particolare di quello VC.

Da un punto di vista architetturale, nella Racc. X.25 e per ambedue i servizi VC e PVC, la funzione di multiplazione appartiene allo strato di rete.

Conseguentemente, in questo strato, l'interfaccia DTE-DCE può supportare una molteplicità di circuiti virtuali contemporanei, tra il DTE locale e DTE remoti, ognuno su un diverso canale logico. Tali circuiti virtuali sono multiplati su un'unica DL-connessione tra il DTE locale e il proprio DCE (Fig. IV.10).

Le funzioni di N-strato nella Racc. X.25, che sono in parziale accordo con le funzionalità offerte dall'N-servizio in un ambiente OSI, possono allora essere distinte con riferimento a ogni singolo canale logico e all'intera interfaccia.

Per ogni singolo canale logico, si svolgono le funzioni riguardanti:

- 1) il trattamento della chiamata, limitatamente al caso di servizio VC e con possibilità di negoziare la fornitura di specifiche prestazioni;
- 2) il trasferimento dei dati, con controllo di flusso;
- 3) l'inoltro di dati veloci con priorità, e cioè tramite pacchetti brevi, che non sono soggetti a controllo di flusso;
- 4) il controllo della corretta sequenzialità dei pacchetti nell'ambito dell'interfaccia;
- 5) la re-inizializzazione del circuito virtuale che utilizza il canale logico considerato.

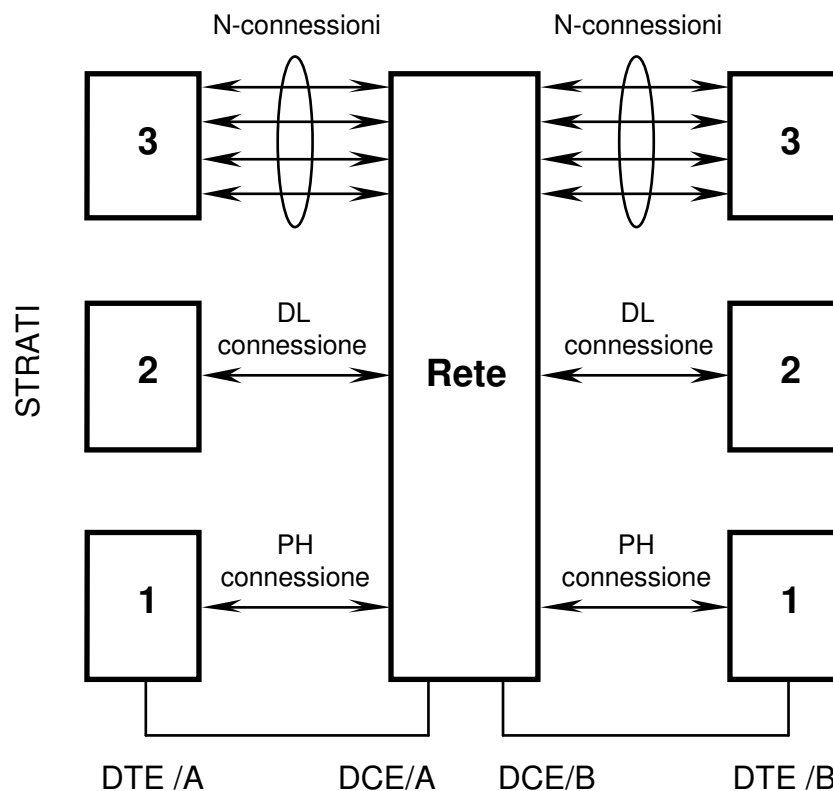


Fig. IV.10 - Multiplazione di più circuiti virtuali (N-connessioni) su un'unica DL-connessione.

Per l'intera interfaccia, si svolgono invece le seguenti ulteriori funzioni:

- a) la re-inizializzazione di tutti i circuiti virtuali precedentemente attivati sull'interfaccia;
- b) l'indirizzamento da estremo e estremo;
- c) il controllo della corretta sequenzialità dei pacchetti da estremo a estremo.

Per lo svolgimento di queste funzioni, la Racc. X.25 prevede varie *procedure* dello strato di rete. Tra queste, quelle valide solo per il servizio VC riguardano l'*instaurazione* e l'*abbattimento* del circuito virtuale in corrispondenza a ogni chiamata. Invece le procedure valide in ambedue i casi VC e PVC sono quelle di:

- A) controllo del trasferimento dei dati;
- B) controllo di flusso;
- C) interrupt;
- D) reset;
- E) restart,

con scopi che sono evidenti per le procedure di cui in A) e in B). La procedura di cui in C) si riferisce allo svolgimento della funzione di cui in 3), mentre le procedure di cui in D) e in E) riguardano le funzioni di cui in 5) ed in a), rispettivamente.

Per l'attuazione di queste procedure sono utilizzati vari tipi di pacchetti, di cui in Fig. IV.11 è fornita una possibile classificazione e di cui è opportuno conoscere il formato prima di passare all'esame dei relativi meccanismi procedurali.

In Fig. IV.11 vengono dapprima distinti i *pacchetti di controllo* e i *pacchetti di dati*. Tra i primi possono individuarsi i *pacchetti di segnalazione*, preposti alle procedure di trattamento di chiamata e ovviamente definiti per il solo servizio VC, e i *pacchetti per il controllo dei circuiti virtuali dell'interfaccia*.

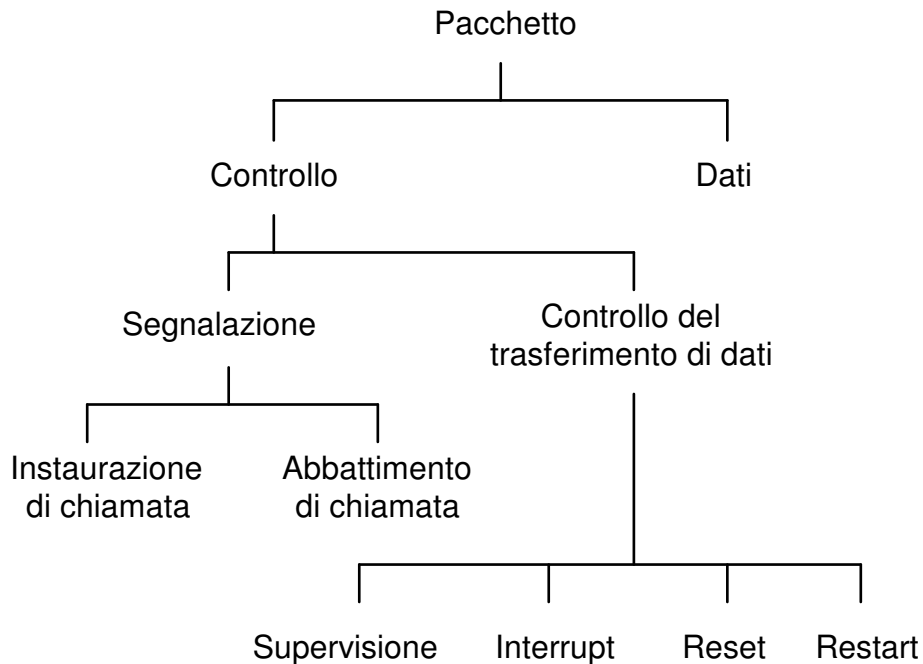


Fig. IV.11 - Classificazione dei pacchetti definiti nel protocollo X.25 di livello 3.

Nell'ambito di questi ultimi si hanno poi i *pacchetti di supervisione*, atti alla procedura di controllo del trasferimento dei dati e a quella di controllo di flusso, nonché i *pacchetti di interrupt*, i *pacchetti di reset* e i *pacchetti di restart*, impiegati nelle procedure omonime.

Le Fig. IV.12 e Fig. IV.13 mostrano i formati di due pacchetti tipici, appartenenti alle categorie dei pacchetti di controllo e dei pacchetti di dati, rispettivamente. Tali formati sono rappresentati, sulla verticale, da una sequenza di *ottetti* (byte) nell'ordine di trasmissione. Le cifre binarie di ogni ottetto, rappresentate sull'orizzontale, sono numerate da 1 a 8, anch'esse nell'ordine di trasmissione.

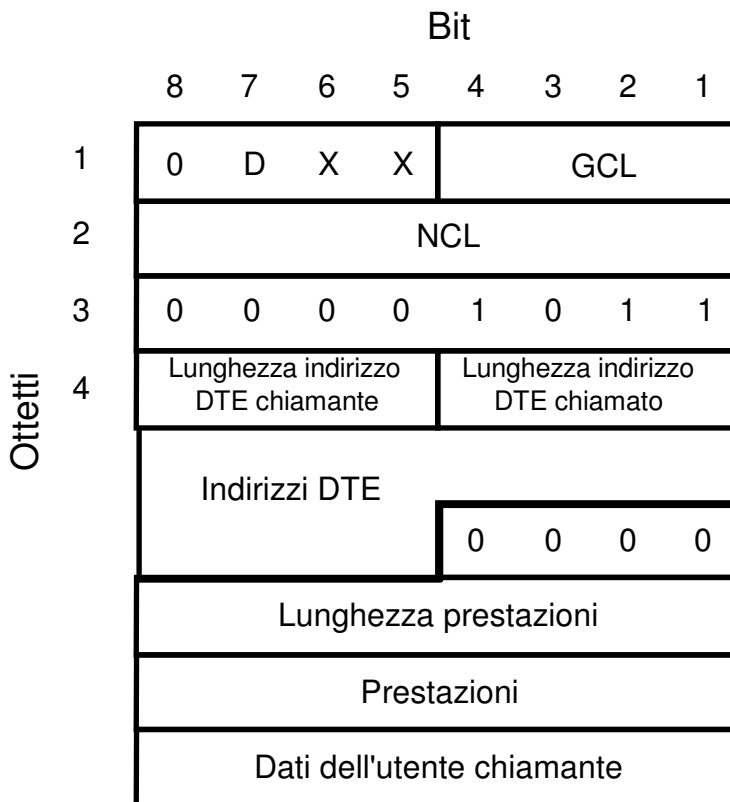
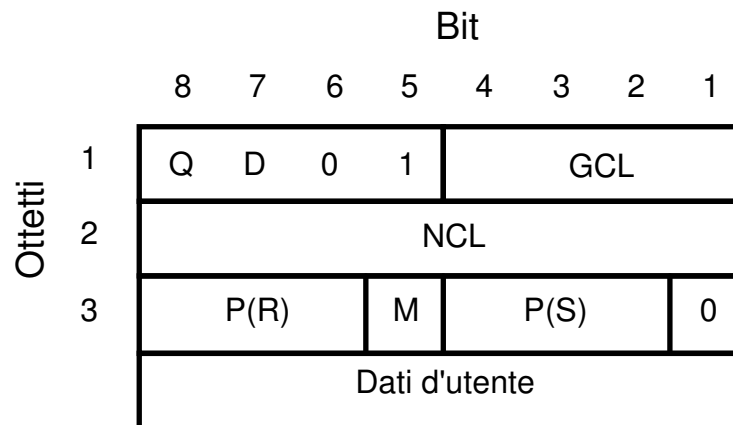


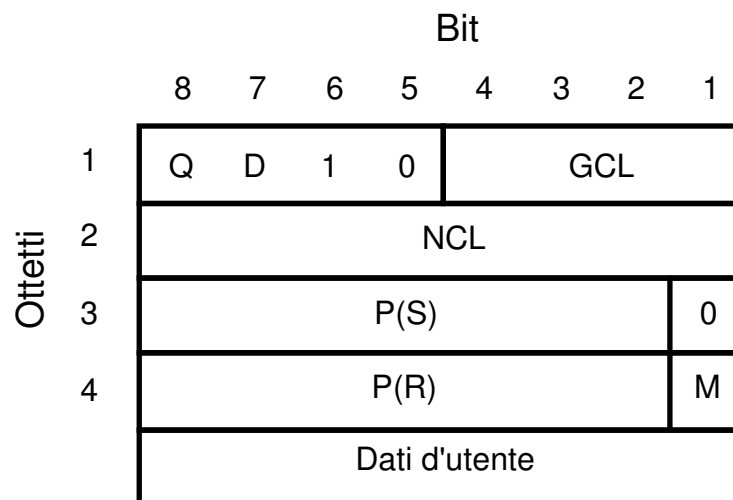
Fig. IV.12 - Formato dei pacchetti di "richiesta di chiamata" (CAR) e di "chiamata entrante" (INC) nel protocollo X.25 di livello 3.

Ciascun pacchetto è costituito, in generale, da una *intestazione* (header), sempre presente anche se con lunghezza variabile, e da un *campo informativo*, che può mancare. Questo secondo campo, se presente, contiene i *dati di utente*, risultanti dalla segmentazione dei messaggi. Nei pacchetti di dati la lunghezza massima del campo informativo è di 128 ottetti, salvo accordo diverso nella fase di instaurazione.

In ogni caso l'intestazione comprende almeno i primi tre ottetti del pacchetto, che sono suddivisi in tre campi: 1) l'*identificatore generale di formato* (General Format Identifier o *GFI*); 2) l'*indicatore di canale logico* (Logical Channel Indicator o *LCI*); 3) l'*identificatore del tipo di pacchetto* (Packet Type Identifier o *PTI*). A questi tre primi campi ne possono seguire altri per il trasferimento di *informazione aggiuntiva di intestazione*.



(a)



(b)

Fig. IV.13 - Formato dei pacchetti di dati (D) nel protocollo X.25 di livello 3, nei casi di: (a) numerazione modulo 8; (b) numerazione modulo 128.

Il campo GFI è costituito dalle ultime quattro cifre binarie del primo ottetto. Di queste le prime due sono utilizzate, nei pacchetti di dati e in quelli di supervisione, per specificare il modulo dei numeri di sequenza contenuti nel campo PTI. Questo modulo può essere uguale a 8 (configurazione 01) o a 128 (configurazione 10).

La terza cifra, indicata come *bit D*, interviene nei pacchetti di segnalazione e in quelli di dati. Infine la quarta cifra, indicata come *bit Q*, trova impiego nei soli pacchetti di dati. Nei pacchetti che non contengono uno o entrambi tra questi bit, le posizioni corrispondenti nel campo GFI sono fissate a 0.

Il bit *D* (*conferma di consegna*) è usato per indicare, con  $D = 1$ , che il DTE emittente desidera un riscontro dei dati emessi, effettuato da estremo a estremo. Se invece viene posto  $D=0$ , è richiesto un riscontro solo al livello di interfaccia locale DTE-DCE.

Il bit *Q* (*qualificatore di dati*) consente di trasferire, nell'ambito dello stesso circuito virtuale, due distinte categorie di pacchetti di dati. Per connessioni virtuali tra DTE a pacchetto (e cioè in grado di pacchettizzare l'informazione), il significato e la gestione di tale

distinzione sono lasciati agli utenti del servizio di rete.

Il campo LCI è costituito da 12 cifre binarie, le prime quattro appartenenti al primo ottetto e le altre otto al secondo. Esso consente quindi di identificare fino a 4095 canali logici distinti, dato che, nella numerazione dei canali, è esclusa la parola di codice costituita da tutti zeri.

Nel caso di servizio VC, il campo LCI è assegnato, per ogni chiamata, durante la fase di instaurazione del circuito virtuale. Nel caso, invece, di servizio PVC, il campo LCI viene pre-assegnato (ad esempio, all'atto della stipula del contratto di fornitura del servizio) e rimane lo stesso per tutta la durata del servizio.

Per ciò che riguarda, in particolare, il campo PTI, la sua prima cifra binaria discrimina tra pacchetti di controllo (determinazione "1") e pacchetti di dati (determinazione "0"). Nel caso di pacchetti di controllo, con l'esclusione dei pacchetti di supervisione, segue una parola di sette cifre binarie, che è utilizzata per codificare il tipo di pacchetto.

Nel caso, invece, dei pacchetti di dati (cfr. Fig. IV.13), le successive posizioni del campo PTI contengono i due *numeri di sequenza in trasmissione e in ricezione*, unitamente alla posizione per il bit  $M$ .

Con riferimento al pacchetto in cui sono contenuti, il numero di sequenza in trasmissione  $P(s)$  e quello in ricezione  $P(r)$  hanno lo stesso significato già chiarito per i numeri  $N(s)$  e  $N(r)$  relativi a una trama  $I$  del livello 2 del protocollo X.25. In particolare il numero  $P(s)$ , che è contenuto nei soli pacchetti di dati, indica il numero di sequenza del pacchetto nell'ambito della chiamata virtuale, mentre il numero  $P(r)$ , contenuto nei pacchetti di dati e in quelli di supervisione, indica ricezione da parte dell'emittente di tutti i pacchetti fino a quello numerato con  $P(s)$  uguale a  $P(r)-1$ .

I numeri  $P(s)$  e  $P(r)$  sono generati con un meccanismo del tutto analogo a quello dei numeri  $N(s)$  e  $N(r)$ . Il loro conteggio può essere con modulo 8 (Fig. IV.13a) o 128 (Fig. IV.13b), come già precisato in precedenza.

I numeri di sequenza  $P(s)$  e  $P(r)$  sono utilizzati per controllare la corretta sequenzialità dei pacchetti e per eseguire il controllo di flusso, secondo le modalità che verranno chiarite nel seguito.

Il bit  $M$  (*dati a seguire*) è impiegato per concatenare una sequenza di pacchetti trasferiti sulla stessa connessione virtuale. Dato che il campo informativo in ogni pacchetto di dati è limitato a una fissata lunghezza massima (che è normalmente di 128 ottetti e che può essere variata, con accordo tra le parti, durante la fase di instaurazione), i DTE che desiderano trasferire un messaggio di lunghezza maggiore di questo limite, debbono, come già detto, segmentare il messaggio e inviarlo in una cosiddetta *sequenza completa* di  $H$  pacchetti.

Tale sequenza viene identificata da  $M = 1$  nei primi  $H - 1$  pacchetti e da  $M = 0$  nel pacchetto rimanente. I pacchetti con  $M = 1$  debbono, per definizione, essere tutti della lunghezza massima consentita.

Per ciò che riguarda invece i pacchetti di segnalazione, ne esistono di due tipi (cfr. Fig. IV.11), quelli preposti alla instaurazione della chiamata virtuale e quelli utilizzati nel suo abbattimento.

Per l'instaurazione di una chiamata virtuale sono previsti quattro pacchetti di segnalazione e cioè quelli di: 1) *richiesta di chiamata* (CAll Request o *CAR*); 2) *chiamata entrante* (INcoming Call o *INC*); 3) *chiamata accettata* (Call ACcepted o *CAC*); 4) *chiamata connessa* (Call cONnected o *CON*).

Relativamente poi all'abbattimento di chiamata, si distinguono i pacchetti di: a) *richiesta di svincolo* (CLear Request o *CLR*); b) *indicazione di svincolo* (CLear Indication o *CLI*); c) *conferma di svincolo* (CLear Confirmation o *CLC*).

La Fig. IV.12 mostra il formato dei pacchetti CAR ed INC. Questi sono contraddistinti da una specifica parola di codice nel campo PTI e comprendono, come informazione addizionale di intestazione, anche:

- a) gli indirizzi del DTE chiamante e di quello chiamato, unitamente all'entità della *lunghezza* (in cifre binarie) di questi indirizzi;
- b) la specificazione di *prestazioni opzionali* richieste dal DTE chiamante, unitamente all'entità della *lunghezza* (in cifre binarie) di questa informazione;
- c) un campo per *dati di utente*, che ha lunghezza massima di 16 ottetti e che è utilizzabile dal DTE chiamante per indicare, ad esempio, a quale processo del DTE chiamato desidera essere connesso ovvero per comunicare una parola chiave di accesso a un elaboratore remoto (login pass-word).

Tra le prestazioni opzionali più significative, ottenibili su base chiamata secondo la Racc. X.25, ma non necessariamente fruibili in una rete a pacchetto che rispetta questa raccomandazione, si possono citare, a titolo di esempio: 1) la negoziazione dei parametri della procedura di controllo di flusso a finestra, ivi inclusa la lunghezza massima dei pacchetti di dati; 2) la negoziazione della classe di portata (cfr. Tab. II.2); 3) l'addebito della chiamata a carico del destinatario; 4) la prestazione di selezione veloce, di cui parleremo nel seguito; 5) la richiesta di documentazione degli addebiti di chiamata; 6) la scelta di una classe di priorità.

La prestazione di *selezione veloce* (fast select) è specificabile dal DTE chiamante nel pacchetto CAR e consente di scambiare dati contestualmente alle fasi di instaurazione e di abbattimento.

In particolare, consente al DTE di includere nei pacchetti CAR e CLR un campo di dati di utente fino a 128 ottetti. Consente altresì al DCE di inoltrare pacchetti CON e CLI, contenenti un analogo campo di dati di utente.

La prestazione prevede anche la cosiddetta *restrizione in risposta*, secondo cui, se essa è selezionata dal DTE, il DCE può rispondere esclusivamente con un pacchetto CLI. In tal

senso la prestazione di selezione veloce può vedersi come un servizio senza connessione di tipo confermato.

La Racc. X.25 prevede anche prestazioni opzionali, da definire *su base contrattuale*. Tra queste possono essere citate le prestazioni:

- 1) chiamate entranti sbarrate o chiamate uscenti sbarrate, che consentono a un DTE di agire, rispettivamente, solo come chiamante o come chiamato;
- 2) gruppo chiuso di utenti, che permette di definire, nell'ambito di una rete pubblica, una struttura di comunicazione con caratteristiche paragonabili a quelle di una rete privata; in particolare un DTE appartenente al gruppo chiuso può porsi in comunicazione con un qualunque componente del gruppo, ma non può fare altrettanto con DTE esterni;
- 3) gruppo di linea ad indirizzo unico (hunt group), che consente di raggiungere un gruppo di DTE con un solo indirizzo; 4) trasferimento di una chiamata ad altro numero, che consente una prestazione dello stesso tipo di quella offerta dalle reti telefoniche private con il servizio "seguimi".

Riassumendo, in Tab. IV.2 è riportato l'elenco dei tipi di pacchetti definiti nella Racc. X.25, con l'esclusione di quelli opzionali che non sono stati qui menzionati. I pacchetti sono raggruppati in base alle procedure in cui vengono impiegati. Inoltre i pacchetti sono distinti a seconda del verso con cui vengono trasferiti sull'interfaccia DTE-DCE. Infine viene precisata l'utilizzazione o meno dei vari pacchetti nei servizi VC e PVC.

Tipo di pacchetto		Servizio	
dal DCE al DTE	dal DTE al DCE	VC	PVC
<b>Instaurazione ed abbattimento</b>			
Chiamata entrante (INC)	Richiesta di chiamata (CAR)	X	
Chiamata connessa (CON)	Chiamata accettata (CAC)	X	
Indicazione di svincolo (CLI)	Richiesta di svincolo (CLR)	X	
Conferma di svincolo (CLC)	Conferma di svincolo (CLC)	X	
<b>Dati ed interrupt</b>			
Dati (D)	Dati (D)	X	X
Interrupt (INT)	Interrupt (INT)	X	X
Conferma di interrupt (INTC)	Conferma di interrupt (INTC)	X	X
<b>Controllo di flusso e reset</b>			
Pronto a ricevere (RR)	Pronto a ricevere (RR)	X	X
Non pronto a ricevere (RNR)	Non pronto a ricevere (RNR)	X	X
Indicazione di reset (REI)	Richiesta di reset (RES)	X	X
Conferma di reset (REC)	Conferma di reset (REC)	X	X
<b>Restart</b>			
Indicazione di restart (RTI)	Richiesta di restart (RTR)	X	X
Conferma di restart (RTC)	Conferma di restart (RTC)	X	X

Tab. IV.2 - Tipi di pacchetti definiti nel protocollo X.25 di livello 3.

Dalla Tab. IV.2 si può notare che, oltre ai pacchetti di dati (*D*), sono definiti:

- 1) due *pacchetti di interrupt*, uno *INT* (INTerrupt), che è utilizzato come richiesta e come indicazione, e l'altro *INTC* (INTerrupt Confirmation), che agisce come conferma;
- 2) due *pacchetti di supervisione*, *RR* (Receive Ready) e *RNR* (Receive Not Ready);
- 3) tre *pacchetti di reset*, e cioè *RES* (RESet Request), *REI* (REset Indication) e *REC* (REset Confirmation); mentre i primi due sono utilizzati come richiesta di reset da parte del DTE e del DCE, rispettivamente, il terzo ha il ruolo di conferma, nei confronti di ognuna delle due richieste;
- 4) tre *pacchetti di restart*, e cioè *RTR* (ResTart Request), *RTI* (ResTart Indication) e *RTC* (ResTart Confirmation); il loro impiego è analogo a quello dei pacchetti di reset.

Esaminiamo ora le procedure cominciando da quelle di instaurazione e di

abbattimento di una chiamata virtuale. Tali procedure sono illustrate nelle Fig. IV.14, Fig. IV.15 e Fig. IV.16.

Nella Fig. IV.14 vengono precisate le modalità secondo cui i pacchetti di segnalazione, preposti allo svolgimento delle due procedure e contraddistinti con gli acronimi riportati in Tab. IV.1, vengono scambiati tra DTE e DCE nell'ambito delle due interfacce *locale* e *remota*. Con questi termini indichiamo, rispettivamente, l'interfaccia relativa al DTE che inizia la procedura (*DTE iniziatore*) e quella relativa all'altro DTE (*DTE remoto*). Nel caso della procedura di instaurazione, l'iniziatore è il *DTE chiamante*, mentre quello remoto è il *DTE chiamato*.

Le Fig. IV.15 e Fig. IV.16 mostrano invece i diagrammi di stato semplificati, relativi all'interfaccia DTE-DCE e corrispondenti, rispettivamente, alle procedure di instaurazione e di abbattimento di una chiamata virtuale. Il simbolismo impiegato in tali figure è analogo a quello adoperato in precedenza per diagrammi dello stesso tipo. In particolare, per ogni transizione, sono indicati i tipi di pacchetti transitanti sull'interfaccia locale o remota e l'entità (DTE o DCE) che li ha emessi.

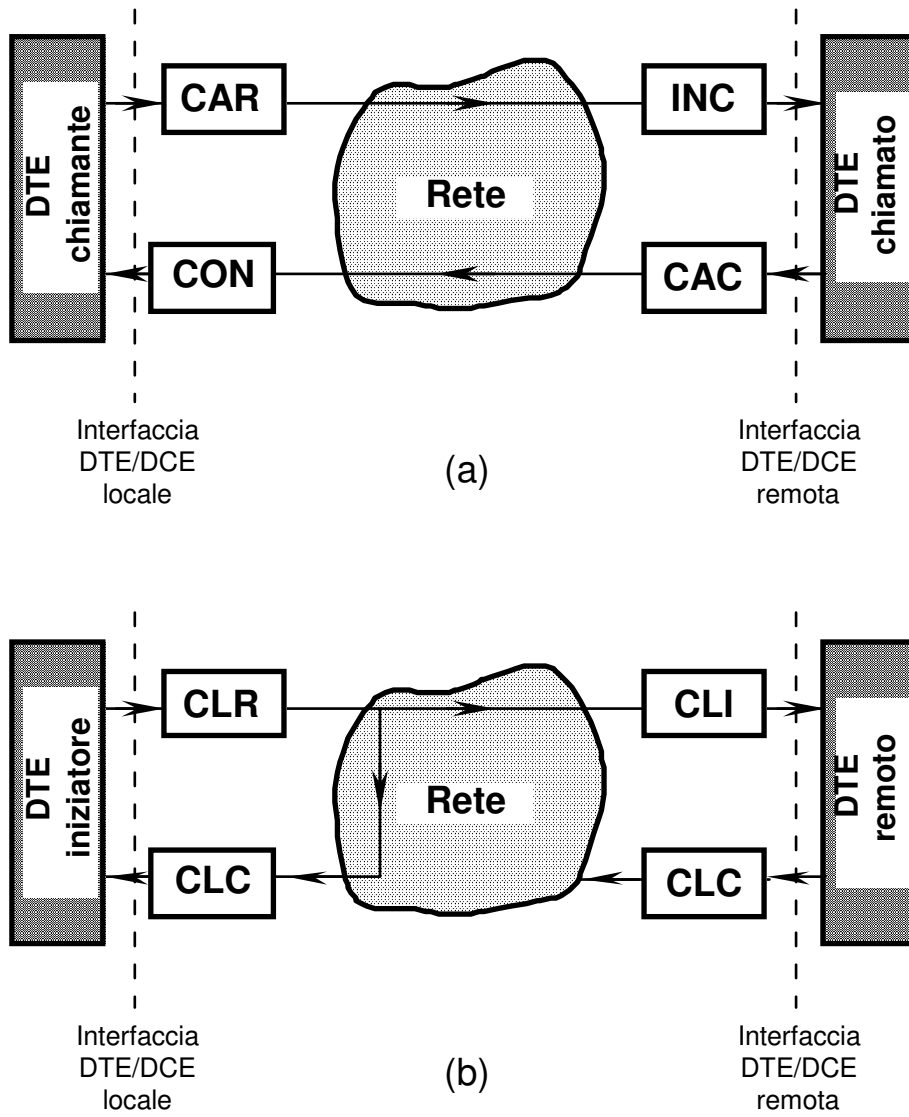


Fig. IV.14 - Procedure di instaurazione (a) e di abbattimento (b) di un circuito virtuale nel protocollo X.25 di livello 3.

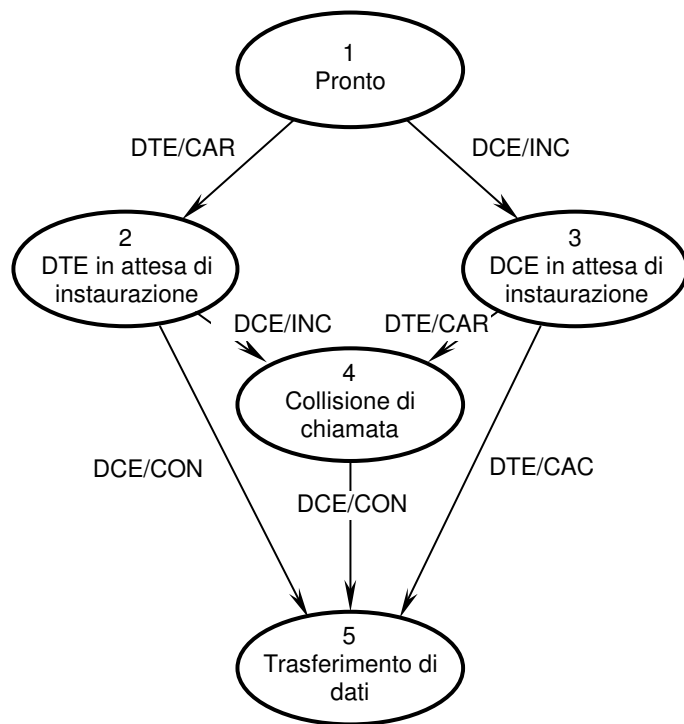


Fig. IV.15 - Diagramma semplificato di transizione di stato, relativo all'interfaccia DTE-DCE durante l'instaurazione di una chiamata virtuale nel protocollo X.25 di livello 3.

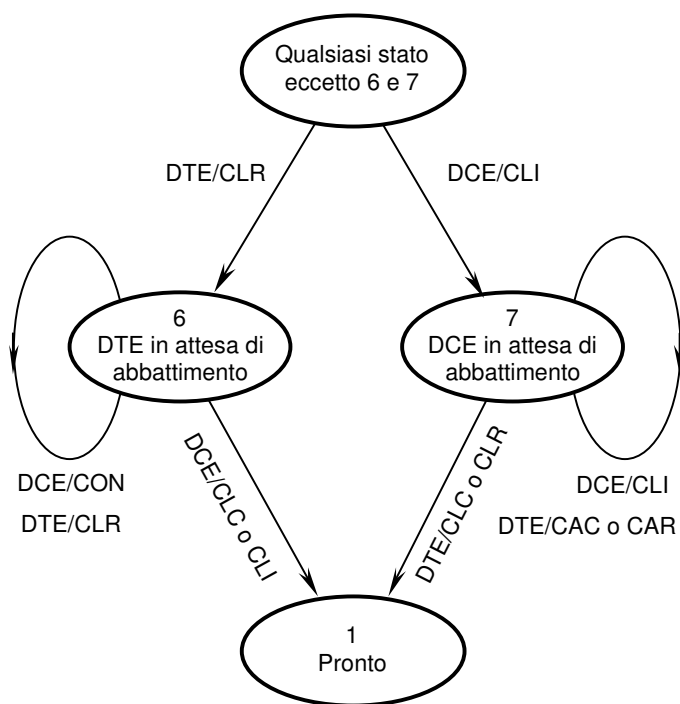


Fig. IV.16 - Diagramma semplificato di transizione di stato relativo all'interfaccia DTE-DCE durante l'abbattimento di una chiamata virtuale nel protocollo X.25 di livello 3.

Le procedure di instaurazione e di abbattimento, nonostante siano descritte in modo completo dalle Fig. IV.14, Fig. IV.15 e Fig. IV.16, vengono illustrate,

per maggiore chiarezza qui di seguito.

### *PROCEDURE DI INSTAURAZIONE E DI ABBATTIMENTO*

L'instaurazione di una chiamata virtuale ha l'obiettivo di rendere disponibile a questa chiamata una connessione virtuale, su cui i pacchetti possono essere scambiati in ambedue i versi tra le parti in comunicazione (*connessione bidirezionale*). È anche prevista, a livello opzionale, la possibilità di una *connessione unidirezionale*, ma, per brevità, ci riferiremo solo al primo caso.

La procedura evolve in due fasi principali, partendo dallo stato 1 (*Pronto*) dell'interfaccia locale e di quella remota (cfr. Fig. IV.15); in tale stato le interfacce sono disconnesse.

All'inizio della prima fase, il DTE chiamante costruisce un pacchetto CAR e lo invia alla rete tramite il proprio DCE (*DCE di origine*). Dopo questa richiesta, l'interfaccia locale passa nello stato 2 (*DTE in attesa di instaurazione*). La rete provvede allora a dare comunicazione della richiesta al DCE del DTE chiamato (*DCE di destinazione*). Ciò viene fatto con meccanismi interni, che non sono oggetto di questa procedura.

La rete completa la prima fase della procedura, emettendo verso il DTE chiamato, per il tramite del DCE di destinazione, un pacchetto INC, che viene interpretato come *indicazione* della chiamata entrante. A seguito dell'emissione di questo pacchetto, l'interfaccia remota passa nello stato 3 (*DCE in attesa di instaurazione*).

A conclusione della prima fase, dunque, le due interfacce locale e remota sono negli stati 2 e 3, rispettivamente. A questo punto comincia la seconda fase. Se il DTE chiamato è disposto ad accettare la chiamata, invia verso la rete il pacchetto CAC, come *risposta* all'indicazione fornita dal pacchetto INC. L'interfaccia remota passa allora dallo stato 3 allo stato 5 (*Trasferimento dei dati*) e la procedura di instaurazione si chiude per il DTE chiamato.

Alla ricezione del pacchetto CAC, che avviene per opera del DCE di destinazione, la rete comunica l'accettazione della chiamata al DTE chiamante, tramite il DCE di origine. Ciò si attua con l'invio di un pacchetto CON. L'interfaccia locale passa allora nello stato 5 e la procedura di instaurazione si conclude anche per il DTE chiamante quando quest'ultimo riceve il pacchetto CAC, interpretandolo come *conferma* dell'attuazione del servizio richiesto.

Nella prima fase della procedura di instaurazione vengono definiti, in particolare, i numeri di canale logico *N1* e *N2*, che identificano il circuito virtuale bidirezionale, in corso di instaurazione, sulle interfacce DTE-DCE locale e remota, rispettivamente.

Il numero *N1* è scelto dal DTE chiamante tra i numeri dei canali logici che sono liberi sull'interfaccia locale, e cioè non precedentemente assegnati ad altre chiamate in corso. Tale numero è inserito nel campo apposito del pacchetto CAR (cfr. Fig. IV.12).

Invece, la scelta del numero *N2* è di competenza del DCE di destinazione, che lo individua tra quelli liberi sull'interfaccia remota e che provvede a inserirlo nel campo apposito del pacchetto INC. Osserviamo che tale numero *N2* può essere, in generale, diverso da *N1*.

Come appare dalla Fig. IV.15, le transizioni dagli stati 2 e 3 possono portare anche nello stato 4 (*Collisione di chiamata*), a cui si perviene quando si verifica uno dei seguenti due eventi:

- 1) il DTE chiamante, dopo aver emesso il pacchetto CAR (stato 2) e a seguito di una nuova chiamata entrante nell'interfaccia locale, riceve dal DCE di origine un pacchetto INC sullo stesso canale logico prescelto per la chiamata uscente in corso di instaurazione;
- 2) il DCE di destinazione, dopo aver emesso il pacchetto INC (stato 3) e a seguito di una nuova chiamata uscente dall'interfaccia remota, riceve dal proprio DTE un pacchetto CAR

sullo stesso canale logico prescelto per la chiamata entrante in corso di instaurazione.

Lo stato di collisione viene risolto dalla rete. Infatti dallo stato 4 si esce favorendo sempre l'instaurazione della chiamata uscente. Quindi, nel caso dell'evento 1, il DCE di origine blocca la nuova chiamata entrante e trasferisce al DTE chiamante un pacchetto CON se la chiamata in corso di instaurazione è andata a buon fine.

Invece, nel caso dell'evento 2 viene bloccata la chiamata in corso di instaurazione e viene lasciata evolvere la nuova chiamata uscente. Questo risultato è ottenuto con una procedura del tutto analoga a quella relativa all'evento 1, ma con riferimento all'interfaccia remota per la chiamata in corso di instaurazione.

La procedura di abbattimento è sinteticamente illustrata dalla fig. Fig. IV.14b. Il DTE iniziatore può essere sia il chiamante che il chiamato. Esso invia alla rete un pacchetto CLR, che rappresenta la *richiesta* del servizio al livello di interfaccia locale. La rete provvede allora a darne comunicazione al DTE remoto con un pacchetto CLI. Nei pacchetti CLR e CLI è specificata la causa dell'abbattimento.

La procedura di abbattimento, per il DTE iniziatore, termina quando la rete gli invia un pacchetto CLC, mentre, per il DTE remoto, la procedura si conclude quando tale DTE invia alla rete un pacchetto CLC. Notiamo che i pacchetti CLC hanno significato esclusivamente a livello di interfaccia DTE-DCE.

L'analisi del diagramma di Fig. IV.16 consente di descrivere la procedura di abbattimento con maggiore dettaglio, in modo da tenere conto dei vari eventi che si possono presentare a livello di interfaccia DTE-DCE.

Osserviamo che lo stato di partenza della procedura può essere uno qualunque degli stati 1, 2, 3 e 4 dell'interfaccia, considerati in Fig. IV.15. Da questo stato di partenza esistono due possibili transizioni, quella verso lo stato 6 (*DTE in attesa di abbattimento*) e quella verso lo stato 7 (*DCE in attesa di abbattimento*).

Consideriamo dapprima la transizione verso lo stato 6, a cui si passa quando il DTE iniziatore emette un pacchetto CLR. Il caso di transizione dallo stato di partenza allo stato 7 verrà considerato nel seguito.

Partendo dallo stato 6 si possono verificare:

1) una transizione verso lo stato 1 (che corrisponde a una situazione di disconnessione), quando il DTE iniziatore riceve dal proprio DCE un pacchetto CLC, come conferma alla richiesta di abbattimento presentata dal DTE iniziatore, ovvero un pacchetto CLI, a seguito di una richiesta di abbattimento presentata indipendentemente da un DTE remoto;

2) una prima permanenza di stato, quando il DTE iniziatore, che avvia l'abbattimento subito dopo una richiesta di instaurazione, riceve dal proprio DCE un pacchetto CON quale conferma alla suddetta richiesta di instaurazione;

3) una seconda permanenza di stato, quando il DTE iniziatore emette una ulteriore richiesta di abbattimento (pacchetto CLR), una volta trascorso un certo intervallo temporale dalla precedente richiesta; ciò può avvenire in quanto quest'ultima non è andata a buon fine entro quell'intervallo.

Con riferimento sempre alla procedura di abbattimento e al diagramma di Fig. IV.16, analizziamo la transizione dallo stato di partenza allo stato 7.

Questa transizione avviene quando il DCE locale emette un pacchetto CLI, a seguito di una richiesta di abbattimento presentata da un DTE remoto. Partendo dallo stato 7 si possono verificare una transizione verso lo stato 1 ovvero due permanenze di stato.

La transizione dallo stato 7 allo stato 1 si verifica quando il DCE locale riceve dal

proprio DTE (DTE locale) un pacchetto CLC, quale risposta positiva alla richiesta di abbattimento presentata dal DTE remoto, ovvero un pacchetto CLR, quale richiesta di abbattimento presentata indipendentemente dal DTE locale.

La prima permanenza di stato si ha quando il DCE locale riceve dal proprio DTE pacchetti CAR o CAC, che, con riferimento alla connessione in corso di abbattimento, vengono interpretati, rispettivamente, come richiesta di connessione presentata dallo stesso DTE locale ovvero come accettazione di una richiesta di connessione presentata dal DTE remoto. Il primo caso si verifica quando il DTE locale non accetta la richiesta di abbattimento, che è stata presentata dal DTE remoto e che ha determinato il passaggio nello stato 7. Il secondo caso si presenta invece quando il DTE remoto ha avviato un abbattimento subito dopo una richiesta di instaurazione e quindi il DTE locale dà risposta favorevole a tale richiesta.

Infine la seconda causa di permanenza di stato si presenta in condizioni del tutto analoghe a quelle del secondo tipo di permanenza nello stato 6.

In Fig. IV.17 è riportato un diagramma tempo-spazio che rappresenta in modo semplificato, l'evoluzione di una chiamata virtuale per mezzo dei pacchetti scambiati in corrispondenza delle due interfacce di rete DTE-DCE.

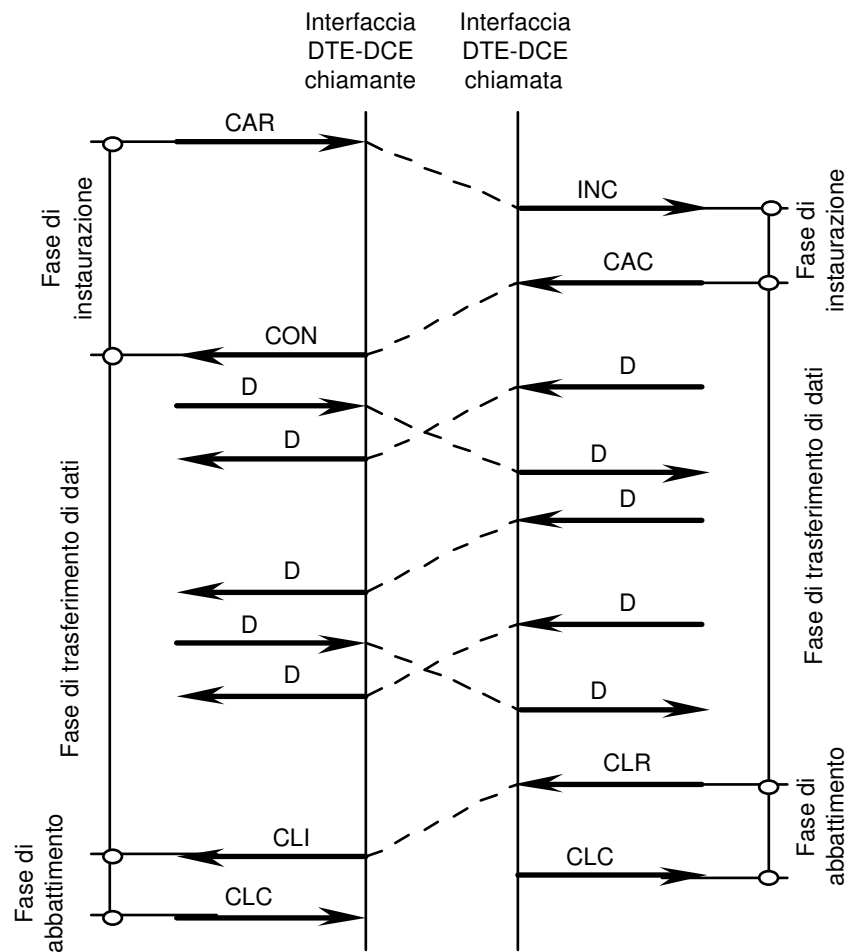


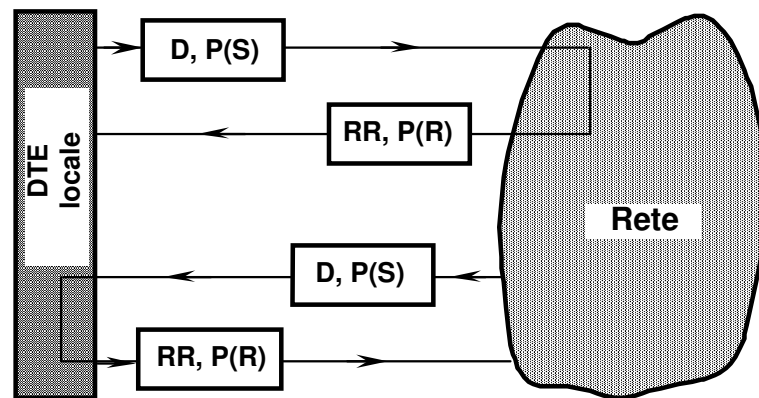
Fig. IV.17 - Evoluzione di una chiamata virtuale nel protocollo X.25 di livello 3.

Passiamo ora alle procedure di controllo del trasferimento dei dati, di controllo di flusso e di interrupt, che si applicano tutte, in maniera indipendente *su ciascun canale logico* esistente all'interfaccia DTE-DCE.

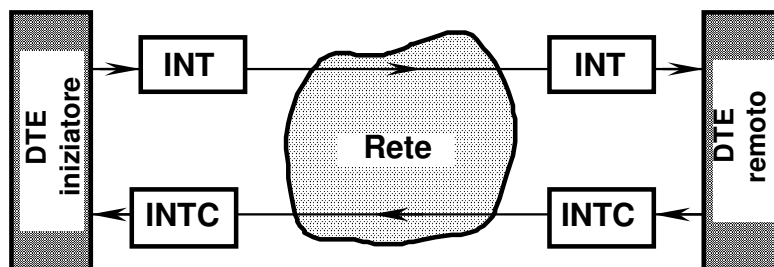
In questo ambito, ciascun pacchetto di dati, trasferito bidirezionalmente sull'interfaccia DTE-DCE in una chiamata virtuale o in un circuito virtuale permanente, è numerato sequenzialmente con i numeri  $P(s)$  e  $P(r)$ , di cui si è già chiarito il significato.

Facciamo riferimento a un'interfaccia DTE-DCE e alle due unità che si interaffacciano su di essa. Per comodità di esposizione, chiamiamo A e B queste unità, con l'intesa che, se A ha il ruolo di DTE, B deve essere il DCE, e viceversa. Indichiamo poi con  $X(s)$  e  $X(r)$  le due variabili di stato in trasmissione e in ricezione, rispettivamente. Infine denotiamo con un pedice A o B le variabili di stato e i numeri di sequenza, a seconda dell'unità a cui queste variabili e questi numeri si riferiscono.

Supponiamo ora che l'unità A emetta, verso l'unità B, pacchetti di dati. Ognuno di questi contiene il numero di sequenza in trasmissione  $P(s)_A$ , che indica il numero d'ordine del pacchetto nell'ambito della chiamata virtuale. L'unità B provvede a riscontrare positivamente questi pacchetti, utilizzando il numero di sequenza in ricezione  $P(r)_B$ , contenuto in pacchetti di dati trasferiti da B verso A ovvero in pacchetti supervisivi RR o RNR emessi appositamente da B. Tale numero  $P(r)_B$  indica la ricezione, da parte di B, di tutti i pacchetti emessi da A con un numero di sequenza  $P(s)_A$  fino a  $P(r)_B - 1$ . Questo meccanismo è illustrato schematicamente nella Fig. IV.18a.



(a)



(b)

Fig. IV.18 - Procedure di trasferimento dei dati (a) e di interrupt (b) nel protocollo X.25 di livello 3.

La disponibilità di una numerazione dei pacchetti consente di svolgere due funzioni principali: 1) il controllo della corretta sequenzialità dei pacchetti; 2) il controllo del loro flusso, mediante la *tecnica a finestra*.

Con questo ultimo termine, come si è visto nel par. III.3, si identifica una tecnica di controllo che può essere utilizzata, in generale, per la regolazione del flusso di unità informative (e quindi, in particolare, di pacchetti) tra due unità A e B, separatamente per ciascun verso di trasferimento. Essa si realizza attraverso l'adozione di due dispositivi, l'uno posto nella estremità emittente (*finestra di trasmissione*), l'altro posto nella estremità ricevente (*finestra di ricezione*). Questi dispositivi possono entrambi risiedere nella stessa unità, quando questa ha il doppio ruolo di emittente e di ricevente.

Con riferimento al caso specifico di nostro interesse, si definisce *finestra di trasmissione con larghezza  $W_T$*  l'insieme ordinato di  $W_T$  consecutivi valori dei numeri di sequenza in trasmissione  $P(s)$  relativi ai pacchetti di dati che sono autorizzati ad essere *emessi* dall'unità in cui è contenuto il dispositivo a finestra. Il più basso valore di  $P(s)$  entro la finestra è chiamato *limite inferiore della*

*finestra* e verrà qui denotato con  $L_T$ .

Se A è l'unità che contiene la finestra di trasmissione, come  $L_T$  viene assunto l'ultimo  $P(r)_B$  che ha attraversato l'interfaccia nel verso da B ad A. Di conseguenza, nelle condizioni sopra ipotizzate, l'unità A è autorizzata da B a emettere i pacchetti di dati contenenti il numero  $P(s)_A$  di valore compreso nell'intervallo

$$P(r)_B \leq P(s)_A \leq P(r)_B + W_T - 1$$

L'unità A non emette pacchetti contenenti il numero  $P(s)_A$  con valore non compreso nella propria finestra di trasmissione. Il valore del numero  $P(r)_A$ , contenuto in un pacchetto di dati che transita da A verso B deve essere maggiore o uguale al valore del numero  $P(r)_A$  presente nell'ultimo pacchetto emesso nello stesso verso e minore del valore del numero  $P(s)_B$  relativo al prossimo pacchetto di dati, che viene emesso da B verso A.

La *finestra di ricezione* è strutturata in modo del tutto analogo a quella di trasmissione, ma opera in maniera indipendente da questa. Tale finestra, se di *larghezza*  $W_R$ , è l'insieme ordinato di  $W_R$  consecutivi valori dei numeri di sequenza in trasmissione  $P(s)$ , relativi ai pacchetti di dati che sono autorizzati ad essere accettati dall'unità in cui è contenuto il dispositivo a finestra. Il più basso valore di  $P(s)$  entro la finestra è chiamato *limite inferiore della finestra* e verrà qui denotato con  $L_R$ .

Se A è l'unità che contiene la finestra di ricezione,  $L_R$  viene assunto uguale al valore dell'ultimo  $P(r)_A$  che ha attraversato l'interfaccia. Conseguentemente, nelle condizioni poste, l'unità A autorizza se stessa ad accettare, anche in modo non ordinato, i pacchetti di dati contenenti il numero  $P(s)_B$  di valore compreso nell'intervallo

$$P(r)_A \leq P(s)_B \leq P(r)_A + W_R - 1$$

I valori di  $W_T$  e di  $W_R$  sono normalmente identici ( $W_T = W_R = W$ ) e possono al massimo essere uguali a  $M-1$ , se  $P(s)$  e  $P(r)$  sono numerati modulo  $M$ . Nel caso di servizio VC, tali valori vengono fissati uguali a 2, a meno che non venga concordato un valore diverso durante la fase di instaurazione. Nel caso invece di servizio PVC, le larghezze delle due finestre sono stabilite al momento del contratto.

Per chiarire l'importanza della scelta della larghezza  $W$ , ci riferiamo alla finestra di trasmissione. Osserviamo che, in base al vincolo posto da tale finestra, possono essere emessi  $W$  pacchetti consecutivi senza riscontro. Una volta emesso l'ultimo pacchetto, non ne possono essere emessi altri se, nel

frattempo, non è pervenuto il riscontro per il primo dei  $W$  pacchetti inoltrati. Questo meccanismo può provocare un flusso intermittente di pacchetti, che riduce l'utilizzazione della connessione.

Affinché ciò non avvenga è necessario che la larghezza  $W$  della finestra sia maggiore di un *valore minimo*, che è sostanzialmente determinato dal rapporto  $\alpha$  fra il *ritardo di propagazione*  $d$  sul canale trasmissivo che supporta la connessione e la *durata di trasmissione*  $T$  di un pacchetto di dati. Si può dimostrare che tale valore minimo cresce all'aumentare del rapporto  $\alpha$ .

D'altra parte ci si può anche convincere che, all'aumentare della larghezza  $W$  della finestra e in condizioni di pieno carico, cresce il *ritardo di transito* dei pacchetti sulla connessione, e cioè il tempo necessario affinché un pacchetto venga inoltrato con successo a destinazione.

Si conclude che il valore di  $W$  deve essere opportunamente negoziato, soprattutto in connessioni che presentano un elevato ritardo di propagazione.

La Fig. IV.19 mostra, nel caso  $WT = WR = 4$ , un esempio di stato delle due finestre contenute in A, nel caso in cui il modulo di numerazione  $M$  è uguale a 8.

Esaminiamo dapprima l'evoluzione della finestra di trasmissione. Nelle condizioni ipotizzate, A è autorizzato ad emettere verso B un numero massimo di 4 pacchetti, senza che avvenga alcun riscontro. Nello stato assunto in figura, tali pacchetti sono quelli aventi  $P(s)_A$  di valore uguale a 1, 2, 3 e 4.

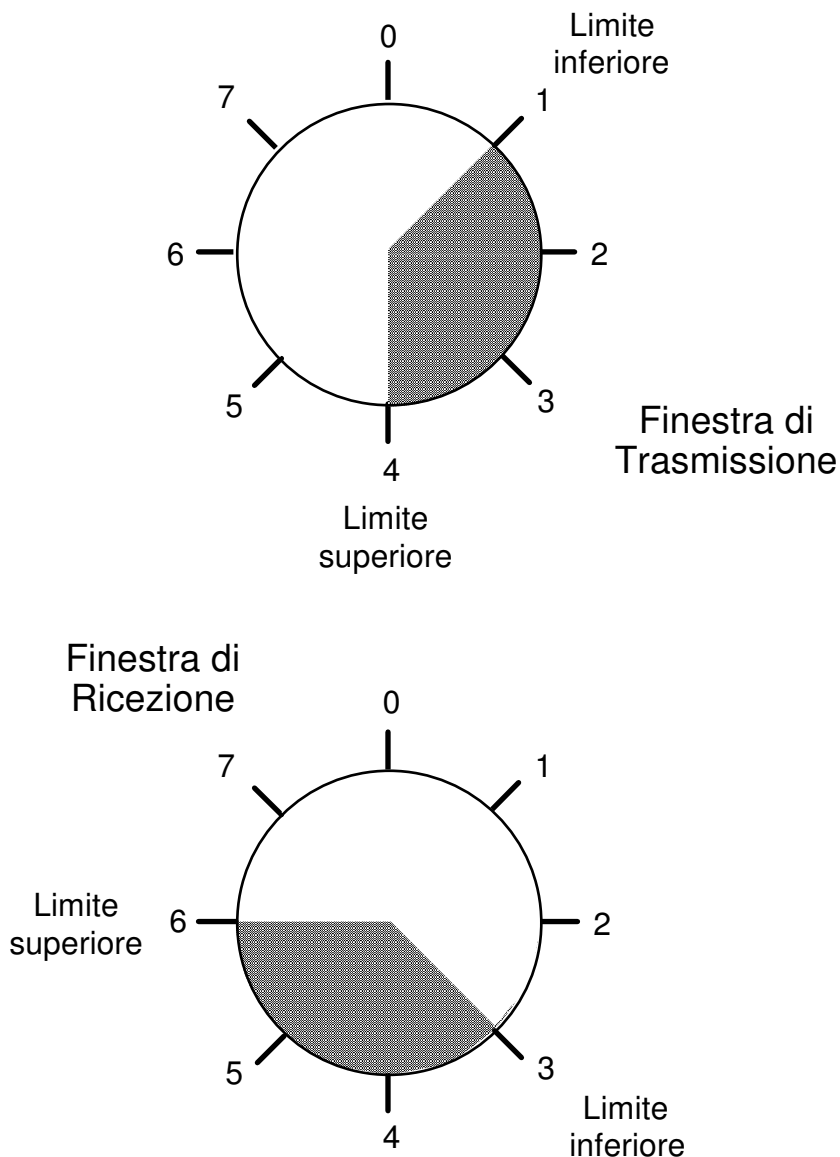


Fig. IV.19 - Esempio di finestre di trasmissione (a) e di ricezione (b) secondo il protocollo X.25 di livello 3.

Dopo aver effettuato l'emissione di questi quattro pacchetti, l'unità A si pone in attesa di riscontro. L'arrivo di un pacchetto (di dati o supervisivo), inviato da B e contenente un numero  $P(r)_B$  uguale, ad esempio, a 3 provoca l'aggiornamento della finestra. In particolare si ha uno spostamento simultaneo dei due limiti, inferiore e superiore, che incrementano il loro valore di una unità. In questa nuova situazione A può riprendere l'emissione solo per inoltrare un pacchetto di dati contenente il numero  $P(s)_A$  uguale a 5.

L'evoluzione della finestra di ricezione può essere descritta in modo analogo. L'unità A può accettare da B solo i pacchetti con  $P(s)_B$  uguale a 3, 4, 5 e 6, anche in modo non ordinato. La ricezione del pacchetto con  $P(s)_B$  uguale a

7 è autorizzata solo quando A emette almeno il riscontro positivo del pacchetto con  $P(s)_B$  uguale a 3.

Osserviamo che si può effettuare un controllo di flusso anche con l'impiego dei pacchetti supervisivi.

Per chiedere la sospensione del flusso di dati, l'unità ricevente invia a quella emittente un pacchetto RNR. Per segnalare poi la disponibilità a ricevere nuovamente, sarà sufficiente inviare un pacchetto RR.

Chiarita la procedura di controllo di flusso, esaminiamo più in dettaglio quella di controllo del trasferimento dei dati. In particolare chiariamo qual è la regola di accettazione dei pacchetti di dati da parte dell'unità B. Tale regola è basata su un criterio rivolto alla identificazione delle condizioni di fuori sequenza e alla rivelazione di quali tra queste non sono recuperabili.

Se il  $P(s)_A$  è uguale al valore attuale della variabile di stato  $X(r)_B$ , e cioè se si è in sequenza, il pacchetto è accettato.

Se invece il pacchetto è fuori sequenza, si può accertare se il pacchetto è contenuto nella finestra di ricezione quando questa è presente. In alternativa l'accertamento si effettua con riferimento alla finestra implicita sempre presente, definita da

$$P(r)_B \leq P(s)_A \leq P(r)_B + M - 2$$

ove  $P(r)_B$  rappresenta l'ultimo  $P(r)$  emesso da B. Nel caso in cui questa condizione sia soddisfatta, il pacchetto di dati viene accettato da B. In caso contrario B dichiara una *condizione di fuori sequenza non recuperabile* nell'ambito della procedura di controllo di trasferimento dei dati. Il recupero, come vedremo, è compito della procedura di reset.

Circa poi la procedura di interrupt, questa è illustrata sinteticamente in Fig. IV.18b. I pacchetti INT (cfr. Tab. IV.2) sono trasferiti, attraverso la rete, con priorità rispetto a ogni altro pacchetto di dati in attesa di trasferimento e vengono accettati da un DTE, anche se questo si è dichiarato indisponibile ad accettare pacchetti di dati (ad esempio tramite l'emissione di un pacchetto RNR).

I pacchetti INT non contengono i numeri  $P(s)$  e  $P(r)$  e quindi non sono soggetti a controllo di sequenza, né a controllo di flusso. Inoltre, l'accettazione di un pacchetto INT deve essere confermata mediante un pacchetto INTC, prima che possa essere inviato un successivo INT.

Per concludere parliamo brevemente delle procedure di reset e di restart.

La procedura di reset si usa nella fase di trasferimento dei dati per ri-inizializzare i numeri  $P(s)$  e  $P(r)$  su un dato canale logico. Essa può essere attuata, in modo indipendente, dai DTE e dalla rete a seguito di errori procedurali non altrimenti recuperabili. Tra questi vanno inserite le condizioni di fuori sequenza non recuperabili, che emergono, come abbiamo visto, dalla procedura di controllo del trasferimento dei dati.

Un DTE, quando avvia la procedura di reset di un canale logico, invia al proprio DCE un pacchetto di richiesta RES. La procedura si conclude, in questo caso, con l'invio, da parte del DCE di un pacchetto di conferma REC.

Il DCE può iniziare il reset di un canale logico a causa di un errore procedurale rivelato all'interno della rete ovvero su sollecitazione del DTE remoto. Il DCE invia al proprio DTE un pacchetto di indicazione REI e la procedura si chiude con l'invio, da parte del DTE, di un pacchetto di conferma REC.

La procedura di restart è impiegata per effettuare l'abbattimento di tutte le chiamate virtuali in corso sull'interfaccia considerata e per operare la ri-inizializzazione di tutti i canali logici impiegati per circuiti virtuali permanenti.

Anche questa procedura può essere attuata, in modo indipendente, dai DTE e dalla rete ed è usata per fronteggiare situazioni anomale particolarmente gravi (ad esempio errori procedurali o congestioni di rete) che coinvolgono più canali logici.

L'attuazione della procedura di restart avviene utilizzando i pacchetti RTR, RTI e RTC, con modalità del tutto analoghe a quelle sopra illustrate per la procedura di reset.



## V STRATO DI TRASPORTO

Lo strato di trasporto è il primo, nell'ambito dell'architettura OSI, procedendo dal basso e in un ordine gerarchico crescente, che abbia un significato *da estremo a estremo (end-to-end)*. I sotto-sistemi di questo strato sono quindi presenti nei sistemi terminali, mentre non esistono nei sistemi intermedi.

Esso ha lo scopo di completare le funzioni di trasferimento dell'informazione di utente, vista come sequenza di segmenti informativi senza rilievo semantico per i processi applicativi. Più in particolare, ha il compito di assicurare che il trasferimento da estremo a estremo avvenga con certe caratteristiche di qualità, indipendentemente dalle risorse di rete utilizzate. In parole più sintetiche, lo strato di trasporto chiude la gerarchia di funzioni legate alle esigenze di trasferimento e apre quella delle funzioni preposte alle esigenze di utilizzazione.

Per conseguire questi obiettivi lo strato di trasporto deve colmare eventuali deficienze e fluttuazioni della qualità di servizio offerta dagli strati inferiori e ottimizzarne l'uso in modo da fornire la qualità di servizio richiesta al costo minimo. Se la qualità di servizio offerta dallo strato di rete è insufficiente, lo strato di trasporto deve provvedere a migliorarla fino al grado richiesto dagli strati superiori.

Le sue principali funzioni sono, ad esempio:

- il trattamento dei propri segmenti informativi (visti come dati di utente nei confronti dello strato di rete) in modo da adattarne il formato a quello delle unità di dati del servizio di strato inferiore;
- il controllo di errore e il controllo di flusso se non previsti negli strati inferiori o, se previsti, in aggiunta a questi;
- le funzioni di multiplazione e di de-multiplazione nel caso in cui il servizio di strato sia con connessione, come quello dello strato inferiore.

Le procedure interne allo strato di trasporto garantiscono la trasparenza dell'informazione di utente da trasferire, cioè non ne limitano né il formato né il contenuto informativo. Nella definizione dello strato si presuppone la possibilità di indirizzare gli utenti del servizio da esso offerto, cioè quella di specificare con modalità indipendenti dallo strato stesso l'informazione necessaria a localizzare entità di sessione.

Nei par. V.1 V.2, che trattano rispettivamente le caratteristiche e le funzionalità del servizio, l'esposizione si può considerare valida, in linea di

principio, per strati di trasporto non necessariamente inquadrati nell'architettura OSI. Tuttavia, per fissare le idee, la trattazione farà riferimento alla Racc. X.214 (che riguarda lo strato OSI). Il successivo paragrafo (V.3.1) è rivolto ai protocolli della pila OSI.

## V.1 Caratteristiche del servizio

Come è noto, un servizio di rete (N-servizio) può essere offerto usando una notevole varietà di risorse di comunicazione; ad esempio possono essere utilizzati i seguenti mezzi:

- circuiti numerici semi-permanenti o commutati, su rete telefonica o su WAN dedicata per dati a circuito;
- circuiti virtuali permanenti o commutati su WAN a pacchetto;
- percorsi di trasferimento su LAN;
- concatenazione di mezzi appartenenti a reti diverse, ecc.

In tutti questi casi la qualità delle connessioni di rete che si possono instaurare in un ambiente "aperto" (nel senso OSI) può essere diversa. Possono variare, da rete a rete o in tempi diversi nella stessa rete, fattori importanti come la capacità di rilevare e correggere alterazioni sul flusso dei dati trasferiti (errori di trasmissione, fuori-sequenza, perdite, duplicazioni), oppure la frequenza di abbattimento delle chiamate per cause interne alla rete o all'interfaccia utente-rete.

Ad esempio, una rete a commutazione di circuito non ha alcun meccanismo logico per rivelare errori di trasmissione. Una rete a commutazione di pacchetto a circuito virtuale offre migliori garanzie contro gli errori di trasmissione, ma non protegge contro l'eventuale perdita di informazione in seguito a procedure di *re-inizializzazione*. Una rete a commutazione di pacchetto di tipo datagramma può dar luogo alla duplicazione di pacchetti. Una rete con servizio orientato alla connessione potrebbe imprevedibilmente abbattere connessioni in corso, per recuperare una situazione di congestione o per errori di protocollo (si pensi alla procedura di *restart* nel protocollo X.25).

Lo strato di trasporto ha lo scopo di mettere a disposizione delle entità di sessione (*TS-utenti*) un percorso per dati con prestazioni statistiche indipendenti dalle risorse di rete effettivamente utilizzate. Le entità di trasporto utilizzano connessioni di rete di qualità variabile di caso in caso e offrono ai TS-utenti *connessioni di trasporto* (brevemente *T-connessioni*) la cui qualità può essere prefissata in fase di instaurazione. Le procedure interne allo strato di trasporto hanno il compito di colmare la differenza tra la qualità richiesta e quella fornita

dalla connessione di rete. Vedremo come sia possibile caratterizzare quantitativamente la qualità di servizio di una connessione di trasporto.

Lo strato di trasporto si può paragonare agli ammortizzatori di un'automobile. Come il compito di questi dispositivi è assicurare un assetto costante dell'abitacolo in presenza di dislivelli e di asperità del fondo stradale, così le procedure dello strato di trasporto hanno il compito di compensare alterazioni sui dati e cadute temporanee della linea, in modo da assicurare una risorsa virtuale per il trasferimento affidabile di informazione di utente. Naturalmente, come esistono percorsi stradali talmente accidentati da mandare fuori uso gli ammortizzatori, così è inevitabile che le procedure di trasporto non siano in grado di fornire la qualità richiesta in presenza di connessioni di rete con qualità estremamente scadente. Questo evento verrà opportunamente segnalato ai TS-utenti.

## V.2 Funzionalità del servizio

Con riferimento a uno strato di trasporto con connessione, il servizio corrispondente offre ai suoi utenti i mezzi (*elementi di servizio*) per:

- a) *instaurare* connessioni di trasporto;
- b) richiedere una prefissata *qualità di servizio* per il trasferimento delle T-SDU;
- c) trasferire informazione, vista dal TS-fornitore come *unità di dati del servizio di trasporto (T-SDU)* da trattare trasparentemente;
- d) controllare, in qualità di riceventi, il ritmo con cui gli utenti corrispondenti inviano dati (*controllo di flusso*);
- e) trasferire tempestivamente T-SDU con carattere di urgenza, evitando il controllo di flusso (*dati espressi*);
- f) *abbattere* connessioni di trasporto, in modo incondizionato, eventualmente distruttivo per dati ancora in transito.

La Tab. V.1 riporta tutti gli elementi del servizio di trasporto, dei quali parleremo nelle sezioni che seguono: T-CONNECT (§ V.2.1), T-DATA e T-EXPEDITED-DATA (§ V.2.2), T-DISCONNECT (§ V.2.3).

<i>Nome del servizio</i>	<i>Tipo del servizio</i>
--------------------------	--------------------------

**Instaurazione di connessione**

T-CONNECT	confermato
-----------	------------

**Terminazione di connessione**

T-DISCONNECT	non confermato
--------------	----------------

**Trasferimento di informazione**

T-DATA	non confermato
T-EXPEDITED-DATA	non confermato

Tab. V.1 - Elenco degli elementi di servizio di trasporto

*V.2.1 Instaurazione di connessioni con prefissata qualità di servizio*

Le funzioni di cui ai punti *a)* e *b)* all'inizio del paragrafo possono essere invocate per mezzo dell'elemento T-CONNECT, che è di tipo confermato e che quindi si articola in quattro primitive: *richiesta*, *indicazione*, *risposta*, *conferma*.

La procedura di instaurazione di una connessione è avviata per iniziativa di un qualsiasi TS-utente che abbia esigenza di comunicare con un altro TS-utente. A questo fine l'entità di sessione che assume l'iniziativa esegue una primitiva T-CONNECT-richiesta, nella quale specifica diverse informazioni:

- il suo indirizzo (*indirizzo chiamante*);
- l'indirizzo dell'entità di sessione remota con cui desidera colloquiare (*indirizzo chiamato*);
- un'eventuale *opzione per dati espressi*;
- un eventuale insieme di *parametri di qualità di servizio*;
- un eventuale campo *dati di utente*.

L'aggettivo "eventuale" significa "utilizzabile a discrezione del TS-utente".

I primi tre parametri sono di immediata comprensione. Si rimanda a § 0 per la trattazione sui parametri di qualità di servizio. Illustriamo invece subito

l'utilizzazione del campo opzionale *dati di utente*. Il protocollo di sessione potrebbe prevederne l'utilizzazione per trasferire le S-PDU di richiesta di connessione (a livello di sessione). Utilizzando questa possibilità la connessione di trasporto e la connessione di sessione vengono richieste contemporaneamente.

Se la T-CONNECT-richiesta può essere soddisfatta, il TS-fornitore inoltra una T-CONNECT-indicazione al TS-utente chiamato, il quale può accettare la connessione emettendo una T-CONNECT-risposta. All'interfaccia del chiamante viene emessa una T-CONNECT-conferma. Si parla allora di *procedura di connessione eseguita con successo* e può iniziare il trasferimento dei dati. Il ciclo descritto è rappresentato nella Fig. V.1.

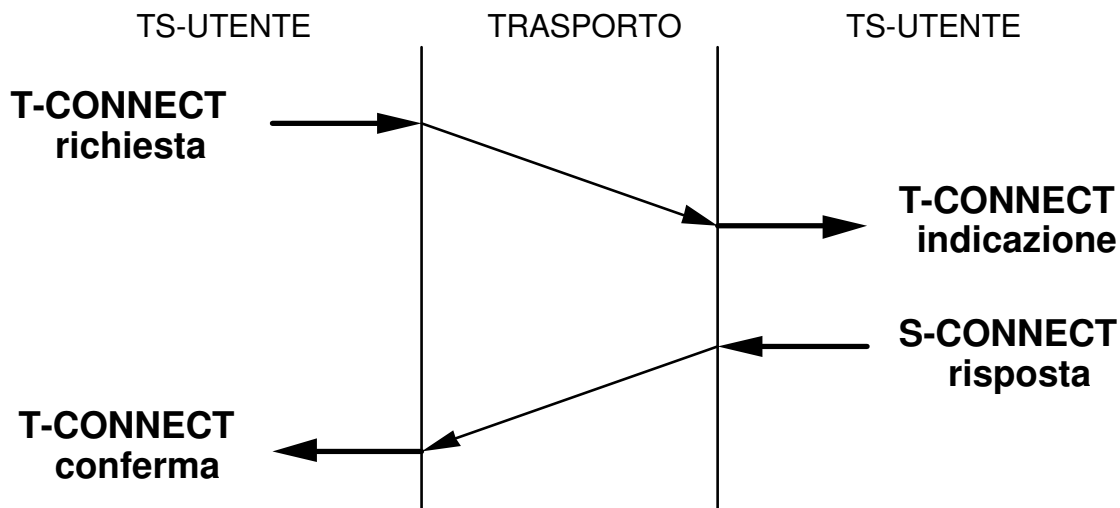


Fig. V.1- Procedura di connessione eseguita con successo

L'utente chiamato può rifiutare la connessione emettendo una T-DISCONNECT-richiesta. All'interfaccia del chiamante viene emessa una T-DISCONNECT-indicazione. Si parla allora di *procedura di connessione fallita per rifiuto del chiamato*. Il ciclo descritto è rappresentato nella Fig. V.2.

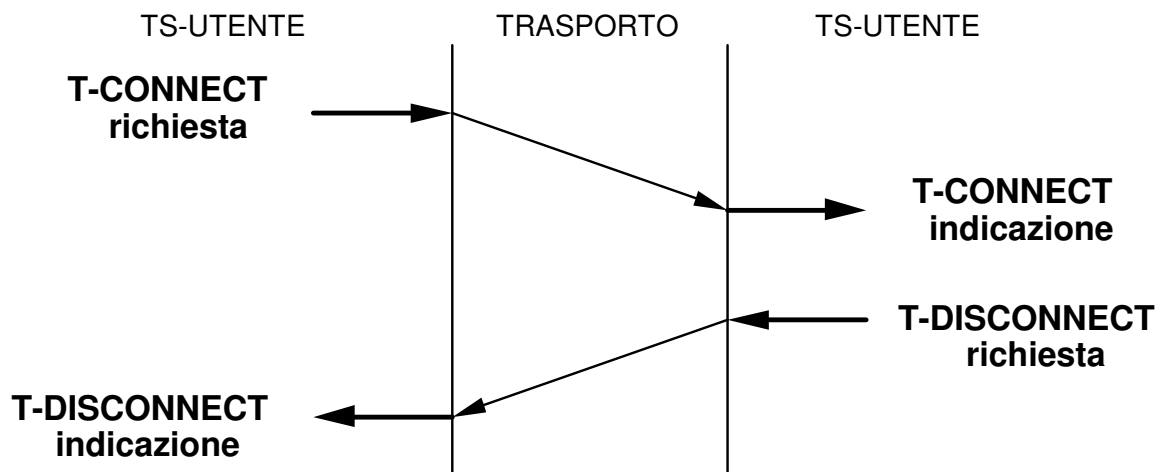


Fig. V.2 - Procedura di connessione fallita per rifiuto del chiamato

Se il servizio di rete non è disponibile, o se è impossibile garantire la qualità richiesta, il fornitore del servizio emette una T-DISCONNECT-indicazione verso l'utente chiamante. Si parla allora di *procedura di connessione fallita per rifiuto del fornitore*. Il ciclo descritto è rappresentato nella Fig. V.3.

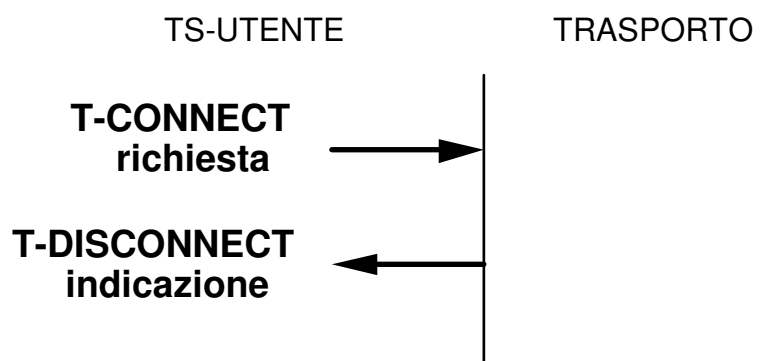


Fig. V.3 - Procedura di connessione fallita per rifiuto del fornitore

La discriminazione tra i due casi di rifiuto (del fornitore o dell'utente chiamato) è possibile per l'entità che riceve una T-CONNECT-indicazione, grazie ad un parametro *motivo del rifiuto* trasferito con la primitiva stessa.

### V.2.2 Trasferimento di dati e controllo di flusso

Le funzioni di cui ai punti *c), d), e)* all'inizio di questo paragrafo vengono invocate mediante gli elementi T-DATA, per quanto riguarda i *dati ordinari*, e T-EXPEDITED-DATA, per quanto riguarda i *dati espressi*.

Entrambi gli elementi sono di tipo non confermato, constano cioè delle sole primitive *richiesta* e *indicazione*. Tali primitive hanno un solo parametro, i

*dati di utente*. Un TS-utente che voglia inviare una stringa di bit  $\sigma$  all'altro TS-utente presenta una T-DATA-richiesta (oppure una T-EXPEDITED-DATA-richiesta) con parametro  $\sigma$ . L'altra entità riceverà una T-DATA-indicazione (oppure una T-EXPEDITED-DATA-indicazione) con identico parametro.

I dati ordinari possono avere lunghezza arbitraria e vengono trasferiti in sequenza dal T-SAP locale al T-SAP remoto. I dati espressi vengono trasferiti con priorità sui dati ordinari, ma hanno lunghezza relativamente limitata.

I TS-utenti scelgono di utilizzare uno o l'altro dei due servizi con criteri stabiliti nel loro protocollo di comunicazione.

Circa il contenuto del parametro *dati di utente*, il valore di tale parametro può rappresentare una singola S-PDU (unità di dati del protocollo di sessione), una parte di S-PDU o una sequenza di S-PDU. Negli ultimi casi il protocollo di sessione opera "segmentazione" o "concatenazione" di S-PDU, rispettivamente.

### V.2.3 *Abbattimento di una connessione*

Una T-connessione viene abbattuta, di norma, per iniziativa di una qualsiasi delle entità di sessione poste in corrispondenza dalla connessione. Questa funzione viene attivata mediante l'elemento T-DISCONNECT, che è di tipo non confermato: l'entità che richiede la disconnessione emette una primitiva T-DISCONNECT-richiesta; il fornitore del servizio emette all'altro estremo una primitiva T-DISCONNECT-indicazione. Questo ciclo, detto di *disconnessione per iniziativa di un utente*, è descritto nella Fig. V.4.

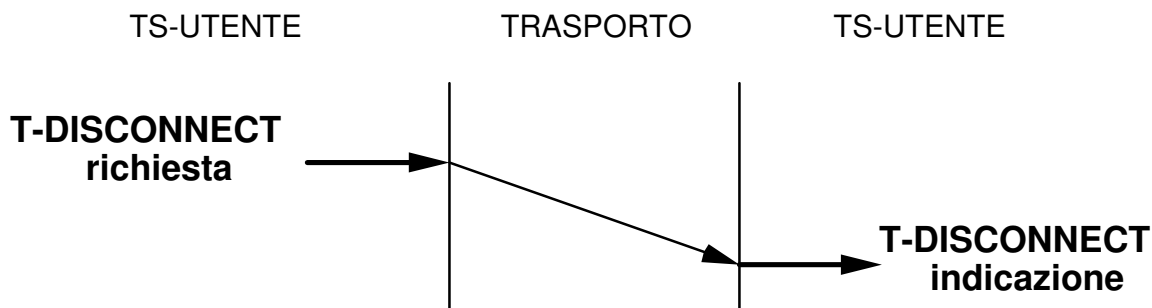


Fig. V.4 - Procedura di disconnessione per iniziativa di un utente

Una T-connessione può anche essere abbattuta, indipendentemente dalla volontà dei TS-utenti, nei casi in cui si verifichi una indisponibilità del servizio di rete o una qualità di servizio inadeguata oppure si presentino errori di protocollo, non recuperabili, all'interno dello strato di trasporto. In questo caso, il TS-fornitore emette verso entrambi gli utenti una primitiva T-DISCONNECT-indicazione. Questo ciclo, detto di *disconnessione per iniziativa*

del fornitore, è descritto nella Fig. V.5.

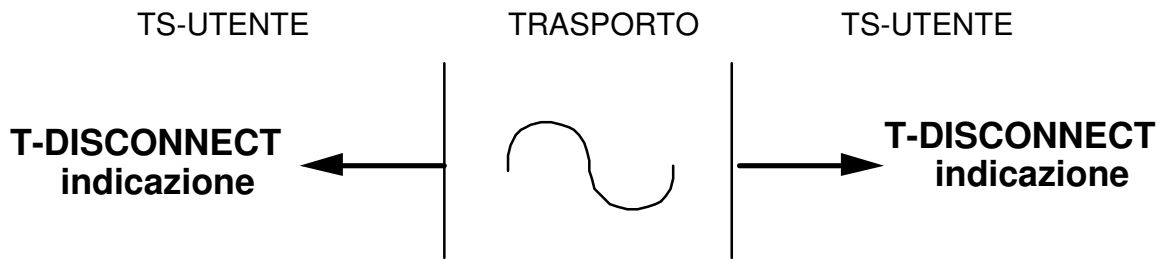


Fig. V.5 - Procedura di disconnessione per iniziativa del fornitore

Illustriamo una situazione che provoca la disconnessione per iniziativa del fornitore. Supponiamo che il protocollo di trasporto utilizzato effettui la rivelazione di errori sulle unità di dati trasferite (tieni presente però che non tutti i protocolli di trasporto hanno questa prestazione). Quando l'entità ricevente rivela un errore su un'unità di dati, essa ne richiede la ritrasmissione da parte dell'altra entità. Per evitare una situazione di stallo, l'attesa di ritrasmissione viene temporizzata. Allo scadere della temporizzazione, l'entità in attesa ripete la richiesta di ritrasmissione. Naturalmente, per evitare un ciclo operativo senza fine, occorre limitare il numero di tentativi di ritrasmissione ad un valore prefissato. Se tale valore viene superato, la situazione di errore non è più recuperabile e occorre imporre agli utenti del servizio di trasporto una disconnessione forzosa.

La discriminazione tra i due casi di disconnessione (per iniziativa dell'utente remoto e per iniziativa del fornitore) è possibile per l'entità che riceve una T-DISCONNECT-indicazione, grazie ad un parametro *motivo della disconnessione* trasferito con la primitiva stessa. Nella disconnessione per iniziativa di un utente è possibile trasferire con le primitive T-DISCONNECT un campo *dati di utente*  $\sigma$ . Questo equivale a chiedere al fornitore: "abbatti la connessione, ma prima trasferisci il blocco di dati  $\sigma$ ".

Ci si può chiedere il motivo per cui anche nella instaurazione di una connessione, regolata dall'elemento T-CONNECT, interviene l'elemento T-DISCONNECT. In altre parole, che senso ha parlare di disconnessione quando ancora non è stata instaurata la connessione? L'obiezione è più che giusta. In realtà, per "risparmiare" sul numero di elementi di trasporto, si è utilizzato un unico oggetto sintattico (T-DISCONNECT) per descrivere due realtà semantiche distinte. Tale oggetto, nell'instaurazione vale come "rifiuto" (da parte del fornitore o dell'entità chiamata), nella disconnessione vale come "abbattimento" (per iniziativa del fornitore o dell'entità remota). Vedremo più avanti come, negli strati di livello più elevato, a questi due concetti corrispondano elementi

distinti. Vedremo inoltre che, per il concetto di abbattimento, ai due casi "iniziato dal fornitore" o "iniziato dall'utente remoto" si faranno corrispondere due primitive distinte.

#### V.2.4 Parametri di qualità di servizio

Questo paragrafo non è strettamente indispensabile per la comprensione del seguito; è tuttavia estremamente formativo perché mostra come si possano definire alcuni "indici", che caratterizzano statisticamente la qualità di una T-connessione e che vengono per l'appunto chiamati *parametri di qualità di servizio* o *parametri QOS* ("Quality of Service"). Osserviamo che tali parametri non servono soltanto a misurare le prestazioni di una T-connessione, ma anche a richiedere al TS-fornitore una T-connessione con una prefissata qualità di servizio. Come già detto nel § V.2.1, ciò è possibile grazie ad un campo opzionale presente nelle primitive T-CONNECT. Ci riferiremo ai parametri presenti nella Racc.X.214, di cui richiameremo solo i seguenti:

- ritardo nell'instaurazione della T-connessione;
- tasso di insuccesso nell'instaurazione della T-connessione;
- ritardo di trasferimento;
- portata;
- tasso di errore residuo;
- tasso di insuccesso nel trasferimento;
- resilienza della T-connessione;
- ritardo nel rilascio della T-connessione.

Vediamo di descrivere brevemente i parametri appena citati.

Il *ritardo nell'instaurazione della T-connessione* è il massimo ritardo ammesso tra una richiesta T-CONNECT e la corrispondente conferma T-CONNECT.

Il *tasso di insuccesso nell'instaurazione della T-connessione* è il rapporto tra il numero totale di insuccessi nell'instaurazione di una T-connessione e il numero totale di tentativi in un campione di misura. Naturalmente, vanno portati in conto solo gli insuccessi dovuti a cause interne al fornitore, non quelli dovuti al reciproco comportamento dell'utente chiamante e dell'utente chiamato.

Il *ritardo di trasferimento* è il tempo trascorso tra una T-DATA-richiesta e la corrispondente T-DATA-indicazione, calcolato solo per T-SDU trasferite con successo. Una T-SDU è considerata trasferita con successo tra due TS-utenti quando essa perviene all'utente ricevente senza errori, in sequenza, prima del rilascio della T-connessione da parte di quest'ultimo.

La *portata (throughput)* si definisce considerando una sequenza di  $n$

T-SDU (con  $n \geq 2$ ) presentate al TS-fornitore in stretta continuità temporale, trasferite con successo e non sottoposte a controllo di flusso da parte del ricevitore. Il valore della portata è il più piccolo dei due rapporti:

(numero di ottetti di dati del TS-utente contenuti nelle ultime  $n-1$  T-SDU)/(tempo tra la prima e l'ultima T-DATA-richiesta nella sequenza)

(numero di ottetti di dati del TS-utente contenuti nelle ultime  $n-1$  T-SDU)/(tempo tra la prima e l'ultima T-DATA-indicazione nella sequenza)

La portata è specificata separatamente per ogni direzione di trasferimento.

Il *tasso di errore residuo* è il rapporto tra il numero totale di T-SDU affette da errore, perdute o duplicate e il numero totale di T-SDU trasferite attraverso l'interfaccia del TS-fornitore durante un periodo di misura.

Il *tasso di insuccesso nel trasferimento* è il rapporto tra il numero totale di insuccessi nel trasferimento e il totale di campioni di trasferimento, osservato in un periodo di misura.

La *resilienza di una T-connessione* è il complemento a uno del tasso di rilascio, su iniziativa del TS-fornitore e durante uno specificato intervallo di tempo, di una T-connessione già stabilita.

Il *ritardo nel rilascio di una T-connessione* è il massimo ritardo accettabile tra una T-DISCONNECT-richiesta iniziata dall'utente e il rilascio con successo della T-connessione. Tale ritardo è specificato indipendentemente per ogni utente. Il rilascio si considera avvenuto per un utente quando esso è in uno stato in cui può iniziare una nuova connessione.

I valori dei parametri negoziati all'instaurazione della connessione sono puramente indicativi per il TS-fornitore: servono soltanto a dimensionare automaticamente le risorse di comunicazione e la scelta di opzioni di protocollo. Non c'è garanzia che tali valori rimangano invariati per tutta la durata della connessione e che le loro variazioni siano segnalate dal TS-fornitore. Se il TS-fornitore include un modulo per la misurazione in linea dei parametri QOS, diventa possibile prevedere l'abbattimento di T-conessioni in seguito a degrado dei valori di tali parametri.

Nella normativa internazionale non si specificano né particolari valori né intervalli di valori per i parametri QOS. A titolo di curiosità possiamo riportare per i seguenti tre parametri i valori massimi codificabili nei rispettivi campi. Si prendano questi valori con il beneficio d'inventario perché non rappresentano affatto valori tipici ma valori compatibili con la tecnologia prevedibile a lungo

termine.

portata:	$2^{24}$ ottetti/secondo = 16 mega-ottetti/secondo
tasso errore residuo:	$10^{-256}$ (valore tipico $10^{-16}$ )
ritardo di trasferimento:	$2^{16}$ millisecondi $\approx$ 65 secondi

#### V.2.5 Metodi di specifica del servizio di trasporto

Il servizio di trasporto può essere specificato graficamente, come abbiamo fatto nelle Fig. V.1, Fig. V.2, Fig. V.3, Fig. V.4 e Fig. V.5, mediante *diagrammi temporali*. Questa rappresentazione enumera *esplicitamente* tutte le sequenze di primitive che è possibile osservare all'interfaccia tra lo strato di trasporto e lo strato di sessione. Ogni sequenza mette in relazione di causa-effetto le primitive osservabili ad un estremo con altre osservabili allo stesso estremo e con altre osservabili all'altro estremo. Ad esempio, è possibile dire che ad una T-CONNECT-richiesta ad un estremo segue in stretta successione una T-CONNECT-indicazione all'altro estremo.

Una rappresentazione più compatta è costituita da un *diagramma di stato dell'interfaccia utente-servizio*, che enumera *implicitamente* tutte le sequenze di primitive che è possibile osservare in un qualsiasi punto di accesso al servizio. Per il servizio di trasporto, il diagramma di stato è quello di Fig. V.6.

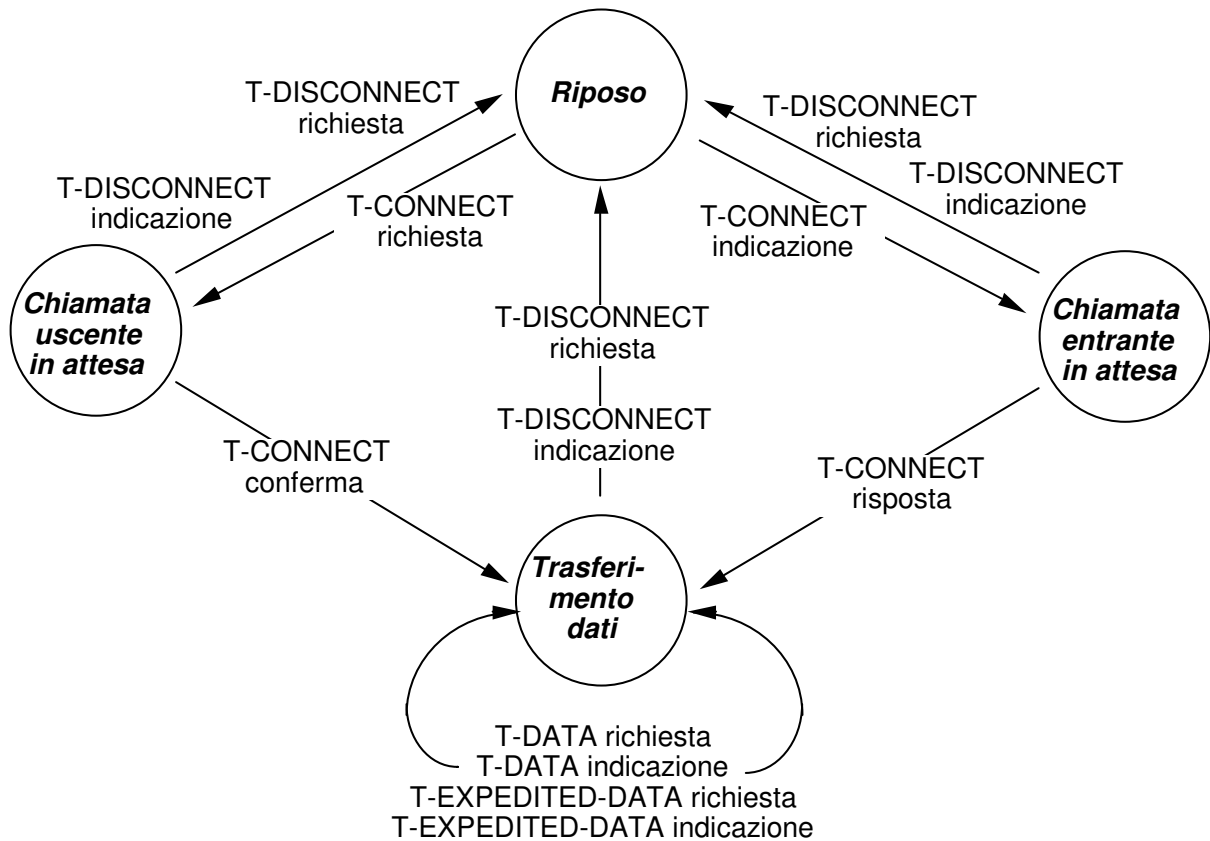


Fig. V.6 - Diagramma di stato del servizio di trasporto

Con riferimento a nozioni elementari di teoria degli automi e dei linguaggi formali, potrebbe essere utile paragonare questi diagrammi a quelli usati per rappresentare automi generatori/riconoscitori di stringhe di un linguaggio regolare. Il diagramma di Fig. V.7 rappresenta il generatore delle stringhe  $ab^nc$  con  $n \geq 0$ , se  $S_0$  è lo stato iniziale.

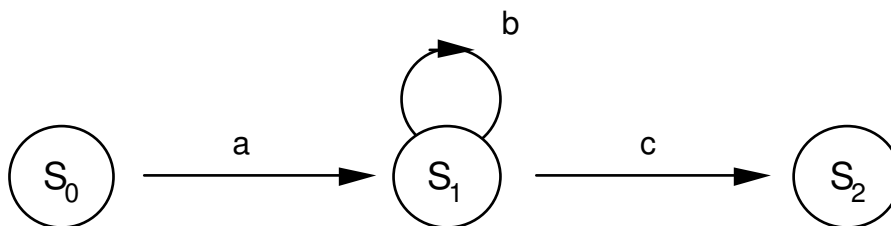


Fig. V.7 - Un esempio di automa generatore di stringhe

Si osserva che in un diagramma di stato dell'interfaccia non è rappresentata alcuna relazione di causa-effetto: tutte le primitive innescano transizioni di stato perché, rispetto alla rappresentazione, sia l'utente sia il fornitore del servizio sono considerati ambiente esterno.

Né i diagrammi temporali né i diagrammi di stato, di cui abbiamo appena

parlato, risultano soddisfacenti per una descrizione completa e compatta del servizio. Se vogliamo ottenere una specifica formalmente completa, dobbiamo costruire un modello astratto del fornitore del servizio che descriva le relazioni tra le primitive trasferite ad ognuno degli estremi in una qualsiasi sequenza di interazioni. È possibile esprimere un tale modello utilizzando speciali tecniche formali, che sono state appositamente definite per i processi di comunicazione. Tecniche di questo tipo sono ad esempio le *Reti di Petri*, il linguaggio *SDL* ("Specification and Description Language"), normalizzato dal CCITT, i linguaggi *LOTOS* ("Language Of Temporal Ordering Specification") ed *Estelle* ("Extended Transition Language"), normalizzati dall'ISO.

È opportuno notare che tali tecniche si prestano alla descrizione di un qualsiasi servizio e di un qualunque protocollo dell'architettura OSI. La padronanza di almeno una di tali tecniche è di indubbio vantaggio per un approfondire le tematiche oggetto di questa trattazione.

### **V.3      Protocolli di strato di trasporto**

Lo scopo principale di un protocollo di trasporto è, per quanto visto sul servizio di trasporto, quello di fornire T-conessioni con un prefissato grado di servizio e con determinate prestazioni, compensando convenientemente grado di servizio e prestazioni delle N-conessioni. Qualità di servizio e prestazioni sono dunque i due fattori determinanti per il progetto di un protocollo di trasporto.

Un primo dato di progetto da prendere in considerazione è rappresentato dal comportamento della rete nei confronti degli errori di trasferimento. Per compensare tali errori può essere necessario prevedere meccanismi di *controllo di errore* e di *controllo di sequenza* delle unità di dati.

Tali meccanismi sono concettualmente simili a quelli dello strato di collegamento, ma hanno una diversa "gittata": nello strato di collegamento essi si applicano alle unità di dati su ogni ramo del percorso seguito nella rete dall'origine alla destinazione; nello strato di trasporto si applicano al flusso dati risultante tra due terminali di una rete di comunicazione aperta, indipendentemente dai percorsi utilizzati dalle singole unità di dati.

Un secondo dato progettuale rilevante è la resilienza di una tipica N-conessione messa a disposizione dalla rete, cioè la probabilità (a posteriori) che essa non cada durante il suo utilizzo da parte di una T-conessione. Se tale probabilità non è soddisfacente (nella normativa, tuttavia, non viene precisato alcun valore di soglia), conviene prevedere meccanismi di *recupero* che

consentano la prosecuzione di una T-connessione in caso di caduta della N-connessione da essa utilizzata.

Per quanto riguarda gli obiettivi di prestazioni, occorre tener presente che le N-conessioni sono caratterizzate da una certa *portata* la quale può essere sovrabbondante o insufficiente rispetto a quella richiesta dai TS-utenti. Nel primo caso, per ottimizzare i costi di una determinata N-connessione, conviene convogliare in essa i flussi relativi a più T-conessioni, cioè è consigliabile effettuare una *multiplazione*. Nel secondo caso, per soddisfare le esigenze dei TS-utenti occorre operare una *suddivisione* del flusso relativo ad una T-connessione in un certo numero di sotto-flussi attraverso altrettante N-conessioni.

Nel caso della multiplazione è poi opportuno un *controllo di flusso* sulle singole T-conessioni per far sì che l'unica N-connessione sia condivisa in maniera statisticamente equilibrata.

Un altro requisito del servizio di trasporto è il trasferimento di dati di lunghezza arbitraria. Poiché la lunghezza dei dati accettati dal servizio di rete è limitata, può essere necessario frazionare le T-PDU prima di affidarle al servizio di rete, cioè può convenire operare una *frammentazione*. Viceversa, per ottimizzare l'uso delle risorse di rete si potrebbe trasferire una sequenza di T-PDU sufficientemente "corte" in un'unica N-SDU, cioè si potrebbe effettuare una *concatenazione*.

### V.3.1      *Protocolli OSI*

Tener conto di tutti i fattori sopra elencati significherebbe definire protocolli con funzionalità estremamente complesse. Risulta dunque più conveniente definire, nell'ambito di una stessa architettura di comunicazione, vari protocolli di trasporto dotati di diversi gradi di complessità, ipotizzando reti con differenti qualità di servizio e mirando a diversi livelli di prestazioni. In particolare, nell'architettura OSI, si è convenuto che tali protocolli siano omogenei per i formati delle unità di dati e per un insieme di procedure di base. Si è così giunti a parlare di "varianti" o di *classi di protocollo*.

In questa trattazione, per maggiore chiarezza e per motivi di concretezza, faremo riferimento allo standard internazionale descritto, indifferentemente, nella Raccomandazione CCITT X.224 e nella norma ISO IS-8073.

La trattazione prosegue con i seguenti argomenti: tipi di reti e classi di protocollo, unità di dati del protocollo di trasporto, procedure comuni a tutte le classi, procedure specifiche delle varie classi.

### *TIPY DI RETI E CLASSI DI PROTOCOLLO*

Allo scopo di definire varie classi di protocollo, la normativa ITU-T e quella ISO distinguono tre *tipi di reti*, in base ad alcuni parametri di qualità delle connessioni che esse offrono. Precisamente, vengono distinte:

- *reti di tipo A*, quando il *tasso di errore residuo* (cioè la misura percentuale degli errori non segnalati dal servizio di rete) e il *tasso di errore segnalato* (cioè la misura percentuale degli errori dovuti a RESET o RESTART notificati dalla rete) sono entrambi accettabili (i valori numerici delle soglie di accettabilità non sono definiti nella normativa);
- *reti di tipo B*, quando risulta accettabile il tasso di errore residuo ma non il tasso di errore segnalato;
- *reti di tipo C*, quando risulta inaccettabile il tasso di errore residuo (a prescindere dall'accettabilità del tasso di errore segnalato).

Un altro elemento di classificazione del protocollo di trasporto è rappresentato dall'eventuale esigenza di moltiplicazione.

In base alle precedenti considerazioni, si distinguono cinque classi di protocollo: *C0 (classe semplice)*, *C1 (classe semplice con recupero)*, *C2 (classe con moltiplicazione)*, *C3 (classe con recupero e con moltiplicazione)* e *C4 (classe con rivelazione e con recupero di errore)*.

La Fig. V.8 illustra come ogni classe sia concepita per un tipo di rete. Su una rete di tipo A e su una rete di tipo B sono previste rispettivamente classi senza moltiplicazione (C0 e C1) e classi con moltiplicazione (C2 e C3). Su una rete di tipo C è sempre possibile la moltiplicazione ed è prevista un'unica classe (C4). La medesima figura mostra qualitativamente la diversa complessità funzionale delle classi e suggerisce l'idea che tale complessità sia in misura inversa al livello di prestazioni del servizio di rete.

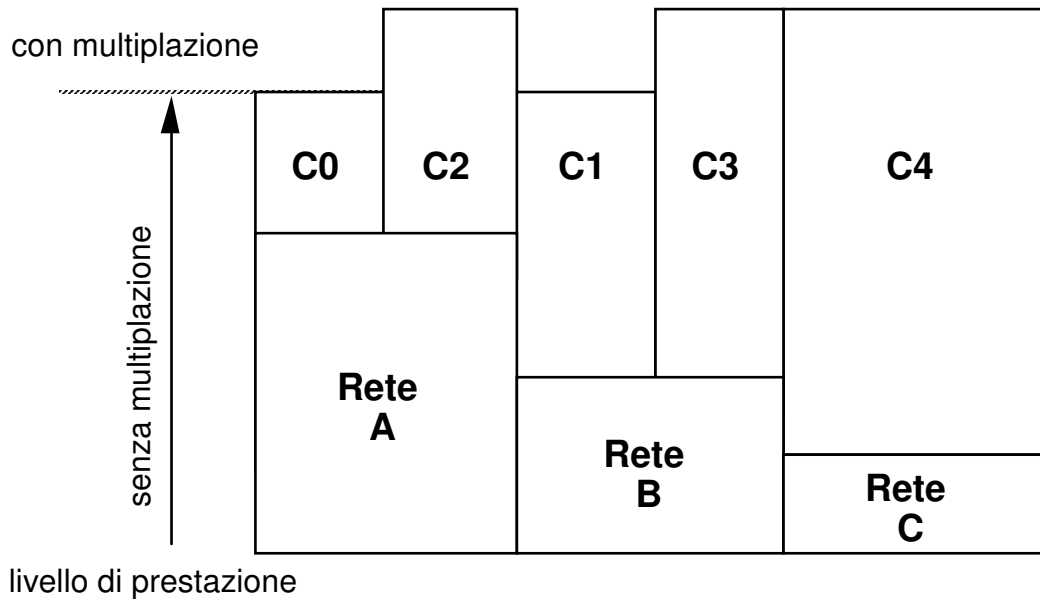
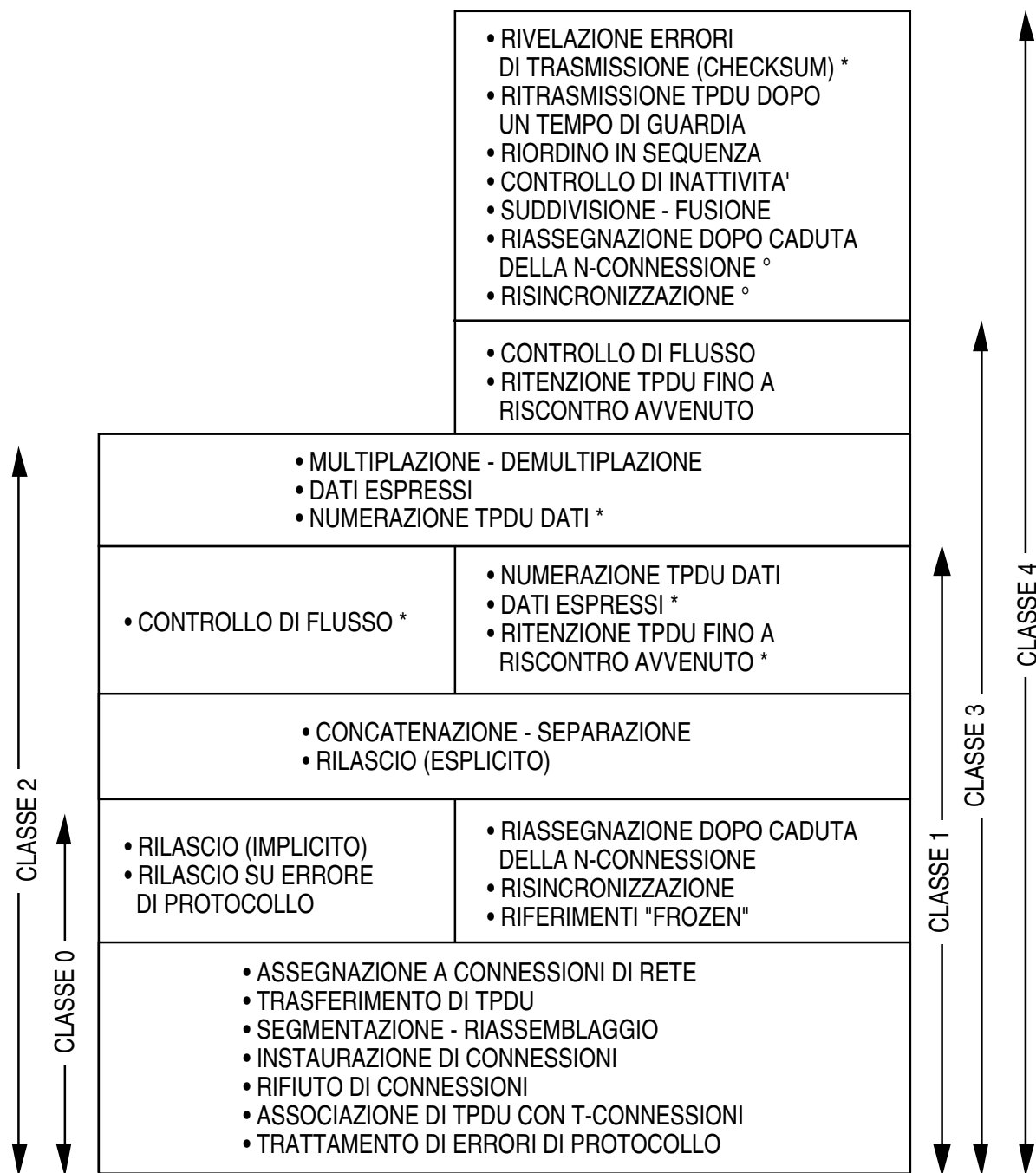


Fig. V.8 - Relazione tra tipi di reti e classi di protocollo di trasporto

La Fig. V.9 offre un quadro d'insieme delle procedure previste per le varie classi di protocollo di trasporto. Come possiamo vedere, esiste un nucleo di procedure comuni a tutte le classi; le altre procedure sono distribuite in modo tale che si abbiano due relazioni di inclusione:  $C2 \supset C0$  e  $C4 \supset C3 \supset C1$ . Di quasi tutte le procedure citate daremo un cenno nel seguito del paragrafo.



\* Funzione soggetta a negoziazione fra le entità

° Funzione realizzata con procedure diverse da quelle usate nelle classi precedenti.

Fig. V.9 - Funzioni di trasporto e classi di protocollo

**UNITÀ DI DATI DI PROTOCOLLO**

La Tab. V.2 elenca le *unità di dati del protocollo di trasporto (T-PDU)*. Osserviamo che non tutte le T-PDU sono usate in tutte le classi. Aggiungiamo che quelle usate in tutte le classi non sempre sono usate con le stesse modalità.

Sigla	Significato	Validità per classe				
		0	1	2	3	4
CR	Richiesta Connessione	x	x	x	x	x
CC	Conferma Connessione	x	x	x	x	x
DR	Richiesta Disconnessione	x	x	x	x	x
DC	Conferma Disconnessione	-	x	x	x	x
DT	Dati	x	x	x	x	x
ED	Dati "Expedited" (veloci)	-	x	nf	x	x
AK	"Acknowledgement" (riscontro)	-	nc	nf	x	x
EA	"Expedited Acknowledgement"	-	x	nf	x	-
RJ	"Reject" (Rigetto)	-	x	-	x	-
ER	Errore di procedura	X	X	X	X	X
NOTE:		x        valida -        non valida nf       valida solo in presenza dell'opzione controllo di flusso nc       valida solo con certe condizioni di negoziazione				

Tab. V.2 - Elenco delle unità dati di protocollo di trasporto

La Fig. V.10 illustra il formato generale di una T-PDU. Il campo LI, detto *identificatore di lunghezza*, indica il numero di ottetti che compongono la T-PDU. La parte variabile dipende dalla classe e da alcune opzioni stabilite in fase di apertura della T-connessione.

LI	parte fissa	parte variabile	campo dati
----	----------------	--------------------	---------------

Fig. V.10 - Formato generale di una T-PDU

**PROCEDURE COMUNI A TUTTE LE CLASSI.**

L'elenco delle procedure comuni a tutte le classi può essere desunto dalla già citata Fig. V.9. Accenniamo ad ognuna di esse.

*Assegnazione a connessioni di rete.* Questa procedura viene invocata ogni volta che occorre una nuova N-connessione per supportare una T-connessione. In assenza di moltiplicazione (classi C0 e C1), essa viene invocata ad ogni richiesta di T-connessione a meno che non esista già una N-connessione aperta

ma non assegnata. In presenza di multiplazione (classi C2, C3 e C4), la procedura viene invocata in seguito ad una richiesta di T-connessione solo se non esiste una N-connessione di rete già aperta e in grado di supportare la T-connessione da instaurare.

Ad esempio, siano *A* e *B* due generici indirizzi di trasporto, *a* e *b* due corrispondenti indirizzi di rete. Intendendo con la notazione *sap-x ! p* dire "al *sap-x* si osserva il trasferimento della primitiva *p*", descriviamo la procedura di assegnazione di T-conessioni a connessioni di rete come serie di eventi osservabili ai T-SAP e agli N-SAP. Supponendo che la richiesta di connessione parta dal T-SAP-A, abbiamo:

T-SAP-A ! T-CONNECT-richiesta

N-SAP-a ! N-CONNECT-richiesta

N-SAP-b ! N-CONNECT-indicazione

N-SAP-b ! N-CONNECT-risposta

N-SAP-a ! N-CONNECT-conferma

Questo esempio mostra come un elemento di procedura non implica necessariamente lo scambio di unità di dati di protocollo.

*Trasferimento di T-PDU.* L'invio di una T-PDU consiste nel trasferire una primitiva N-DATA-richiesta con parametro *dati di utente* uguale alla stringa di bit che costituisce la T-PDU. La ricezione di una T-PDU consiste nell'effettuare una primitiva N-DATA-indicazione e nel prelevare da questa il valore del parametro *dati di utente*. L'invio (oppure la ricezione) di T-PDU può avvenire anche attraverso la primitiva N-DATA-EXPEDITED-richiesta (oppure N-DATA-EXPEDITED-indicazione). La scelta del tipo di "veicolo" è stabilita per ogni T-PDU dalla classe selezionata.

Instaurazione di connessioni di trasporto. Viene invocata in seguito ad una richiesta di T-connessione. Se non esiste una connessione di rete già disponibile, si assegna prima una N-connessione (vedi procedura precedente). Segue, quindi, la negoziazione dei parametri della T-connessione: innanzitutto, classe preferita e, compatibilmente con la classe, una serie di opzioni primarie e una serie di opzioni secondarie.

La Fig. V.11 illustra il formato delle parti fisse delle unità CR e CC. I campi

SRC-REF e DST-REF, detti riferimento origine e riferimento destinazione, servono per identificare univocamente la T-connessione da parte del chiamante e da parte del chiamato. Il campo CDT è usato nelle classi con multiplazione per gestire il controllo di flusso, mentre è nullo in classe 0 e classe 1. Il campo classe e opzioni contiene la classe che il chiamante intende utilizzare e le opzioni primarie (ad esempio l'opzione per il controllo di flusso).

1110	CDT	DST-REF 00000000 00000000	SRC-REF	CLASSE, OPZIONI
------	-----	------------------------------	---------	--------------------

parte fissa della tpdu CR

1101	CDT	DST-REF	SRC-REF	CLASSE, OPZIONI
------	-----	---------	---------	--------------------

parte fissa della tpdu CC

Fig. V.11 - Formato delle parti fisse delle T-PDU CR e CC

La parte variabile dipende dal contenuto del campo classe e opzioni e contiene le opzioni secondarie. Opzioni e parametri della CR e della CC sono riportati nella Tab. V.3, la quale illustra altresì come vengono "negoziati" i vari parametri. Lo spazio a disposizione non ci consente ulteriori dettagli. Basta solo tener presente che la CC può eventualmente restringere le opzioni della CR. In particolare, può indicare una classe più semplice di quella richiesta. Ciò è dovuto all'eventualità che l'entità chiamata non supporti la classe richiesta.

Parametro	componenti	significato o valore nella CR	classi di applicabilità	significato/ valore nella CC
CDT (credito iniziale controllo di flusso)		0000 (C0, C1)	C2-C4	0000 (C0, C1)
DST-REF (riferimento destinazione)		00000000 00000000	tutte	indicato
SRC-REF (riferimento origine)		indicato	tutte	replicato
classi e opzioni	<i>classe</i>	preferito	v.note	scelto
	<i>formato esteso</i>	preferito	C2-C4	scelto
	<i>controllo di flusso</i>	preferito	C2	scelto
TSAP-ID chiamante		fornito dall'utente	tutte	replicato
T-SAP-ID chiamato		fornito dall'utente	tutte	replicato
lunghezza TPDU		proposto	tutte	scelto
numero di versione		indicato	C1-C4	-
riservatezza		fornito dall'utente	C1-C4	fornito dall'utente
checksum		calcolato	C4	calcolato
altre opzioni	<i>N-dati veloci</i>	preferito	C1	scelto
	<i>N-conferma ricezione</i>	preferito	C1	scelto
	<i>non uso checksum</i>	preferito	C4	scelto
	<i>espressi trasporto</i>	preferito	C1-C4	scelto
classi di protocollo alternative		lista di classi	C1-C4	-
tempo di riscontro		indicato	C4	indicato
qualità di servizio	<i>portata</i>	proposto	C1-C4	scelto
	<i>tasso errore residuo</i>	proposto		scelto
	<i>priorità</i>	proposto		scelto
	<i>ritardo di transito</i>	proposto		scelto
tempo di riassegnazione		indicato	C1, C3	-
dati utente		fornito dall'utente	C1-C4	fornito dall'utente

NOTA: la "classe" costituisce il parametro di partenza per il criterio "Classi di applicabilità"

VALORE PREFERITO: apre la negoziazione in un insieme discreto di valori  
 VALORE PROPOSTO: apre la negoziazione in un insieme continuo di valori  
 VALORE SCELTO: chiude la negoziazione ("prendere o lasciare")  
 VALORE REPLICATO: identico a quello della corrispondente CR

VALORE CALCOLATO: risultante dall'algoritmo utilizzato per il checksum  
VALORE INDICATO: non soggetto a negoziazione

Tab. V.3 - Elenco dei parametri delle T-PDU CR e CC

Se l'entità chiamante è soddisfatta dell'esito della negoziazione, essa invia una T-CONNECT-conferma al TS-utente e la T-conneessione si considera allora stabilita con successo. Se l'entità chiamante non considera accettabili le restrizioni indicate nella CC, essa avvia la procedura di abbattimento della connessione.

Ad esempio, siano *A* e *B* due generici indirizzi di trasporto, *a* e *b* due corrispondenti indirizzi di rete. Intendendo con la notazione *sap-x ! p* dire "al *sap-x* si osserva il trasferimento della primitiva *p*", descriviamo la procedura di instaurazione di una T-conneessione come serie di eventi osservabili ai SAP di trasporto e di rete. Supponiamo che la richiesta di T-conneessione parta dal T-SAP-A e che non sia necessario procedere all'assegnazione di una connessione di rete. Abbiamo:

T-SAP-A ! T-CONNECT-richiesta

N-SAP-a ! N-DATA-richiesta (CR)

N-SAP-b ! N-DATA-indicazione (CR)

T-SAP-B ! T-CONNECT-indicazione

T-SAP-B ! T-CONNECT-risposta

N-SAP-b ! N-DATA-richiesta (CC)

N-SAP-a ! N-DATA-indicazione (CC)

T-SAP-A ! T-CONNECT-conferma

Segmentazione e ri-assemblaggio. Un'entità che riceve in consegna una T-SDU (unità di dati del servizio di trasporto) può suddividerla, ai fini dell'invio, tra i campi informativi di più T-PDU dati consecutive (segmentazione). In tal caso, l'entità ricevente ricomporrà i vari segmenti in un'unica sequenza da rilasciare al TS-utente (ri-assemblaggio). A tal fine, nelle T-PDU dati (vedi Fig. V.12) esiste un campo EOT (End Of Text, cioè "fine del

testo dell'utente") il cui valore discrimina se il campo dati di utente è o non è l'ultimo segmento di una T-SDU.

LI	1111 0000	TPDU-NR, EOT	campo dati
----	-----------	--------------	------------

formato della tpdu DT in classe 0 e classe 2

LI	1111 0000	DST-REF	TPDU-NR, EOT	parte variabile	campo dati
----	-----------	---------	--------------	-----------------	------------

formato della tpdu DT nelle classi 1, 3, 4

Fig. V.12 - Formato delle T-PDU DT e ED

*Rifiuto di connessione.* Questa procedura viene invocata dall'entità chiamata per rispondere negativamente ad una CR. Se la CR è formalmente valida ma contiene opzioni e valori dei parametri non convenienti all'entità chiamata, essa viene rigettata tramite una DR. Se la CR non è valida nel formato o nelle opzioni o nei valori dei parametri, essa viene rigettata tramite una ER.

*Associazione di T-PDU con T-connessioni.* Per associare univocamente una T-PDU ad una determinata T-connessione si utilizzano i riferimenti di origine e di destinazione (SRC-REF e DST-REF), per essa stabiliti. Il valore DST-REF trasferito con la CC si usa nel corso della connessione per etichettare le T-PDU inviate dall'entità chiamante; il valore SRC-REF trasferito nella CR invece si usa nel corso della connessione per etichettare le T-PDU inviate dall'entità chiamata. Il riferimento presente in una T-PDU diversa dalla CR è sempre chiamato DST-REF.

Premesso quanto sopra, ecco come viene elaborata una T-PDU ricevuta: se essa è una CR, l'entità ricevente attua la procedura di negoziazione relativa ad una nuova T-connessione. In caso contrario essa legge il campo DST-REF e, in base allo stato della connessione da questo identificata, tratta opportunamente la T-PDU ricevuta.

#### PROCEDURE SPECIFICHE DELLA CLASSE 0.

La classe 0 ha due procedure specifiche per il rilascio della connessione: il rilascio implicito e il rilascio su errore.

*Rilascio implicito.* Questo termine significa che non occorre trasferire

alcuna T-PDU di disconnessione per rilasciare una T-connessione. In classe 0, un'entità disconnette la T-connessione semplicemente inviando una N-DISCONNECT-richiesta. L'entità che riceve una N-DISCONNECT-indicazione considera automaticamente chiusa anche la T-connessione.

Rilascio su errore. L'errore di cui si parla è un "errore segnalato dalla rete" e può essere di due tipi: o si tratta di una N-DISCONNECT-indicazione per iniziativa del NS-fornitore, oppure si tratta di una N-RESET-indicazione. Poiché in tali eventualità può esserci stata perdita di dati e poiché la classe 0 non ha alcun meccanismo di recupero da anomalie, non resta che abortire la T-connessione mediante l'invio di una T-DISCONNECT-indicazione all'utente locale.

#### *PROCEDURE SPECIFICHE DELLA CLASSE 1.*

Quattro procedure caratterizzano la classe 1: la possibilità di ri-assegnare una T-connessione dopo caduta della N-connessione, la possibilità di risincronizzare il flusso dati, la congelazione dei riferimenti e la ritenzione delle T-PDU fino a riscontro avvenuto. La prima procedura permette di massimizzare la resilienza di una T-connessione. Le altre tre mirano a minimizzare il tasso di errore residuo di una T-connessione.

*Riassegnazione dopo caduta della N-connessione.* Questa procedura si avvia in seguito ad una N-DISCONNECT-indicazione per iniziativa della rete; consiste nel tentare ripetutamente di richiedere una nuova N-connessione, senza considerare abbattuta la T-connessione; se il ciclo di tentativi, controllato da un tempo di guardia, ha successo, si procede con la ri-sincronizzazione della T-connessione (vedi più avanti), altrimenti la si considera chiusa ma se ne congelano i riferimenti (vedi più avanti).

*Risincronizzazione.* Consiste nello scambio di opportune informazioni di controllo fra le entità, allo scopo di riprendere lo scambio dati da uno stato noto ad entrambe, in seguito a segnalazione di reset dalla rete oppure a ri-assegnazione dopo caduta della N-connessione.

*Congelazione dei riferimenti.* Consiste nel "marcare" i riferimenti di una T-connessione che non può essere più mantenuta; tale marcatura previene la ri-assegnazione di quei riferimenti a nuove T-conessioni. Si evita così il rischio di associare a queste ultime dati appartenenti alla vecchia T-connessione, che dovessero giungere dopo la sua chiusura.

La procedura di "congelazione dei riferimenti" si applica anche nella vita di tutti i giorni. Ad esempio, se un impiegato si dimette, l'azienda non ri-assegnerà immediatamente il suo numero di matricola ad un nuovo assunto, per evitare pericolose confusioni amministrative e contabili.

*Ritenzione delle T-PDU fino a riscontro avvenuto.* Al fine di limitare le necessità di ritrasmissione, le due entità possono concordare di utilizzare, per determinate T-PDU, la tecnica del riscontro di ricezione. Finché non giunge tale riscontro, le T-PDU sono memorizzate in un'area tampone per un'eventuale ritrasmissione. Sono possibili due varianti per effettuare i riscontri: utilizzo della primitiva del servizio di rete N-DATA-ACKNOWLEDGE (variante "conferma di ricezione") oppure utilizzo dell'apposita T-PDU AK (variante "normale"). La prima variante può convenire se all'uso del servizio di riscontro a livello di rete può corrispondere una riduzione del costo delle N-conessioni. La scelta della variante da utilizzare è effettuata a tempo di instaurazione della T-conessione.

Elenchiamo alcune T-PDU che possono essere ritenute finché non arrivi un riscontro. Una CR, in entrambe le varianti, è ritenuta finché non arriva una CC, una DR oppure una ER. Una CC è ritenuta solo nella variante "conferma di ricezione" e finché non arriva una N-DATA-ACKNOWLEDGE-indicazione oppure una delle T-PDU RJ, DT, ED, EA. Una DT è ritenuta anch'essa solo nella variante "conferma di ricezione" e finché non arriva una N-DATA-ACKNOWLEDGE-indicazione relativa alla N-DATA-richiesta su cui viaggiava la DT in questione.

#### *PROCEDURE SPECIFICHE DELLA CLASSE 2.*

Controllo di flusso. Questa procedura opera in aggiunta al controllo di flusso già assicurato dallo strato di rete. È quindi inutile quando c'è corrispondenza uno-a-uno tra connessioni di trasporto e connessioni di rete (classe 0 e classe 1). Essa diventa, invece, fortemente consigliabile in caso di moltiplicazione: secondo la normativa, è opzionale in classe 2, ed obbligatoria nelle classi 3 e 4.

Il controllo di flusso è realizzato con modalità non dissimili da quelle usate per lo strato di rete e non viene qui trattato ulteriormente.

Moltiplicazione-demoltiplicazione. Per ottimizzare i costi del servizio di rete o, semplicemente, per poter attuare T-conessioni in simultanea quando è disponibile un'unica connessione di rete, è possibile l'assegnazione di più T-conessioni ad una singola N-conessione (moltiplicazione). In questo modo,

due entità possono trasferire su una stessa N-connessione T-PDU appartenenti a T-conessioni distinte.

Si dice demultiplazione l'operazione effettuata in ricezione per discriminare T-PDU ricevute sulla stessa N-connessione e appartenenti a T-conessioni distinte, così da ricostruire il flusso dati su ognuna di queste.

#### *PROCEDURE SPECIFICHE COMUNI ALLE CLASSI 1 E 2.*

Concatenazione-separazione. Al fine di ottimizzare la espressità di trasferimento, può convenire trasferire su un'unica N-SDU una sequenza di T-PDU (concatenazione). Le T-PDU concatenate si decodificano in base al loro campo lunghezza. Esistono delle opportune regole di concatenabilità:

- CR, DR, CC, DT , ED non si possono concatenare;
- AK, EA, RJ, ER, DC sono concatenabili solo se appartenenti a connessioni distinte;
- il numero di T-PDU concatenabili non può superare il grado di multiplazione (numero di T-conessioni attive che usano la stessa N-connessione).

L'operazione inversa della concatenazione si dice separazione.

Rilascio esplicito. L'aggettivo "esplicito" oppone questa procedura a quella valida in classe 0. Il rilascio della connessione viene eseguito con lo scambio delle T-PDU DR e DC. Il rilascio di una T-conessione non implica il rilascio della N-connessione.

*Numerazione delle T-PDU dati.* È obbligatoria in classe 1; lo diventa in classe 2, in presenza dell'opzione per il controllo di flusso. Basilare nelle procedure di controllo di flusso, di recupero da errore e di riordino in sequenza, è realizzata con il campo T-PDU-NR della DT (vedi Fig. V.12).

*Dati espressi.* Questa procedura utilizza le T-PDU ED ed EA, convogliandole, dietro opzione, o sulla primitiva N-DATA o sulla primitiva N-EXPEDITED-DATA.

#### *PROCEDURE SPECIFICHE DELLE CLASSI 3 E 4.*

Dal punto di vista dell'insieme delle funzioni offerte, la classe 3 è l'unione delle classi 1 e 2. Perciò, non occorre aggiungere altro a quanto già detto nei paragrafi precedenti. La classe 4 è invece oggetto del resto di questo paragrafo.

Due procedure notevoli della classe 4 sono la *rivelazione di errori non segnalati dalla rete* e la *ritrasmissione di T-PDU*. Esse sono concettualmente analoghe a quelle dello strato di collegamento. Infatti, una rete di tipo C può

considerarsi "scadente" dal punto di vista dell'affidabilità del trasferimento e, dunque, vanifica parzialmente eventuali controlli su errori di trasferimento, effettuati a livello di trama. È logico, dunque, aspettarsi che questi vadano ripetuti nello strato di trasporto, il primo completamente "da estremo a estremo" nell'architettura OSI.

Anche un'altra procedura, detta di *riordino in sequenza*, è caratteristica della classe 4. È necessaria perché le procedure di ri-sincronizzazione e di ri-assegnazione dopo caduta di connessione potrebbero provocare come effetto "collaterale" la perdita di sequenza nei dati trasferiti.

Un'ulteriore procedura, quella di *controllo di inattività*, temporizza gli intervalli tra un evento e l'altro ad ogni estremo della T-connessione. Se la loro durata supera un tempo di guardia, si suppone che abbia avuto luogo una disconnessione a livello di rete non segnalata dal servizio di rete. La T-connessione viene ri-assegnata ad una nuova connessione di rete.

Infine esiste la procedura di *suddivisione-fusione*, che serve per ottenere, su richiesta dell'utente del servizio di trasporto, T-connessioni con portata superiore a quella consentita da una singola connessione di rete. A tale scopo, in corrispondenza di una T-connessione, vengono aperte più connessioni di rete in simultanea. Durante la fase di trasferimento dati, la sequenza di T-PDU relativa alla T-connessione viene scissa in sotto-sequenze disgiunte che vengono convogliate sulle varie connessioni di rete (*suddivisione*). Al lato destinazione, dai sotto-flussi di dati trasferiti sulle varie connessioni di rete viene ricostruito un unico flusso di dati relativo alla T-connessione (*fusione*).

#### *REALIZZAZIONI INDUSTRIALI DEI PROTOCOLLI DI TRASPORTO.*

Raramente un'implementazione industriale dello strato di trasporto include tutte le classi di protocollo. Normalmente viene offerta una classe senza moltiplicazione e una con moltiplicazione, ad esempio C0 e C2 oppure C1 e C3. La classe 4 è raramente implementata, a causa degli elevati costi di sviluppo, della complessità del corrispondente software e della relativa utilità delle procedure C4 (fortunatamente, le reti di tipo C non sono la norma). Inoltre, in vista di alcune applicazioni OSI particolarmente critiche rispetto agli errori di trasferimento (ad esempio FTAM, File Transfer Access and Manipulation), esistono nello strato di sessione meccanismi di sincronizzazione e ri-sincronizzazione, attivabili su opzione, in grado di recuperare gli errori rivelati dal servizio di trasporto.



## VI Le reti in area locale e metropolitana

### VI.1 Topologie

La topologia logica di una LAN, e cioè la configurazione geometrica del suo mezzo di comunicazione, dal punto di vista delle relazioni di traffico consentite, è scelta tra varie possibili soluzioni, che ben rispondono alle esigenze di una infrastruttura a basso costo. Le topologie più largamente utilizzate sono:

- il bus bidirezionale ;
- il bus unidirezionale ;
- il doppio bus unidirezionale;
- l'anello;
- la stella.

#### *BUS BIDIREZIONALE.*

In una LAN con topologia a bus bidirezionale (Fig. VI.1), il mezzo ha una configurazione lineare e le stazioni sono connesse ad esso in derivazione. Un esempio di prodotto commerciale con questo tipo di topologia è fornito dalla *rete Ethernet*.

L'informazione viaggia dalla stazione d'origine, in entrambe le direzioni, verso le estremità del mezzo, raggiungendo quindi tutte le stazioni connesse alla rete. La stazione di destinazione riconosce, in base all'indirizzo trasportato nell'etichetta, le unità informative ad esso indirizzate e ne legge il contenuto durante il loro transito attraverso l'interfaccia tra stazione e mezzo.

Da un punto di vista fisico, la topologia a bus bidirezionale ha la caratteristica fondamentale di avere le interfacce tra le stazioni ed il canale trasmissivo completamente *passive*. Questo elemento aumenta considerevolmente l'affidabilità della struttura poiché, da un lato, diminuisce la possibilità di guasti dell'interfaccia e, dall'altro, anche in presenza di guasti, il funzionamento della rete non è compromesso. È tuttavia evidente che l'assenza di unità di rigenerazione limita la distanza percorribile dal segnale a causa dell'attenuazione del canale trasmissivo. Reti di questo tipo pongono quindi stringenti vincoli sulla loro estensione massima.

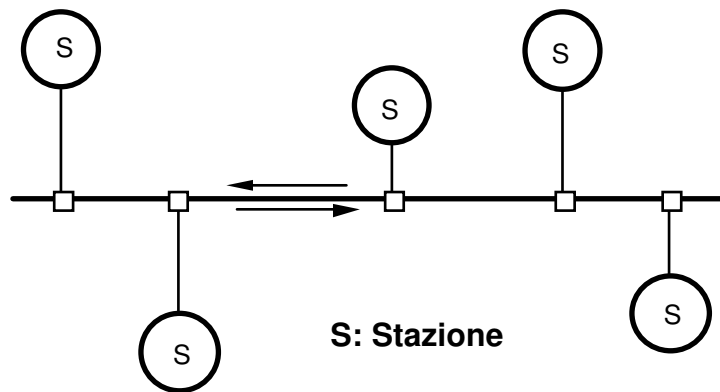


Fig. VI.1 - Topologia a bus bidirezionale.

Con riferimento alla rete Ethernet, in Fig. VI.2 è mostrata la struttura meccanica dell'interfaccia tra MAU (Medium Access Unit) e mezzo trasmissivo. Questa è composta da due blocchi che circondano la coppia coassiale e che provvedono a fornire il contatto elettrico con il conduttore. Una vite esterna assicura la stabilità del contatto.

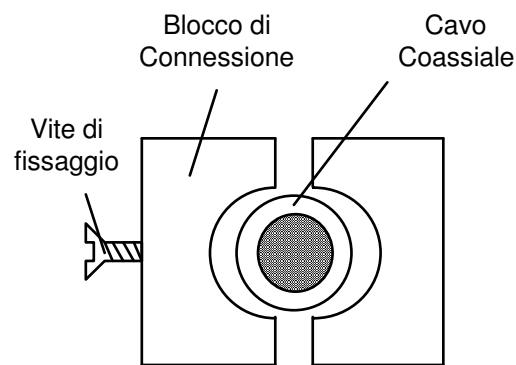


Fig. VI.2 - Struttura dell'interfaccia tra mezzo trasmissivo e MAU nella rete Ethernet.

Un segmento della rete Ethernet senza rigenerazioni ha un'estensione massima di mezzo chilometro e non può contenere più di 100 stazioni. Tali valori sono ottenuti tenendo conto dell'attenuazione tipica del mezzo trasmissivo, della perdita di inserzione causata da ogni interfaccia e del livello minimo di potenza di segnale necessario per il corretto funzionamento delle MAU. I vincoli di estensione massima possono essere superati suddividendo la rete in *sezioni* e dando luogo ad una topologia ad *albero* (Fig. VI.3). In questa le varie sezioni comunicano tra loro tramite dispositivi di interconnessione chiamati "*ponti*".

Un ponte (bridge) ha il compito di filtrare le unità informative in transito in una sezione di rete e, se queste sono dirette verso altre sezioni, di provvedere al loro rilancio. Le singole sezioni della rete hanno quindi un funzionamento indipendente. Il ponte esegue la funzione di instradamento delle unità informative tra sezioni diverse della rete.

La topologia ad albero è particolarmente adatta in casi di interconnessione di sezioni di rete in ambienti con elevata densità di traffico risulta particolarmente efficiente se il traffico di transito tra singole sezioni di rete è sufficientemente limitato in modo da non saturare la capacità dei ponti. Il criterio generale di progetto deve quindi prevedere di individuare aree omogenee di traffico e di dedicare a queste sezioni di rete indipendenti.

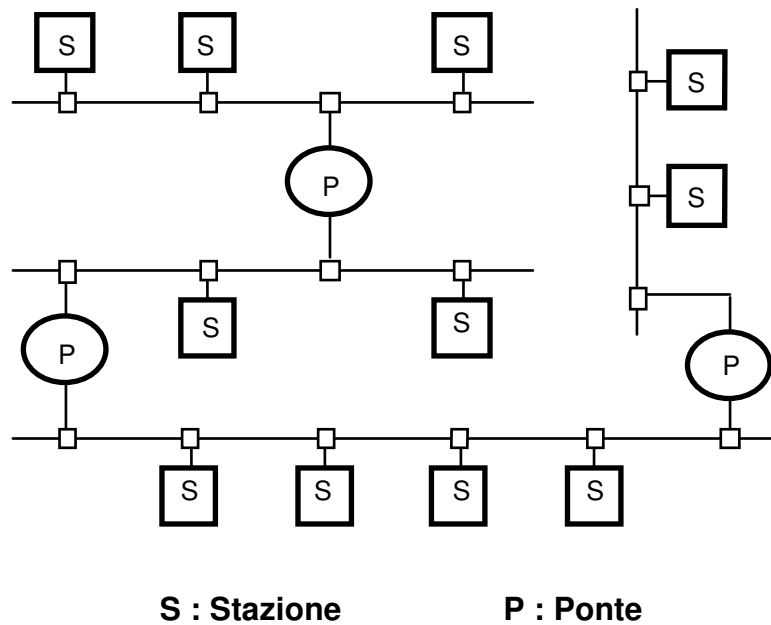


Fig. VI.3 - Topologia ad albero costituita da bus bidirezionali interconnessi.

**BUS UNIDIREZIONALE.**

La topologia a bus unidirezionale si differenzia dalla precedente per il fatto che l'informazione fluisce sul bus esclusivamente in un verso. Affinché sia assicurata la completa connettività tra le stazioni, occorre quindi suddividere il bus in due segmenti, uno *di immissione* e l'altro *di estrazione*. Ogni interfaccia tra mezzo e stazione deve riguardare ambedue i segmenti e risulta quindi funzionalmente composta da due unità distinte: una *di scrittura* e l'altra *di lettura*.

Le due alternative possibili per la creazione dei segmenti di immissione e di estrazione su un unico bus sono illustrate in Fig. VI.4. La prima alternativa (Fig. VI.4a) è un *bus unidirezionale ripiegato*, mentre la seconda (Fig. VI.4b) è un *bus unidirezionale doppiamente ripiegato*. La differenza sostanziale tra le due alternative è che i versi di attraversamento delle stazioni da parte dei segmenti di immissione e di estrazione sono diversi nel caso del bus ripiegato, mentre sono identici nel caso del bus doppiamente ripiegato.

È da osservare che, considerando il verso dei flussi informativi sul mezzo, in entrambe queste topologie una generica stazione è in grado di ricevere l'informazione emessa dalle stazioni "a monte" su entrambi i segmenti del bus, mentre le informazioni emesse dalle stazioni "a valle" sono ricevute solo sul segmento di estrazione. Tale caratteristica è sfruttata per la definizione di opportuni protocolli MAC, che per limiti di trattazione non verranno qui considerati.

Da un punto di vista fisico, l'unidirezionalità della propagazione dei segnali consente di superare i problemi di attenuazione del mezzo trasmissivo. È infatti possibile utilizzare dispositivi di amplificazione e/o di rigenerazione che, a causa della loro intrinseca unidirezionalità, non possono essere utilizzati nel caso della topologia a bus bidirezionale. In questo modo l'estensione massima di ogni sezione di rete può aumentare considerevolmente.

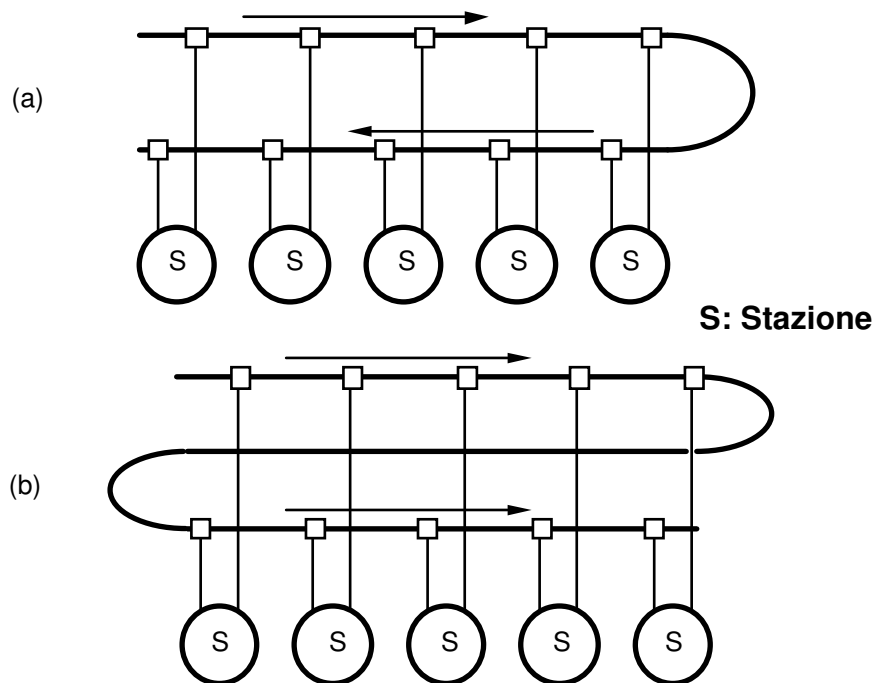


Fig. VI.4 - Topologie a bus unidirezionale: a) bus unidirezionale ripiegato; b) bus unidirezionale doppiamente ripiegato.

Un ulteriore elemento a favore della topologia a bus unidirezionale consiste nella sua affidabilità. Infatti, tramite un'opportuna *procedura di riconfigurazione*, le reti con questa topologia possono mantenere la piena connettività anche in presenza di un guasto in un segmento di rete.

In condizioni normali, la rete viene posta nella configurazione ad anello fisico mostrata in Fig. VI.5a, in cui un'unica stazione (*stazione di testa*) assolve le funzioni di *terminazione* e di *ripiegatura* del bus. Se avviene un guasto, la rete è in grado di isolare il segmento di bus guasto (Fig. VI.5b), sdoppiando le funzioni di terminazione e di ripiegatura del bus in due stazioni diverse. In tal modo la connettività tra le stazioni è conservata e il servizio di rete non deve essere interrotto. Quando il guasto sarà riparato, la rete potrà tornare nella sua configurazione fisica iniziale.

È evidente che condizione necessaria alla realizzazione della procedura di riconfigurazione ora descritta è che tutti le stazioni della rete siano in grado di svolgere, se richiesto, le funzioni di terminazione e di ripiegatura del bus. Ciò ovviamente aumenta la complessità delle stazioni ed il loro costo.

La procedura di riconfigurazione può essere così schematizzata. Il guasto di un segmento di rete viene rivelato dalla stazione a valle del segmento a causa della mancanza di segnale ricevuto. Quando ciò avviene, tale stazione assume il ruolo di stazione di testa ed emette un segnale di controllo per avvertire le altre stazioni dell'avvio della procedura di riconfigurazione. Alla ricezione di tale segnale, la stazione che precedentemente aveva il ruolo di stazione di testa assume la configurazione normale, mentre la stazione a monte del segmento guasto svolge la funzione di ripiegatura. Al termine della procedura la rete riprende il servizio normale.

#### DOPPIO BUS UNIDIREZIONALE.

La topologia a doppio bus unidirezionale (Fig. VI.6) è un'evoluzione di quella a bus

unidirezionale ripiegato. In questo caso i due segmenti corrispondenti ai due versi di trasferimento sono completamente separati. Un segmento attraversa così le stazioni in un verso, mentre l'altro le attraversa in verso opposto. Ovviamente, per utilizzare la rete al meglio delle sue capacità, una stazione d'origine dovrà effettuare l'emissione di una unità informativa solo su quel segmento che consente il trasferimento verso la stazione di destinazione. A tale scopo è necessario che la stazione d'origine conosca la posizione di quello di destinazione.

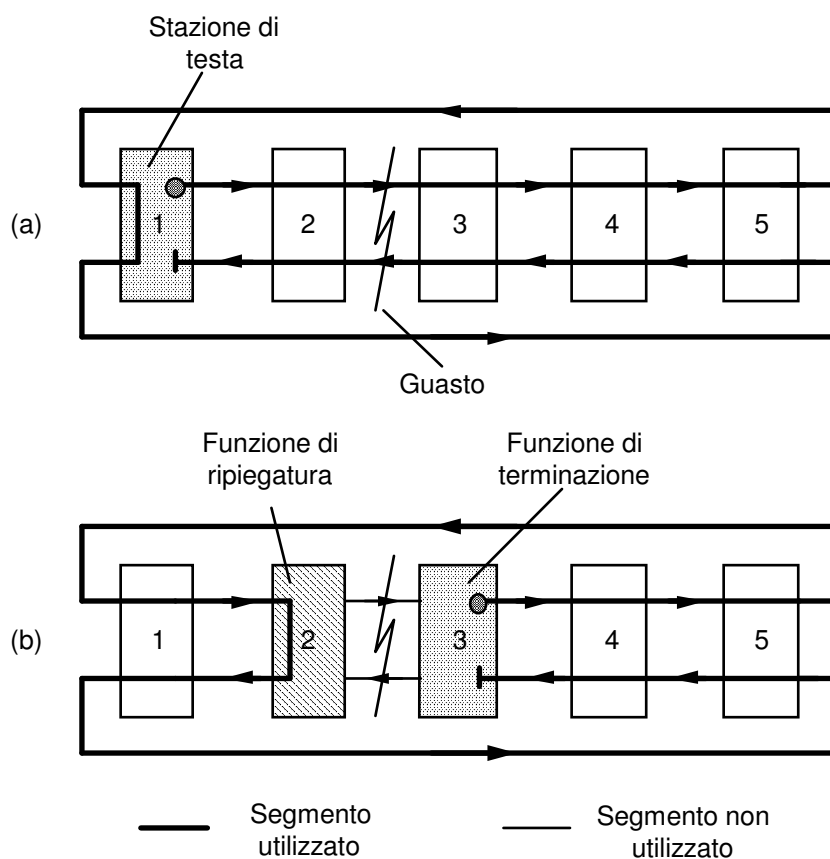


Fig. VI.5 - Riconfigurazione, in caso di guasto, di una rete a bus unidirezionale:  
a) funzionamento normale; b) rete riconfigurata.

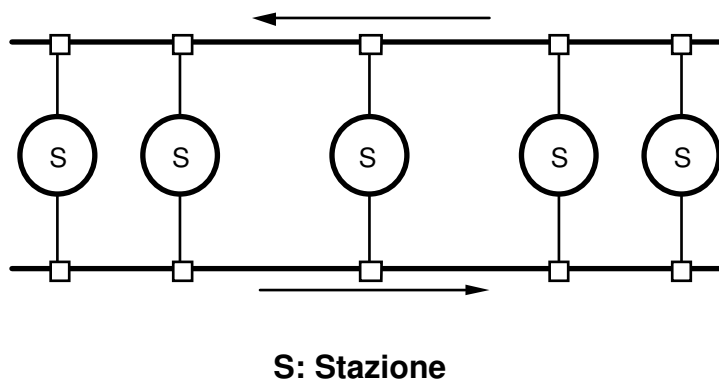


Fig. VI.6 - Topologia a doppio bus unidirezionale.

Nel caso in cui non si ritenga necessario ottenere la massima efficienza dalla rete, il modo più semplice di funzionamento è, naturalmente, quello di consentire l'emissione di una unità informativa su entrambi i bus. Per evitare l'uso di entrambi i bus, occorre invece che ogni stazione abbia a disposizione una tabella nella quale siano memorizzate le posizioni relative di tutti le altre stazioni connesse alla rete. La gestione di queste tabelle è un elemento molto importante per il corretto funzionamento della rete. In particolare, queste tabelle devono essere aggiornate quando una nuova stazione è connessa alla rete, ma anche quando una stazione viene spostata all'interno della rete stessa. A tale scopo deve essere prevista una procedura di controllo specializzata.

Le interfacce di una stazione su entrambi i segmenti sono del tutto identiche e devono consentire operazioni sia di scrittura che di lettura sul mezzo di comunicazione. Inoltre la gestione dell'accesso ai due segmenti è effettuata in modo completamente indipendente. L'intera rete è quindi costituita da due bus unidirezionali controllati in modo indipendente. Ciò evidentemente comporta una complessità doppia dell'interfaccia tra stazione e mezzo di comunicazione rispetto a quella che si avrebbe utilizzando una rete a segmento singolo.

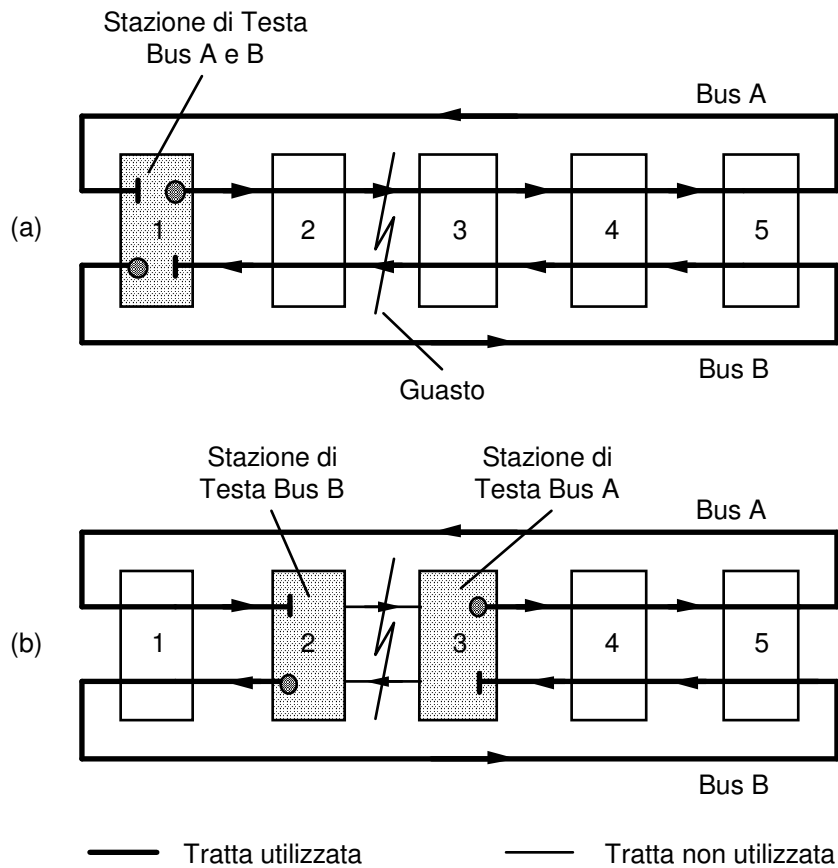


Fig. VI.7 - Riconfigurazione, in caso di guasto, di una rete a doppio bus unidirezionale:  
a) funzionamento normale; b) rete riconfigurata.

I vantaggi della topologia a doppio bus consistono essenzialmente nella sua affidabilità, nella possibilità di utilizzare pienamente le potenzialità delle fibre ottiche e nel fatto che l'effettiva capacità a disposizione delle stazioni è doppia rispetto a quella raggiungibile con reti a mezzo trasmissivo unico. Ciò è evidentemente dovuto al fatto che si ha un segmento dedicato per ogni verso di trasferimento. Per tali ragioni questa topologia ha trovato larga applicazione specialmente nelle reti locali ad alta velocità e nelle MAN.

Anche la topologia a doppio bus può essere riconfigurata in modo da garantirne la piena connettività anche in presenza di un guasto. La configurazione normale è detta a *bus chiuso* (looped bus) (Fig. VI.7a) ed è caratterizzata dal fatto che le funzioni di stazione di testa per entrambi i bus sono concentrate in un'unica stazione. In caso di guasto, la procedura di riconfigurazione deve fare in modo di assegnare a due stazioni diverse, quelli a valle del segmento guasto sui due bus, le funzioni di stazioni di testa (Fig. VI.7b).

#### ANELLO

Nella topologia ad anello, ogni stazione è connessa a due sole altre stazioni, una "a monte" e l'altra "a valle", tramite legamenti unidirezionali, che complessivamente formano un percorso chiuso (Fig. VI.8). Le informazioni sono trasferite in modo sequenziale da una stazione alla successiva.

Ogni stazione dell'anello può essere, oltre che origine e destinazione, anche punto di transito delle unità informative; in quest'ultimo caso, da un punto di vista trasmissivo, una stazione agisce come elemento attivo della rete operando una rigenerazione del segnale.

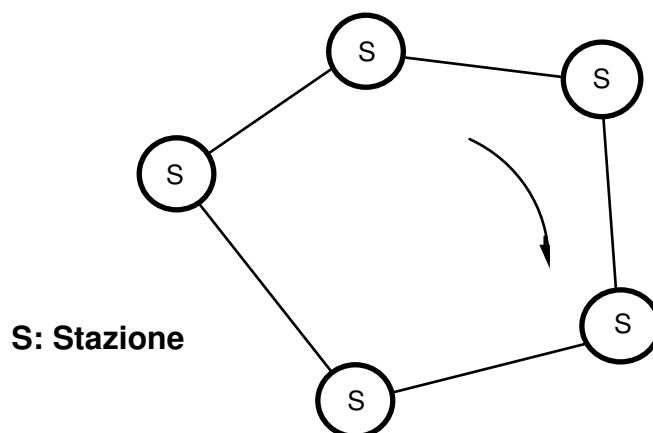


Fig. VI.8 - Topologia ad anello.

Rispetto ad una topologia a bus, unidirezionale o bidirezionale, si può osservare che: a) la presenza di interfacce attive elimina le limitazioni all'estensione massima della rete, ma diminuisce le prestazioni in termini di affidabilità; b) le unità informative entranti in ogni stazione debbono essere memorizzate, totalmente o parzialmente, in modo da riconoscere la loro destinazione e da decidere l'eventuale ri-emissione sul ramo uscente.

Inoltre, mentre nelle LAN con topologia a bus le informazioni che sono trasferite sul mezzo di comunicazione vengono assorbite da opportuni terminatori posti alle estremità del bus, nelle reti ad anello le unità informative circolerebbero indefinitamente sul mezzo a meno di provvedimenti opportuni. È quindi necessario che venga effettuata la rimozione delle unità informative aprendo, in istanti opportuni, l'anello.

L'operazione di rimozione di una unità informativa dall'anello richiede la lettura degli indirizzi della unità informativa stessa. Se l'operazione di rimozione è effettuata dalla stazione d'origine, questo deve controllare che l'indirizzo di sorgente corrisponda al suo. Se invece la rimozione è effettuata dalla stazione di destinazione, quest'ultima deve controllare l'indirizzo di destinazione. Queste operazioni richiedono l'introduzione, in ogni stazione, di un ritardo uguale almeno al numero di bit da controllare. Nella valutazione del ritardo totale di trasferimento di una unità informativa, tale ritardo (*latenza della stazione*) va aggiunto a quello di propagazione fisica dei segnali.

Si può osservare che l'adozione di interfacce di tipo attivo comporta un aggravio in termini di costo di installazione ed una diminuzione dell'affidabilità del sistema. Per ovviare a quest'ultimo inconveniente, spesso si adottano strutture ad *anello duplicato*, che consentono di riconfigurare la rete in modo da isolare l'eventuale segmento guasto.

In conclusione, la topologia ad anello può essere vantaggiosamente utilizzata laddove occorra connettere stazioni localizzate in un'area estesa che richiedono elevate capacità. Ciò si ottiene però al prezzo di una struttura più complessa per ciò che riguarda la gestione dei guasti e delle riconfigurazioni.

La struttura ad anello duplicato più largamente utilizzata è quella mostrata in Fig. VI.9a, che prevede due anelli con versi di trasferimento opposti (*counter-rotate ring*). In condizioni normali di funzionamento la presenza di due anelli raddoppia la capacità della rete. In presenza di un guasto, il segmento guasto viene isolato poiché le stazioni a destra e a sinistra di questo provvedono alla connessione tra i due anelli e quindi a formare un unico anello (Fig. VI.9b). In questo caso però, a differenza delle riconfigurazioni descritte nelle Fig. VI.5b e Fig. VI.7b, la capacità di trasferimento della rete diventa uguale a quella di un unico anello.

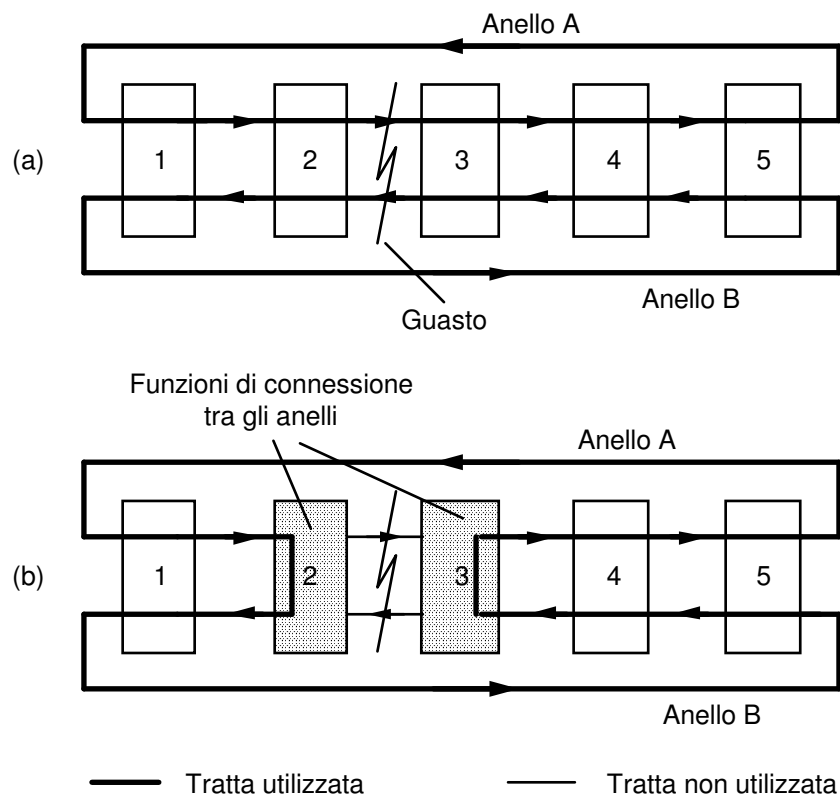


Fig. VI.9 - Riconfigurazione, in caso di guasto, di una rete ad anello: a) funzionamento normale; b) rete riconfigurata.

STELLA

I principi generali su cui si basa la tecnologia delle LAN (ad es. mezzo di comunicazione unico, distribuzione delle funzioni di controllo, ecc.) non sempre si adattano al caso di topologia a stella (Fig. VI.10). Questa risulta infatti semplice e vantaggiosa principalmente quando le stazioni colloquiano in modo prevalente con il centro stella, come avviene, ad esempio, nel caso di un elaboratore centrale connesso ad un certo numero di stazioni. Se il traffico di rete è invece composto prevalentemente da comunicazioni tra

stazioni, come è il caso tipico in una LAN, il centro stella dovrebbe avere una capacità molto elevata in modo da poter smaltire tutto il traffico generato dalle stazioni. Ciò rende talvolta antieconomica tale soluzione poiché costringerebbe ad un sovra-dimensionamento delle risorse e, quindi, ad un elevato costo iniziale di installazione.

La topologia a stella è largamente impiegata nelle reti locali che utilizzano *centralini privati* (PABX, Private Automatic Branch Exchange). I PABX operano secondo il modo di trasferimento a circuito. Normalmente, questi apparati permettono, oltre all'espletamento del servizio telefonico, anche l'effettuazione di trasferimenti di dati tra le varie terminazioni. Tuttavia, poiché il principio di funzionamento di tali sistemi è quello a circuito, la loro utilizzazione come supporto di comunicazione in un ambiente di informatica distribuita non è, in generale, conveniente. Infatti, occorre considerare che la complessità delle procedure di instaurazione di una connessione tra due terminazioni è notevolmente maggiore rispetto a quella prevista in una LAN. Inoltre il tipo di traffico generato da un terminale di dati è tale da condurre ad una sensibile sotto-utilizzazione dei canali di comunicazione.

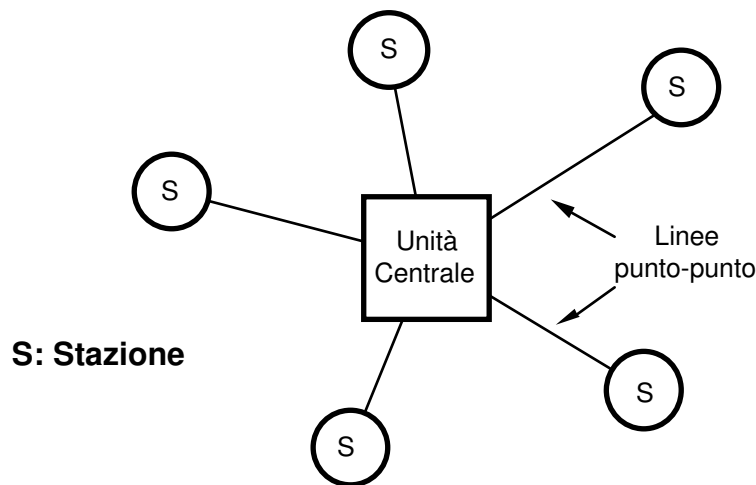


Fig. VI.10 - Topologia a stella.

## VI.2 Architetture di una LAN

La architettura di una stazione connessa ad una LAN è mostrata in Fig. VI.11 e posta a confronto con l'apparecchiatura di un apparecchio terminale che rispetti i principi del modello OSI. Da questo confronto appare che lo strato di collegamento è sdoppiato in due sotto-strati: il sotto-strato di *controllo del collegamento logico* (Logical Link Control - LLC) ed il sotto-strato di *controllo dell'accesso al mezzo* (Medium Access Control - MAC). Nel seguito, per brevità, questi due sottostrati saranno indicati come *strato LLC* e *strato MAC*.

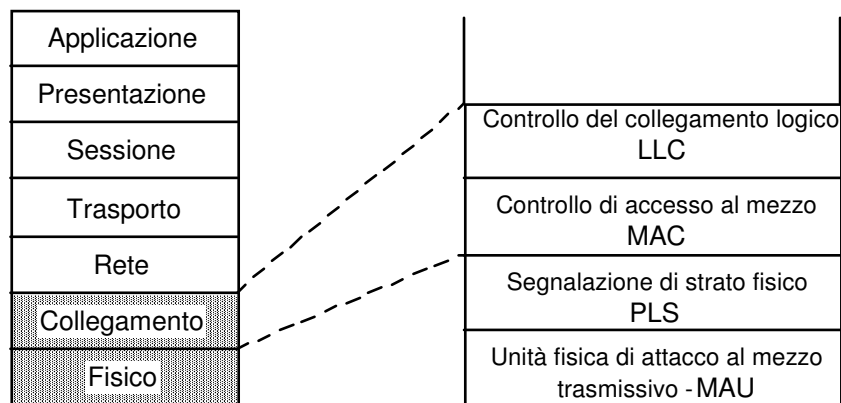


Fig. VI.11 - Architettura logica di un terminale LAN.

Lo strato LLC ha il compito di controllare lo scambio delle unità di dati tra i terminali connessi ad una LAN. In particolare, esso tratta le *unità di dati del protocollo di strato LLC* (LLC Protocol Data Unit - LLC PDU) allo scopo di rendere lo scambio informativo efficiente ed esente da errori.

Lo strato MAC ha invece la funzione di regolare l'accesso al mezzo, regolando eventuali contese tra le stazioni. Il particolare protocollo è il già citato protocollo d'accesso (*protocollo MAC*) ed è uno degli elementi caratteristici di una LAN.

Lo strato fisico è a sua volta suddiviso in un due sotto-strati, uno logico e l'altro fisico. Il primo è il *sotto-strato di segnalazione di strato fisico* (Physical Layer Signaling - PLS), mentre il secondo è l'*unità di attacco al mezzo trasmissivo* (Medium Attachment Unit - MAU).

Il sotto-strato PLS ha il compito di mettere in grado le entità MAC di inviare le loro unità di dati sul canale trasmissivo. La MAU ha invece il compito di consentire l'accoppiamento tra il sotto strato PLS ed il mezzo trasmissivo. Tale unità emette e riceve i segnali su e dal mezzo trasmissivo sotto il comando del sottostrato PLS.

Sebbene in gran parte le LAN siano state originariamente definite e realizzate da aziende private e quindi basate su architetture e protocolli di tipo proprietario, da parte di varie organizzazioni internazionali è stato prodotto un notevole sforzo per arrivare all'approvazione di norme che stabilissero una definizione univoca dei protocolli dei vari strati funzionali di una LAN.

Attualmente, il riferimento in questo campo è costituito dalla famiglia di norme emesse dalla Institution of Electrical and Electronics Engineers (IEEE), conosciute sotto la sigla IEEE 802; queste norme sono poi state fatte proprie dall'ISO nella serie 8802.

Le norme della famiglia IEEE 802 trattano esclusivamente gli strati fisico,

MAC e LLC. In Fig. VI.12 è illustrata la relazione fra queste norme, i cui contenuti sono qui di seguito riassunti brevemente:

- *IEEE 802.1 (ISO 8802.1)*: introduce la struttura e le architetture dei singoli elementi della famiglia IEEE 802;
- *IEEE 802.2 (ISO 8802.2)*: descrive le funzioni associate allo strato LLC, specificando la sua interfaccia (SAP, Service Access Point) con lo strato superiore e con quello MAC;

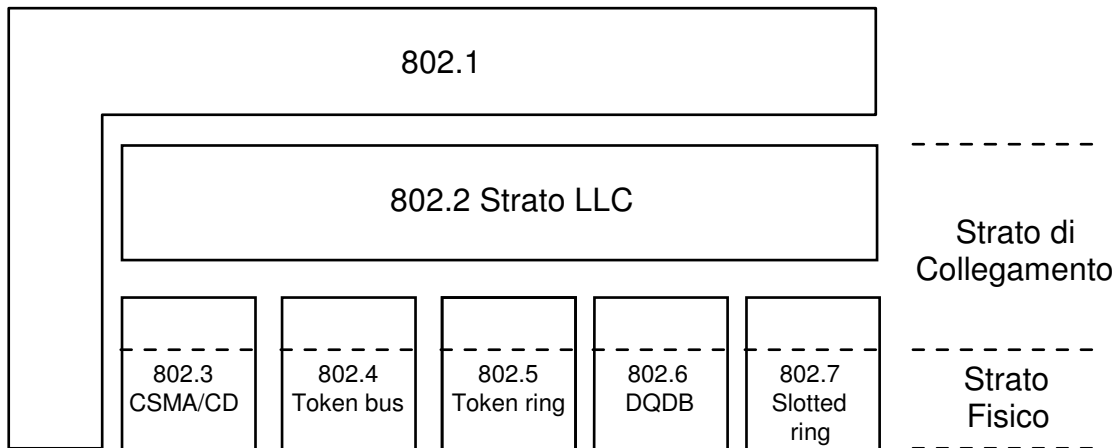


Fig. VI.12 - Norme IEEE 802 e loro relazione con il modello OSI.

- *IEEE 802.3 (ISO 8802.3)* : descrive il protocollo di accesso CSMA/CD, con riferimento a diversi mezzi trasmissivi ed ai ritmi binari utilizzabili (1-20 Mbit/s), ed il relativo strato fisico;
- *IEEE 802.4 (ISO 8802.4)* specifica gli elementi riguardanti il protocollo di accesso a testimone su bus (token bus) ed il relativo strato fisico;
- *IEEE 802.5 (ISO 8802.5)* : specifica gli elementi riguardanti il protocollo di accesso a testimone su anello (token ring) ed il relativo strato fisico;
- *IEEE 802.6 (ISO 8802.6)* : descrive la MAN DQDB, che utilizza il protocollo d'accesso a coda distribuita;
- *IEEE 802.7 (ISO 8802.7)* : riguarda una LAN con topologia ad anello, in cui il protocollo MAC è chiamato “ad anello suddiviso in intervalli temporali” (slotted ring).

### VI.3 Tecniche di trasmissione nelle LAN

Le due metodologie utilizzate per la trasmissione delle informazioni sul mezzo trasmissivo costituente una LAN sono: 1) *la trasmissione numerica in banda base*; 2) *la trasmissione analogica in banda estesa*.

Nel caso di trasmissione in banda base, il segnale numerico in linea occupa

l'intera banda del canale trasmissivo. Quindi solo un terminale alla volta può emettere le proprie unità informative. Si devono quindi prevedere, come già accennato, opportuni schemi di regolazione dell'accesso al mezzo per evitare la sovrapposizione dei segnali.

I vantaggi delle LAN in banda base sono la loro semplicità realizzativa ed il loro basso costo. L'installazione è rapida e semplice ed i sistemi impiegati hanno ormai una tecnologia acquisita. Occorre però tenere conto delle limitazioni riguardanti la loro capacità trasmissiva e le loro possibilità di copertura nel caso in cui la rete debba gestire un traffico elevato oppure coprire superfici estese.

Eventuali limitazioni in estensione possono essere superate utilizzando dispositivi come i *ripetitori*, che provvedono a rigenerare il segnale in transito sul mezzo trasmissivo, oppure come i ponti (cfr § I.1.4), che hanno lo scopo di interconnettere segmenti diversi della stessa rete. Nel primo caso, le varie tratte di ripetizione formano un'unica rete; nel secondo, i vari segmenti hanno un funzionamento indipendente e spetta ai ponti di gestire il traffico che deve passare da un segmento all'altro. È evidente che i ponti, se non sono accuratamente progettati ed utilizzati, possano divenire colli di bottiglia per il traffico in transito sulla LAN, riducendo sensibilmente la portata di rete.

Nella trasmissione in banda estesa viene utilizzata la tecnica di multiplazione a divisione di frequenza, in modo da consentire la trasmissione simultanea da parte di più terminali. La larghezza di banda del mezzo trasmissivo viene suddivisa in un certo numero di sottobande, ognuna delle quali è destinata ad uno specifico servizio di comunicazione (ad es. alla telemetria, alla voce, ai dati, al video, ecc.). In questo modo, ogni sottobanda costituisce un canale indipendente, a cui può accedere, con le modalità precedentemente descritte, una pluralità di utenti.

Se una LAN con trasmissione a banda estesa deve servire una vasta area geografica, tale da richiedere l'uso di amplificatori di linea unidirezionali, la propagazione dei segnali sul mezzo trasmissivo deve avvenire necessariamente in un solo verso, a meno di accorgimenti di tipo particolare. Pertanto, per garantire la piena connettività a tutte i terminali della rete, occorre prevedere un *canale di trasmissione* ed un *canale di ricezione* separati, con versi di propagazione opposti.

L'elemento di giunzione fra i due canali è un dispositivo detto *traspositore*, la cui funzione è quella di raccogliere i segnali in ingresso provenienti dal canale di trasmissione e ritrasmetterli in uscita sul canale di ricezione.

I canali di trasmissione e di ricezione possono essere separati assegnando a

ciascuno di essi un diverso mezzo trasmissivo oppure una diversa sotto-banda dello stesso mezzo trasmissivo. Nel primo caso, il traspositore è un semplice dispositivo di interconnessione e, eventualmente, di amplificazione, fra i due mezzi. Nel secondo caso, i segnali che viaggiano nelle due direzioni sono separati in frequenza da una opportuna banda di guardia. Il traspositore ha quindi il compito di traslare i segnali dalla sotto-banda di trasmissione a quella di ricezione.

Nelle LAN a banda estesa, per ottenere una struttura di interconnessione fisica ad albero, vengono utilizzati anche dispositivi di *accoppiamento* e di *diramazione* dei segnali. Tale tipo di struttura trova largo impiego nei sistemi in cui, ad un certo numero di utenti periferici localizzati in corrispondenza delle foglie dell'albero, viene offerta una pluralità di servizi da un centro localizzato nella radice dell'albero.

Come già accennato, un tipico esempio di LAN a banda estesa con topologia ad albero è rappresentato dalle reti CATV. In questo caso, la banda del mezzo trasmissivo è suddivisa in canali, ognuno dei quali trasporta un distinto segnale televisivo. Accanto a questi possono essere previsti altri canali dedicati alle procedure di controllo. I canali televisivi sono di tipo unidirezionale, mentre i canali di controllo possono essere anche bidirezionali per consentire alle apparecchiature d'utente il colloquio con il centro servizi localizzato nella radice dell'albero.

#### **VI.4 Strato MAC**

I protocolli MAC possono essere raggruppati in due categorie: 1) *protocolli ad accesso casuale*; 2) *protocolli ad accesso controllato*.

Nel primo caso, un terminale può iniziare il trasferimento di un'unità di dati in qualsiasi istante, previa l'eventuale esecuzione di opportune procedure di ascolto del canale. Invece, nel secondo caso, il terminale è vincolato all'emissione solo in opportuni istanti, determinati dall'esecuzione della procedura di accesso.

Per le sue modalità di funzionamento, un protocollo ad accesso casuale non evita che accadano conflitti di accesso al mezzo (*collisioni*). Per tale motivo occorre prevedere opportuni meccanismi, che garantiscano il corretto trasferimento delle informazioni anche se tali conflitti accadono. Un esempio di protocollo ad accesso casuale è il CSMA/CD (cfr. § 0), utilizzato nella rete Ethernet.

Una collisione avviene quando due o più terminali emettono

contemporaneamente una unità informativa sul canale trasmissivo. In questo caso, per effetto della mutua interferenza tra i segnali, il contenuto informativo delle unità informative viene corrotto. Un protocollo ad accesso casuale deve quindi essere in grado di attivare opportune procedure di ritrasmissione in modo da garantire l'affidabilità del trasferimento informativo.

Invece, nel caso di protocolli ad accesso controllato, uno specifico segnale autorizza un terminale ad inviare i propri dati, evitando così l'accesso contemporaneo sul canale trasmissivo e quindi le collisioni. Per tale ragione questa classe di protocolli è anche detta *senza collisioni*.

La gestione del segnale di controllo può essere demandata ad una stazione speciale (*controllo centralizzato*), oppure può essere condivisa fra tutti i terminali del sistema (*controllo distribuito*). Un esempio di protocollo ad accesso controllato con controllo centralizzato è il *protocollo ad interrogazione* (polling), mentre esempi di analoghi protocolli a controllo distribuito sono il *protocollo a testimone* e il *protocollo a coda distribuita*. Nel seguito ci occuperemo esclusivamente dei protocolli a controllo distribuito.

## **VI.5 Protocolli ad accesso casuale**

I protocolli d'accesso di tipo casuale sono a controllo completamente distribuito e caratterizzati dal fatto che l'emissione delle unità di dati da parte delle stazioni connesse alla rete avviene in modo non coordinato: una stazione può emettere, infatti, le proprie unità di dati in qualsiasi istante.

Tale modo di operare non evita il presentarsi di eventi di *collisione*, che accadono quando due o più unità di dati sono emesse simultaneamente da due o più stazioni. Una collisione determina la completa inintelligibilità delle unità di dati in essa coinvolte e quindi la necessità, da parte delle stazioni emittenti, di provvedere alla loro ri-emissione. Quando il traffico offerto alla rete cresce, anche la probabilità di collisione aumenta; le collisioni sono quindi una potenziale causa di instabilità nel funzionamento della rete. Questo fenomeno, come sarà illustrato nel seguito, determina un valore massimo, non superabile, della portata media di rete.

A causa della semplicità delle loro procedure, i protocolli ad accesso casuale sono facilmente realizzabili e relativamente poco costosi ed hanno avuto una larga applicazione in LAN di tipo commerciale. In particolare, essi si adattano ad ambienti caratterizzati da un elevato numero di terminali, ognuno dei quali offre un traffico di tipo interattivo, altamente intermittente.

Esistono numerosi schemi di protocolli ad accesso casuale, che si

differenziano principalmente per le strategie seguite per ridurre il numero di collisioni, o, comunque, per limitarne l'effetto sulle prestazioni della rete. Nel seguito di questo paragrafo saranno presentati, e valutate le prestazioni, dei protocolli ALOHA e slotted ALOHA. Altri protocolli ad accesso casuale saranno trattati nei paragrafi seguenti.

#### VI.5.1 Protocollo ALOHA

Il protocollo ALOHA è stato sviluppato, agli inizi degli anni '70, per una rete multiaccesso a radio frequenza dell'Università delle Hawaii, in cui un insieme di stazioni periferiche erano logicamente connesse, da un unico canale, ad una stazione centrale. Tuttavia, il principio chiave di funzionamento del protocollo si adatta a qualsiasi rete in cui una pluralità di stazioni d'utente competono per l'uso di una risorsa comune, nel nostro caso il canale trasmissivo.

Il principio base del protocollo è estremamente semplice: una stazione può trasmettere non appena una propria unità dati è disponibile. Se la stazione emittente non riceve un riscontro positivo dalla stazione di destinazione entro un determinato intervallo di tempo (*time-out di ritrasmissione*), si assume che si sia verificata una collisione. Le stazioni coinvolte nella collisione devono provvedere a riemettere le proprie unità di dati. Occorre tuttavia fare in modo che gli istanti di ri-emissione non coincidano, in modo da evitare collisioni successive e, quindi, uno stato di stallo del protocollo. Tale obiettivo è raggiunto per mezzo di un opportuno *algoritmo di subentro* (back-off) che ha lo scopo di determinare, in modo casuale, l'istante di ri-emissione da parte di ogni stazione, così da ridurre la probabilità di collisione.

Per la valutazione delle prestazioni del protocollo, supponiamo che: 1) la rete sia composta da  $N$  stazioni; 2) ogni stazione generi le proprie unità di dati (UI) in accordo alla distribuzione di Poisson con frequenza media uguale a  $\lambda$  UI/sec; 3) la lunghezza delle UI sia costante ed uguale a  $L$  bit; 4) il ritmo binario di trasmissione sul canale sia uguale a  $F_c$  bit/s.

Il processo complessivo di arrivo delle UI alla rete è dato dalla sovrapposizione di due processi: uno di tipo poissoniano costituito dall'insieme dei processi relativi alle singole stazioni; il secondo costituito dalle UI che subiscono ritrasmissioni. Per l'analisi che segue, si assume che anche il processo di arrivo delle UI che subiscono ritrasmissione sia poissoniano. Ovviamente, ciò non corrisponde al vero, poiché l'arrivo di tali UI dipende dal verificarsi delle collisioni in rete, ma l'assunzione è tanto più valida quanto più elevato è il tempo intercorrente tra due trasmissioni successive della stessa UI; ovvero tanto

più è grande l'intervallo di ritrasmissione generato dall'algoritmo di subentro.

Indicando con  $\Lambda$  la frequenza media complessiva di interarrivo di UI alla rete, risulta ovviamente  $\Lambda \geq N\lambda$ . Se il traffico è basso ( $\lambda \approx 0$ ), le collisioni sono estremamente rare, quindi si ha  $\Lambda \approx N\lambda$ . Se il traffico è elevato ( $\lambda \gg 0$ ), la probabilità di collisione è conseguentemente elevata e quindi  $\Lambda > N\lambda$ .

Sotto le precedenti ipotesi, il tempo  $T$  necessario per la trasmissione di una UI è dato da  $L/F_c$  ed il numero medio  $G$  di tentativi di trasmissione di UI in un intervallo di tempo pari alla durata di una UI, è uguale a:

$$\text{(Eq. VI.1): } G = T \Lambda$$

Per calcolare la portata media  $S$  della rete, occorre tener conto solo delle trasmissioni andate a buon fine, quindi

$$\text{(Eq. VI.2): } S = G P_0$$

dove  $P_0$  è la probabilità che una UI non subisca collisioni.

Per il calcolo della probabilità  $P_0$  occorre considerare che, indicato con  $t_0$  l'istante di inizio della trasmissione di una UI sul bus, tale UI non subirà collisioni se nell'intervallo  $[t_0 - T; t_0 + T]$  nessun'altra stazione inizia una trasmissione. Tale intervallo, di ampiezza  $2T$ , è detto *intervallo di vulnerabilità* della trasmissione. Quindi, data l'ipotesi di traffico poissoniano, si ha:

$$\text{(Eq. VI.3): } P_0 = \Pr\{\text{nessuna trasmissione in } 2T\} = e^{-2\Lambda T} = e^{-2G}$$

e conseguentemente, dalla (Eq. VI.2):

$$\text{(Eq. VI.4): } S = G e^{-2G}$$

In Fig. VI.13 è mostrato l'andamento determinato dall'(Eq. VI.4, della portata media  $S$  del protocollo ALOHA in funzione di  $G$ . Come è facile verificare, il massimo valore della portata media si ottiene per  $G=0.5$  ed è uguale a  $S = 1/2e \approx 0.184$ ; ciò stabilisce che l'utilizzazione del canale non può andare oltre il 18% della sua capacità. Per  $G > 0.5$ , la portata media di rete diminuisce, dando luogo ad una instabilità della rete. Ciò può essere spiegato dal fatto che le collisioni via via aumentano, determinando una diminuzione progressiva della quota parte di utilizzazione utile del canale.

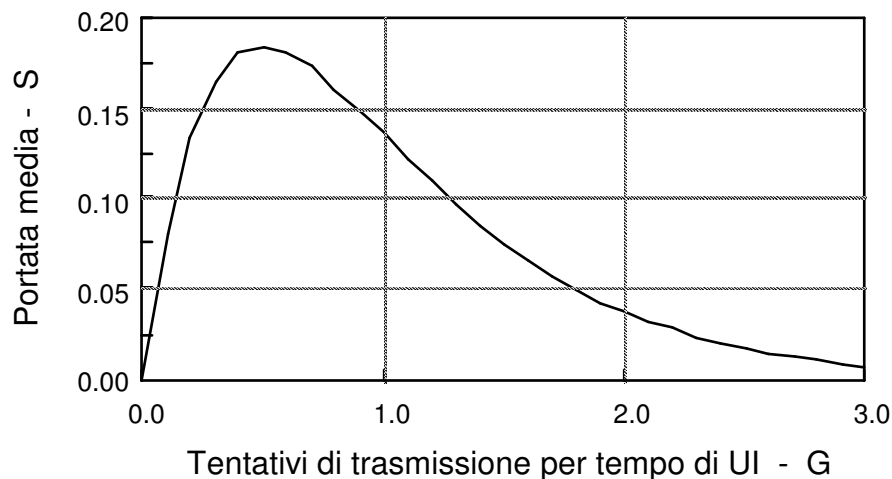


Fig. VI.13 - Portata media del protocollo ALOHA

Per quanto riguarda la valutazione del ritardo medio da accesso  $D$  subito da una UI, definito come il tempo intercorrente tra il primo arrivo della UI alla stazione d'origine e l'istante d'inizio sua trasmissione esente da collisione, occorre tener conto:

- del tempo necessario a rivelare un'eventuale collisione, ovvero la durata del time-out di ritrasmissione;
- dell'intervallo di ritrasmissione determinato dall'esecuzione dell'algoritmo di subentro.

Indicando con  $R_1$  la durata costante del time-out di ritrasmissione, che è supposto iniziare in corrispondenza dell'inizio della trasmissione di una UI, e con  $R_2$  la durata media dell'intervallo di ritrasmissione, si ha

$$(Eq. VI.5): \quad D = T + K (R_1 + R_2)$$

dove  $K$  è il numero medio di trasmissioni con collisione subite da una UI.

Il numero  $K$  può essere calcolato tenendo conto che il rapporto  $G/S$  equivale al numero medio totale di trasmissioni subite da ogni UI, ovvero:

$$(Eq. VI.6): \quad 1 + K = G/S$$

Sostituendo la (Eq. VI.4) nella (Eq. VI.6), si ha:

$$(Eq. VI.7): \quad K = e^{2G} - 1$$

quindi

$$(Eq. VI.8): \quad D = T + (e^{2G} - 1) (R_1 + R_2)$$

In Fig. VI.14 è mostrato il valore medio normalizzato (rispetto a  $T$ ) del ritardo di accesso del protocollo ALOHA, calcolato nel caso di  $T=250$  msec,

$R_1=100$  msec e  $R_2= 750$  msec e in funzione del numero medio  $G$  di tentativi di trasmissione per tempo di UI. Come si può vedere il valore di questo ritardo è trascurabile nell'intervallo di stabilità della rete ( $G<0.5$ ); oltre tale valore si ha una sua rapida crescita, dovuta all'elevato numero di collisioni.

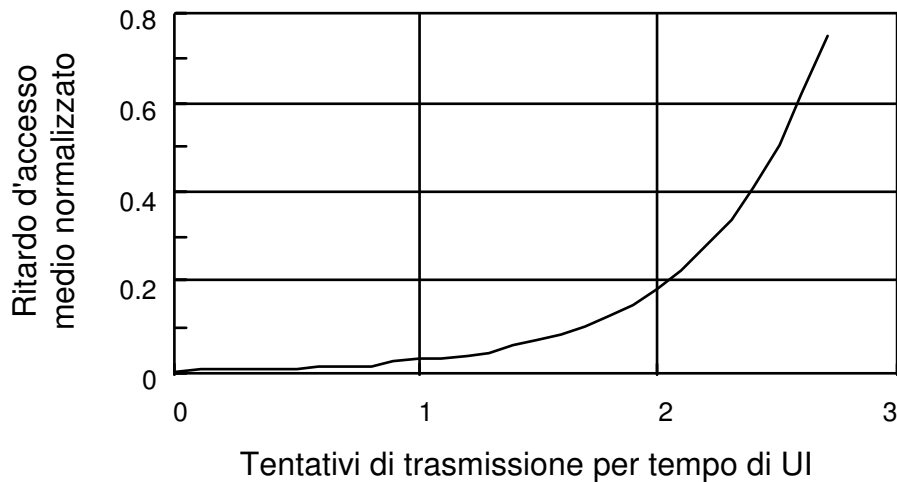


Fig. VI.14 - Ritardo medio di accesso  $D$  del protocollo ALOHA.

#### VI.5.2 Protocollo Slotted ALOHA.

Le scarse prestazioni di efficienza del protocollo ALOHA possono essere migliorate adottando una semplice strategia di sincronizzazione tra le stazioni. L'asse dei tempi è suddiviso in intervalli temporali (*slot*) di durata fissa, uguale al tempo di trasmissione  $T$  di una UI. Ogni stazione è vincolata ad iniziare la trasmissione delle proprie UI esclusivamente in corrispondenza dell'inizio di uno slot. Ovviamente, questo schema di accesso richiede che tutte le stazioni della rete siano sincronizzate tra loro, ovvero siano asservite ad un unico riferimento temporale generato da una stazione master.

Nel protocollo Slotted ALOHA, una UI è trasmessa senza collisioni nel caso in cui nessun'altra UI è trasmessa, nello stesso slot, da un'altra stazione. Tale eventualità si verifica solo se non si hanno nascite di altre UI nello slot precedente a quello di trasmissione. In altre parole, l'intervallo di vulnerabilità della trasmissione si riduce, in questo caso, alla durata  $T$  di uno slot, che è uguale alla metà del valore calcolato nel caso del protocollo ALOHA.

Ripetendo le medesime considerazioni effettuate nel paragrafo precedente, la portata media  $S$  ottenibile mediante il protocollo Slotted ALOHA è dato da:

$$(Eq. VI.9): \quad S = G e^{-G}$$

In Fig. VI.15 è mostrato l'andamento della portata media del protocollo

Slotted ALOHA, in funzione di  $G$ , ed è confrontato con la corrispondente portata media ottenibile nel caso del semplice protocollo ALOHA.

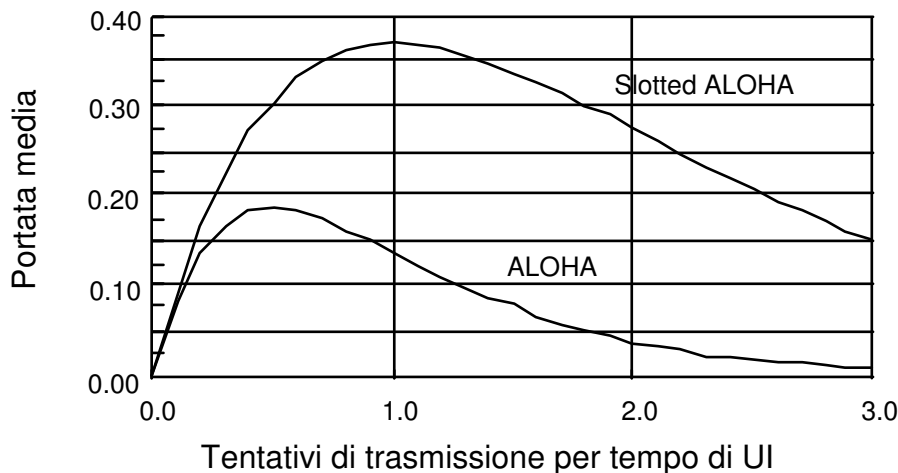


Fig. VI.15 - Portata media del protocollo Slotted ALOHA.

Come risulta evidente, la diminuzione dell'intervallo di vulnerabilità della trasmissione comporta un sensibile miglioramento della portata media di rete. In questo caso,  $S$  raggiunge il suo massimo per  $G=1$  ed il valore corrispondente è uguale al doppio di quello relativo al caso ALOHA, cioè  $S = 1/e \approx 0.368$ .

La valutazione del ritardo medio di accesso  $D$  può essere effettuata in modo del tutto analogo a quello riportato nel caso del protocollo ALOHA. L'unica differenza risiede nel fatto che, al tempo di trasmissione di una UI, occorre aggiungere il tempo medio di attesa tra la nascita di una UI e l'inizio dello slot successivo in cui viene effettuata la trasmissione. Tale tempo, nell'ipotesi di distribuzione uniforme degli istanti di nascita delle UI in un intervallo di tempo di durata finita, è uguale a  $T/2$ . Quindi, in analogia alla (Eq. VI.5):

$$(Eq. VI.10): \quad D = T + T/2 + K (R_1 + T/2 + R_2)$$

Il numero medio  $K$ , di trasmissioni con collisione subite da una UI, può essere ricavato sostituendo la (Eq. VI.9) nella (Eq. VI.6), da cui:

$$(Eq. VI.11): \quad K = e^G - 1$$

Quindi

$$(Eq. VI.12): \quad D = T + T/2 + (e^G - 1) (R_1 + T/2 + R_2)$$

In Fig. VI.16 sono confrontati i ritardi medi di accesso (normalizzati rispetto a  $T$ ) dei due protocolli ALOHA e Slotted ALOHA, calcolati per gli

stessi valori dei parametri utilizzati in Fig. VI.14 e in funzione del numero medio dei tentativi di trasmissione per tempo di UI.

Come risulta evidente, il protocollo ALOHA garantisce un ritardo inferiore a per valori molto bassi di traffico ( $G < 0.1$ ), il protocollo Slotted ALOHA è invece nettamente migliore per valori di traffico più elevati. Ciò può essere giustificato tenendo conto che, sebbene il protocollo Slotted ALOHA garantisce una minore probabilità di collisione, quando l'effetto di queste è trascurabile, cioè a traffico molto basso, il protocollo ALOHA è favorito dal fatto che prevede un accesso istantaneo alla rete.

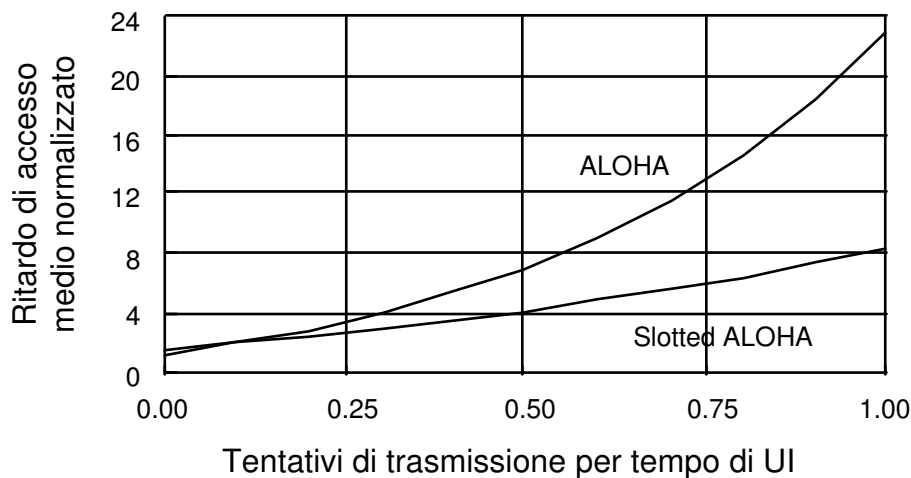


Fig. VI.16 - Ritardo medio di accesso del protocollo Slotted ALOHA.

## VI.6 Protocollo CSMA/CD

Il protocollo CSMA/CD è, come già detto, descritto nello standard IEEE 802.3 ed è utilizzato in LAN con topologia a bus bidirezionale. Esso è basato sulle esperienze derivate dall'utilizzazione dei protocolli ALOHA e sulla precedente definizione del protocollo CSMA (Carrier Sense Multiple Access).

Quest'ultimo, a differenza dei protocolli ALOHA, adotta la strategia "ascolta prima di parlare". Secondo questo schema, un terminale che deve emettere deve preventivamente ascoltare il canale. Se il canale è libero, la emissione può iniziare immediatamente; altrimenti sarà ritardata ad un successivo istante di tempo, la cui occorrenza viene stabilita attraverso l'esecuzione di una procedura di *persistenza*. A tale proposito esistono tre distinte possibilità:

- *1-persistente*;
- *0-persistente*;

- *p*-persistente;

il cui comportamento è mostrato in Fig. VI.17.

La procedura 1-persistente (Fig. VI.17a) prevede che, se al momento dell'ascolto, il canale è occupato, il terminale continui l'ascolto del canale ed esegua la sua emissione non appena il canale diventi libero.

La procedura 0-persistente (fig. Fig. VI.17b) differisce completamente dalla precedente; infatti, se il canale è occupato, il terminale ritarda la emissione di un intervallo determinato attraverso l'esecuzione di un *algoritmo di subentro* (backoff). Questo ha lo scopo di casualizzare il nuovo accesso al canale e quindi di ridurre la probabilità di collisione.

Un comportamento intermedio è fornito dalla procedura *p*-persistente (Fig. VI.17c). In quest'ultimo caso, se il canale è occupato, il terminale attende che ritorni nello stato di libero. A questo punto la emissione avviene con probabilità *p*, mentre con probabilità *1-p* il terminale attende ulteriormente un intervallo la cui durata si calcola con l'algoritmo di subentro.

L'importanza dell'algoritmo di subentro è dimostrata dal seguente esempio. Durante la trasmissione di una PDU più di un terminale può effettuare l'ascolto del canale ottenendo lo stesso risultato di canale occupato. Se il nuovo tentativo di emissione avvenisse dopo un intervallo di tempo fisso, tutti i terminali eseguirebbero le operazioni di ascolto del canale alle stesse distanze temporali reciproche in cui hanno effettuato la medesima operazione la prima volta. Ciò causerebbe un ulteriore ritardo nella emissione per effetto dell'occupazione del canale.

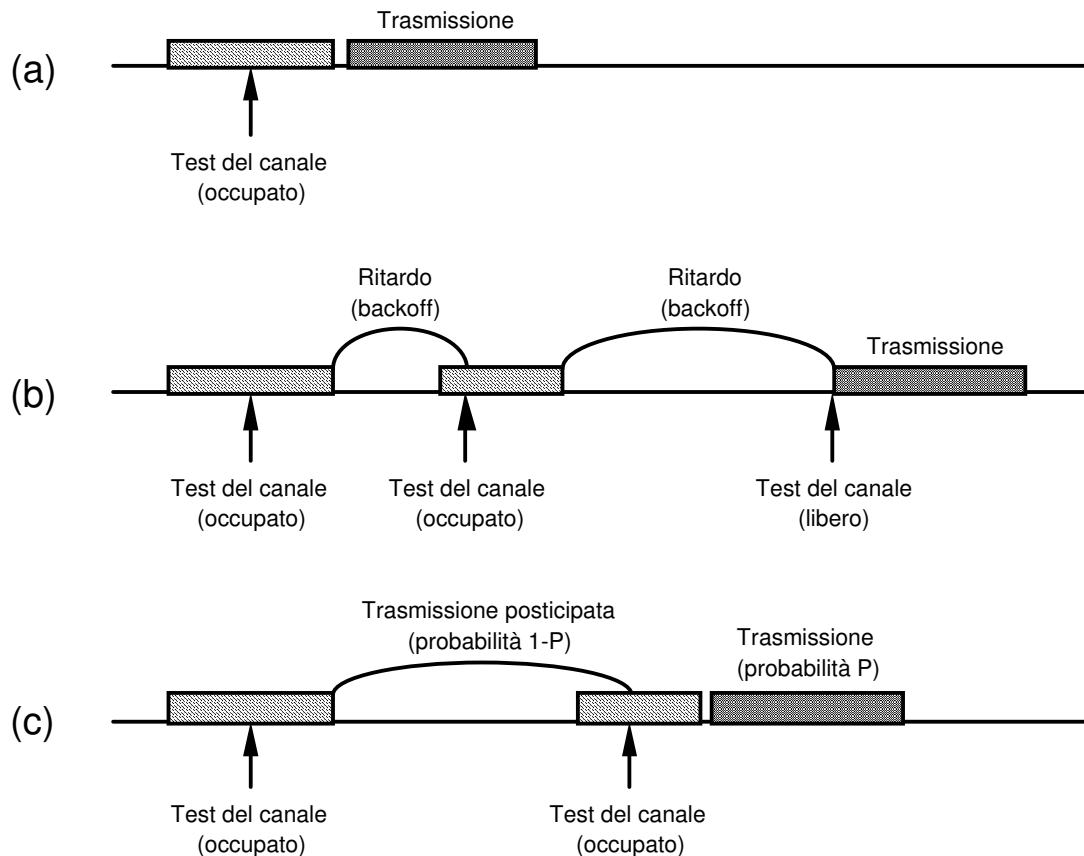


Fig. VI.17 - Effetto delle procedure di persistenza:

- a) *1-persistente*;  
 b) *0-persistente*;  
 c) *p-persistente*.

Confrontando le varie procedure di persistenza, si può osservare che:

- la procedura *1-persistente* potrebbe portare a buoni valori di traffico smaltito in quanto il canale non può andare a riposo finché esiste almeno un terminale che deve emettere, ma ha lo svantaggio di causare una collisione, con probabilità 1, se più terminali ascoltano il canale durante una emissione;
- la procedura *0-persistente* è in grado di ridurre questo svantaggio rendendo casuale l'istante di inizio di un tentativo di trasmissione;
- la procedura *p-persistente* consente, valutando opportunamente il valore del parametro  $p$ , di ottimizzare le prestazioni del protocollo in funzione dei particolari parametri di rete e di traffico incontrati.

Come appare chiaro dalla descrizione del protocollo CSMA, può accadere che, anche se il canale è rivelato libero, due o più terminali emettano contemporaneamente: ciò provoca una collisione e causa quindi la perdita delle informazioni. Infatti, supponendo che il ritardo di propagazione da estremo a estremo sul canale sia uguale a  $T_p$ , un terminale X, che abbia inserito una unità

informativa in rete può subire interferenze da parte dei terminali che hanno ascoltato il canale trovandolo libero e che hanno quindi cominciato a emettere nell'intervallo compreso fra  $-T_p$  e  $+T_p$  rispetto all'inizio della emissione dell'unità informativa da parte del terminale X. Pertanto, nel protocollo CSMA, l'*intervallo di vulnerabilità* e cioè l'intervallo massimo di tempo entro il quale una unità informativa può subire collisione è uguale a  $2T_p$ .

È evidente che l'intervallo di vulnerabilità è un parametro che influenza pesantemente le prestazioni del protocollo, giacché il numero di collisioni, e quindi l'intervallo di tempo necessario a emettere con successo una unità informativa, risulta crescente con l'aumentare dell'intervallo di vulnerabilità e del traffico offerto. Tanto più la rete è estesa, tanto più elevati sono i valori del tempo di vulnerabilità e della probabilità di collisione. Per questa ragione il protocollo CSMA non può essere utilizzato in reti molto estese, pena una sensibile riduzione della loro portata massima.

Il protocollo CSMA/CD (Collision Detection) si distingue dal CSMA per il fatto che i terminali oltre alla strategia "ascolta prima di parlare" utilizzano anche quella "ascolta mentre parli" per tutta la durata della emissione. In questo modo tutte i terminali sono in grado di rivelare le collisioni confrontando la sequenza di bit emessa con quella ricevuta. Tale accorgimento ha il vantaggio di consentire l'interruzione della emissione di una unità informativa nell'istante in cui viene la collisione viene rivelata e quindi di ridurre il tempo di occupazione non utile del canale.

Quando viene rivelata una collisione, la emissione è, come detto, interrotta. L'istante in cui sarà effettuato il tentativo successivo è determinato utilizzando l'algoritmo di subentro.

La durata massima di un periodo di collisione può essere calcolata con riferimento alla Fig. VI.18. Se  $T_1$  è il tempo necessario alla rivelazione di una collisione e se  $T_2$  è il tempo di permanenza del canale nello stato di avvenuta collisione, il tempo totale necessario  $T$  affinché tutte i terminali interrompano la emissione è

$$T = 2 T_p + T_1 + T_2$$

ove  $T_p$  è il ritardo di propagazione da estremo a estremo.

Infatti se i terminali A e B sono poste agli estremi del canale ed iniziano a emettere rispettivamente negli istanti  $t_{0A}$  e  $t_{0A} + T_p$ , il terminale A è in grado di rilevare la collisione nell'istante  $t_{0A} + 2T_p + T_1$  per poi mantenere lo stato di collisione del canale fino all'istante  $t_{0A} + 2T_p + T_1 + T_2$ , istante entro il quale tutte

le emissioni sono interrotte.

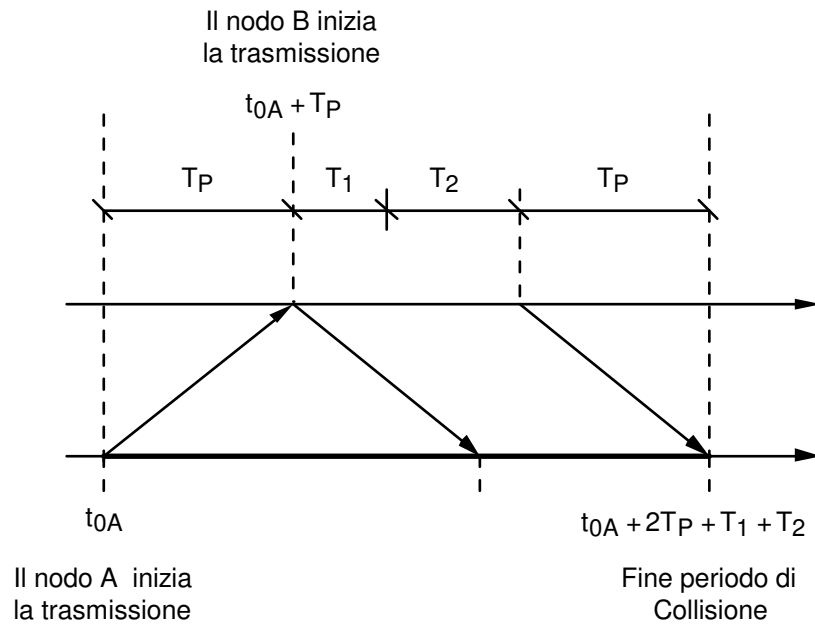


Fig. VI.18 - Illustrazione del tempo di rivelazione di una collisione.

Dall'esame ora effettuato emerge che per mettere in condizione tutte i terminali di rivelare una collisione, una PDU di strato MAC CSMA/CD non può avere lunghezza inferiore ad un certo valore. Tale valore è pari a  $(2T_P + T_1) F_c$ , dove  $F_c$  è il ritmo binario in linea. Infatti, se la PDU avesse una lunghezza inferiore, nel caso venissero a collisione i terminali posti a distanza massima sul bus, questi non avrebbero la possibilità di rivelare la collisione stessa e quindi non avrebbero modo di effettuare la ri-emissione della PDU.

Il problema più importante che occorre considerare nell'uso del protocollo CSMA/CD, ed in generale in tutti i protocolli ad accesso casuale, riguarda la degradazione delle prestazioni ad alto traffico. In particolare, questa si verifica poiché, quando il traffico offerto al bus cresce, il numero delle collisioni tende ad aumentare e quindi aumenta il tempo in cui il canale non è disponibile per le emissioni. Se il traffico supera un determinato valore di soglia, la durata di occupazione del canale per le collisioni diviene via via preponderante rispetto a quella spesa in emissioni utili, fino ad impedire le emissioni stesse. Per questa ragione il rendimento di utilizzazione del canale trasmissivo non può andare oltre a valori nell'intervallo 0.6÷0.7.

Con riferimento allo standard IEEE 802.3, le funzioni principali del livello MAC-CSMA/CD, per quanto riguarda la emissione e la ricezione delle MAC PDU, sono le seguenti:

a) nella emissione delle MAC PDU:

- accettare i dati dallo strato LLC e costruire la PDU;
- presentare un flusso di dati seriale allo strato fisico per la codifica e per la successiva trasmissione;
- ritardare la emissione se il canale è occupato;
- se è rivelata una collisione, interrompere la emissione, svolgere una procedura di *imposizione di collisione* (collision enforcement) per segnalare l'evento agli altri terminali ed eseguire l'algoritmo di subentro per decidere quando dovrà essere riemessa la PDU;
- assicurare che due PDU consecutive siano separate da un intervallo di durata non inferiore ad un valore specificato (*tempo di intertrama*) e che le PDU abbiano lunghezza non inferiore ad un valore opportuno, in modo che ogni terminale possa rivelare eventuali collisioni.

b) nella ricezione delle MAC PDU:

- ricevere un flusso seriale di dati dallo strato fisico;
- presentare allo strato LLC le PDU indirizzate al terminale locale (direttamente o in modo diffusivo);

L'algoritmo di subentro definito nello standard IEEE 802.3 è di tipo *esponenziale binario troncato* (truncated binary exponential backoff) ed ha lo scopo di determinare l'intervallo di tempo al termine del quale avviene un nuovo tentativo di emissione dopo una collisione. Assumendo come unità di tempo l'intervallo temporale corrispondente al ritardo di propagazione sul bus da estremo ad estremo, l'algoritmo di subentro fornisce, per la  $n$ -ma ri-emissione, un numero casuale intero uniformemente distribuito nell'intervallo  $(0, 2^k)$  dove  $k=[0, \min(n,10)]$

La procedura di imposizione di collisione ha lo scopo di garantire che la durata della collisione sia sufficiente a che tutti i terminali coinvolti in essa siano in grado di rivelare la collisione. La configurazione di bit da emettere per segnalare l'avvenuta collisione non è specificata e può essere qualunque, la sua lunghezza è invece fissata in 32 bit.

In Fig. VI.19 è riportata la struttura delle MAC PDU del protocollo CSMA/CD.

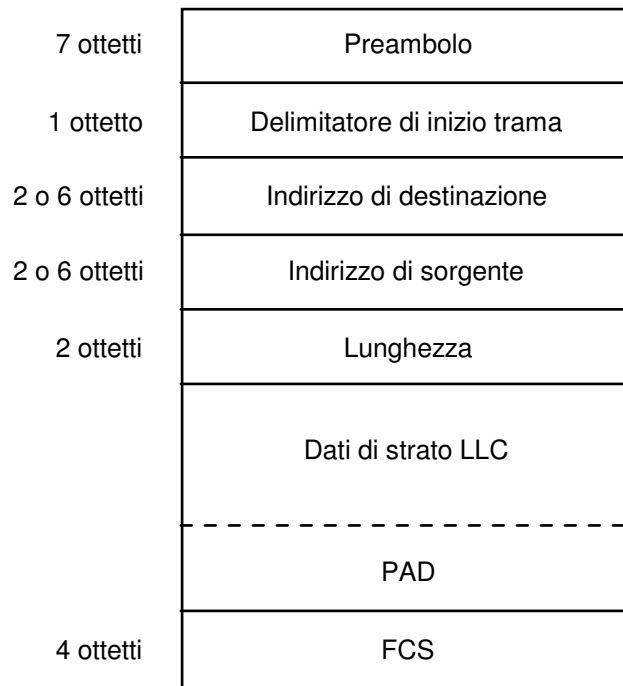


Fig. VI.19 - Struttura della MAC PDU del protocollo CSMA/CD.

In particolare il significato dei singoli campi è il seguente:

- *Preambolo*: ha lunghezza uguale a sette ottetti ed assicura la sincronizzazione del terminale con l'inizio della MAC PDU;
- *Delimitatore di inizio trama*: è una sequenza di otto bit (10101011) che indica l'inizio della MAC PDU;
- *Indirizzi di destinazione e di sorgente*: indicano rispettivamente il terminale di destinazione e di sorgente; possono avere lunghezza uguale a 16 o 48 bit;
- *Lunghezza*: indica la lunghezza (in ottetti) del campo dati di strato LLC e comprende anche la lunghezza dell'eventuale campo PAD;
- *Dati di strato LLC*: questo campo contiene i dati dello strato LLC;
- *PAD*: è inserito nel caso che la lunghezza del campo dati di strato LLC non sia sufficiente per raggiungere la lunghezza minima di PDU necessaria per assicurare la correttezza delle operazioni di riconoscimento delle collisioni;
- *FCS*: ha una lunghezza di 32 bit con l'usuale significato chiarito nell'Unità 7, con riferimento alle PDU dello strato di collegamento; è calcolato sulla base dei bit della trama compresi tra il campo delimitatore di inizio trama ed il campo FCS.

Per quanto riguarda lo stato fisico, lo standard IEEE 802.3 prevede un ritmo binario uguale a 10 Mbit/s con trasmissione numerica in banda base. Un singolo segmento di rete deve avere una lunghezza inferiore a 500 metri, mentre

la distanza massima tra due terminali non può superare 1500 metri. Per aumentare questa distanza massima si possono utilizzare ripetitori. In questo caso tra due MAU non possono essere inseriti più di 4 ripetitori e la distanza tra di essi non può superare 1000 metri. Il numero massimo di terminali collegabili alla rete è 1024.

Infine lo standard prevede che sia comunque rivelata una collisione se la potenza del segnale ricevuto è superiore alla somma di quelle emesse da due terminali. La trasmissione è effettuata utilizzando il codice di Manchester (cfr. par. II.3) in modo da facilitare l'estrazione della temporizzazione da parte dei terminali.

## VI.7 Protocollo a testimone su bus

Il protocollo a testimone su bus (token bus) è ad accesso controllato in quanto consente ai terminali connessi ad una rete a bus di accedere al mezzo solo quando sono in possesso di una particolare PDU di controllo chiamata *testimone*. È descritto nello standard IEEE 802.4.

Il testimone è trasferito da un terminale all'altro lungo un *anello logico* sul bus. La definizione della sequenza dei terminali sull'anello logico è completamente indipendente dalla loro posizione fisica sul bus. Un terminale viene in possesso il testimone quando rivela, sul bus, il transito del testimone recante il proprio indirizzo.

Durante il funzionamento normale, la rete passa ciclicamente attraverso due fasi: 1) la *fase di emissione*, in cui il terminale che possiede il testimone emette la propria unità informativa; 2) la *fase di trasferimento del testimone* in cui il terminale che ha concluso la sua emissione trasferisce il testimone al terminale che lo segue sull'anello logico, dopo aver posto l'indirizzo di questo all'interno del testimone stesso.

Per consentire il trasferimento corretto del testimone tra i vari terminali, ognuno di questi deve conoscere il proprio indirizzo, quello del terminale che lo precede e quello del terminale che lo segue sull'anello logico.

In Fig. VI.20 è rappresentato un possibile anello logico ricavato tra i terminali di un bus bidirezionale.

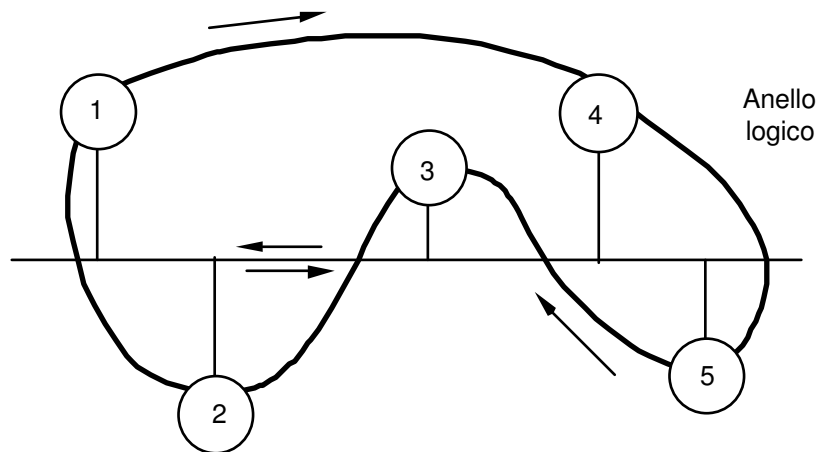


Fig. VI.20 - Creazione di un anello logico su un bus bidirezionale.

È evidente che la sequenza dei terminali sull'anello logico può essere qualsiasi e non dipende in alcun modo dalla disposizione fisica di questi sul bus. La circolazione del testimone avviene nel seguente modo. Un terminale, ad esempio il 5, rivela sul bus la presenza del testimone a lui indirizzato; se ha informazione da emettere, esegue la emissione. Al termine di questa rilascia il testimone con l'indirizzo del terminale 3, che è quello successivo nell'anello logico.

Una caratteristica fondamentale del protocollo a testimone su bus è che il *tempo di ciclo*, e cioè il tempo che intercorre tra due successive visite del testimone ad uno stesso terminale e quindi tra due successive opportunità di emissione, ha un valore massimo finito ed uguale alla somma dei tempi di trasferimento del testimone e di emissione delle MAC PDU in tutti i terminali. Tale caratteristica consente quindi di garantire ad un terminale, in qualsiasi condizione, un valore minimo di traffico smaltibile dalla rete.

Da tale caratteristica emerge che il principale vantaggio offerto dal protocollo a testimone su bus, rispetto a quello CSMA/CD, consiste nell'assenza di fenomeni degenerativi delle prestazioni. I protocolli a testimone riescono infatti a garantire un valore stabile di traffico smaltito anche se il traffico offerto supera la portata massima della rete.

I vantaggi precedentemente elencati sono ottenuti al prezzo di una considerevole complicazione della struttura della rete e delle funzionalità di ogni singolo terminale. Infatti, ognuno di questi, oltre a quelle di accesso, deve eseguire le seguenti funzioni di gestione:

- inizializzazione dell'anello logico e creazione del testimone;
- gestione del testimone, cioè sua ri-emissione in caso di distruzione, riconoscimento e cancellazione di testimoni multipli, ecc.;

- inserimento dei terminali nell'anello logico e loro rimozione;
- recupero delle funzionalità di rete in caso di guasti;
- riconoscimento ed aggiornamento degli indirizzi.

La gestione del testimone è l'elemento più delicato del protocollo. Infatti per cause accidentali può accadere che il testimone sia perso, causando l'interruzione del servizio di rete. Le ragioni che possono portare alla perdita del testimone sono da individuarsi nel guasto di un terminale, che non rilancia il testimone, o negli errori di linea, che rendono inintelligibile il contenuto del testimone stesso. Per fronteggiare situazioni di questo tipo occorre prevedere opportune procedure di controllo della rete.

Il formato generale delle MAC PDU impiegate nel protocollo a testimone sul bus è riportato in Fig. VI.21.

X ottetti	Preambolo
1 ottetto	Delimitatore di inizio trama
1 ottetto	Controllo
2 o 6 ottetti	Indirizzo di destinazione
2 o 6 ottetti	Indirizzo di sorgente
	Dati di strato LLC
4 ottetti	FCS
1 ottetto	Delimitatore di fine trama

Fig. VI.21 - Struttura della MAC PDU del protocollo a testimone su bus.

I campi in essa contenuti hanno il seguente significato:

- *preambolo*: è una sequenza trasmessa prima di qualsiasi MAC PDU per sincronizzare il terminale remoto; deve avere lunghezza tale da assicurare una durata di almeno 2  $\mu$ s (per esempio per ritmi binari di 10 Mbit/s deve essere lunga almeno 3 ottetti);
- *delimitatore di inizio trama*: un ottetto avente struttura fissa che indica l'inizio della trama ed è codificato come *NNNN000*, dove *N* è un simbolo

MAC non utilizzato per i dati e ottenuto con la tecnica della violazione del codice di Manchester utilizzato nella trasmissione;

- *controllo*: un ottetto utilizzato per distinguere tra PDU di controllo MAC, PDU di dati di strato LLC e PDU di gestione dei terminali;
- *indirizzi di destinazione e di sorgente*: identificano rispettivamente il terminale di destinazione e di sorgente; possono avere lunghezza uguale a 16 o 48 bit.
- *dati di strato LLC*: contiene i dati dello strato LLC;
- *FCS*: è lungo quattro ottetti per la rivelazione degli errori di trasmissione;
- *delimitatore di fine trama*: un ottetto, che indica la fine della PDU, codificato come *NNINNIIE*, dove *N* è lo stesso simbolo utilizzato nel delimitatore d'inizio di trama, *1* è un bit uguale ad "1", *I* è il *bit intermedio* ed *E* è il *bit d'errore*.

I bit *I* ed *E* sono gestiti dai ripetitori di rete ed hanno il seguente significato:

*bit I*: se posto ad 1 indica che la stazione in trasmissione deve emettere altre PDU (questo bit aiuta il ripetitore nella determinazione di che cosa segue il campo delimitatore di fine trama);

*bit E*: è posto ad 1 dal primo ripetitore che riveli un errore nella PDU.

A differenza del protocollo CSMA/CD, nel protocollo a testimone su bus non è necessario che la PDU abbia una lunghezza minima. In questo caso, infatti, non essendo presenti collisioni, una emissione ha sempre esito positivo. La lunghezza del bus non deve essere quindi tenuta in conto. Ciò, evidentemente, aumenta l'efficienza del protocollo.

Il formato del testimone è mostrato in Fig. VI.22. Esso differisce da quello di una generica MAC PDU per la mancanza del campo di dati di strato LLC.

X ottetti	Preambolo
1 ottetto	Delimitatore di inizio trama
1 ottetto	0 0 0 0 1 0 0 0
2 o 6 ottetti	Indirizzo di destinazione
2 o 6 ottetti	Indirizzo di sorgente
4 ottetti	FCS
1 ottetto	Delimitatore di fine trama

Fig. VI.22 - Formato della PDU token nel protocollo a testimone su bus.

Lo standard IEEE 802.4 stabilisce in dettaglio la procedura di trasferimento del testimone. In ogni istante un solo terminale è in possesso del testimone ed ha quindi il diritto di emettere. Alla fine della emissione dei dati, il testimone è trasferito al terminale successivo. Tale operazione è ripetuta se, entro un opportuno intervallo di tempo, il terminale non rivela nessuna PDU valida sul bus. Se anche il secondo tentativo non ha successo, il terminale in questione inizia un'opportuna procedura di recupero.

La procedura di recupero del testimone è basata sulla emissione di una PDU di controllo, detta *Who\_Follows*, che contiene l'indirizzo del terminale che non ha fornito risposta. Tutti i terminali della rete confrontano il valore del campo di dati della PDU *Who\_Follows* ricevuta con l'indirizzo del proprio predecessore (cioè del terminale che normalmente gli spedisce il testimone). Il terminale il cui predecessore ha indirizzo uguale a quello letto diviene il nuovo successore del terminale in possesso del testimone. Questo terminale trattiene il testimone e risponde alla sorgente della PDU *Who\_Follows* con la PDU *Set\_Successor*, nel cui campo di dati è scritto il suo indirizzo. Risultato di questa procedura di recupero è l'esclusione del terminale in errore o guasto dalla rete.

## VI.8 Protocollo a testimone su anello

Il protocollo a testimone su anello (token ring) è basato sullo stesso principio di quello a testimone su bus, ma è definito per essere utilizzato su reti con topologia da anello. È descritto nello standard IEEE 802.5.

Il testimone circola nell'anello dando ciclicamente ad ogni terminale

l'opportunità di accedere alla rete (Fig. VI.23a). Un terminale che è in attesa di emissione può impadronirsi del testimone e iniziare la emissione quando il testimone transita attraverso la propria interfaccia (Fig. VI.23b). Il terminale che inserisce una unità informativa in rete deve successivamente rimuoverla ed emettere un nuovo testimone (Fig. VI.23c).

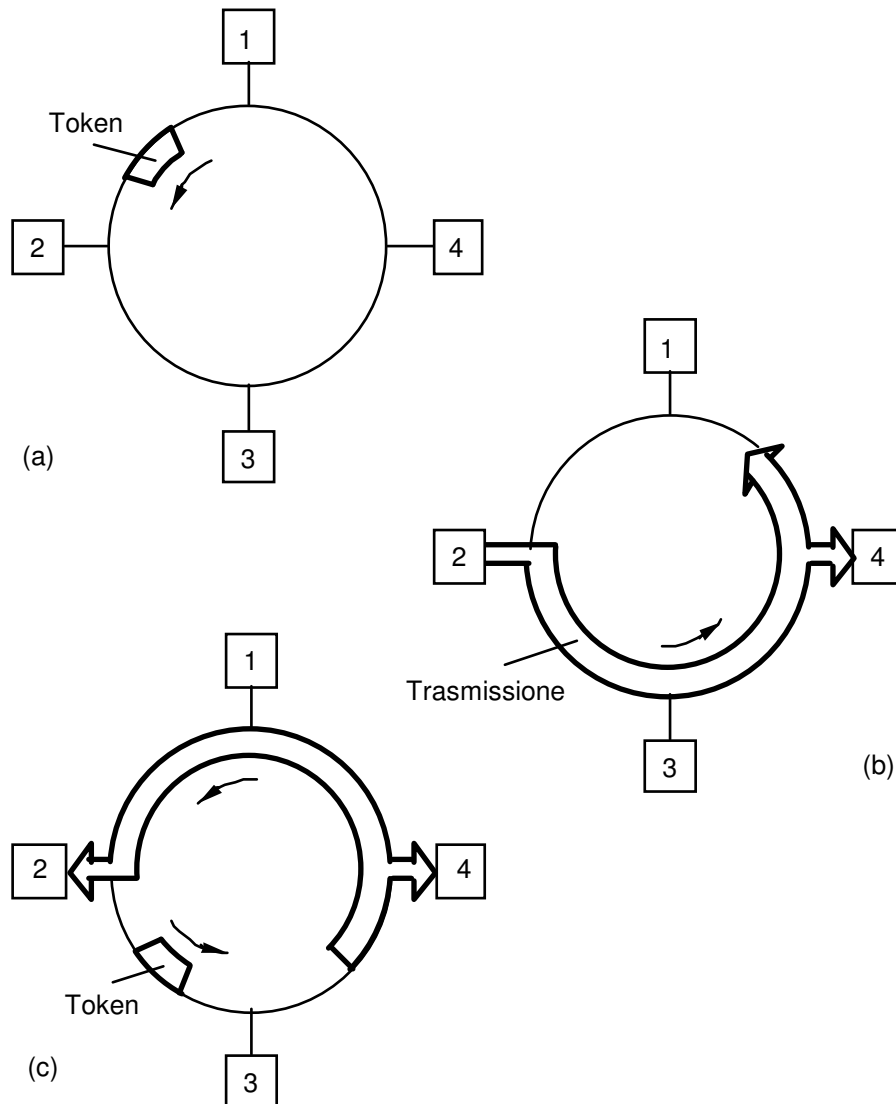
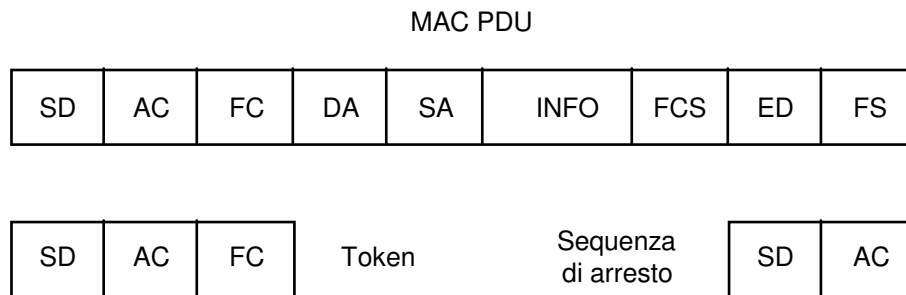


Fig. VI.23 - Fasi del protocollo a testimone su anello:

- a) *circolazione del token;*
- b) *cattura del token da parte di un terminale;*
- c) *riemissione del token.*

L'unità di dati nell'anello è, come per gli altri protocolli, la MAC PDU, la cui struttura è rappresentata in Fig. VI.24. In tale figura sono inoltre mostrati i formati del testimone e della *sequenza di arresto*. Quest'ultima è utilizzata per segnalare l'interruzione prematura della emissione di una trama, e può essere introdotta in qualsiasi momento durante la emissione.



SD : Delimitatore d'inizio trama (1 ottetto);  
 AC : Controllo d'accesso (1 ottetto);  
 FC : Controllo di trama (1 ottetto);  
 DA : Indirizzo di destinazione (2 o 6 ottetti);  
 SA : Indirizzo di sorgente (2 o 6 ottetti);  
 INFO : Informazioni;  
 FCS : Frame Check Sequence (4 ottetti);  
 ED : Delimitatore di fine trama (1 ottetto);  
 FS : Stato della trama (1 ottetto).

Fig. VI.24 - Struttura della MAC PDU, del token e della sequenza di arresto del protocollo a testimone su anello.

Il formato e la funzione di ciascun campo della MAC PDU è riportato in Fig. VI.25.

In particolare si possono distinguere:

- *delimitatore di inizio trama* (SD, Starting Delimiter): è una sequenza di otto simboli; quattro simboli di questo campo sono codificati attraverso una opportuna violazione del codice di Manchester;
- *controllo di accesso* (AC, Access Control): contiene le informazioni per la gestione delle priorità e l'identificatore del testimone; in particolare il significato dei bit è il seguente:
  - i primi tre bit indicano la *priorità* del testimone e pertanto quale classe di utenti è autorizzata ad utilizzarlo;
  - il quarto bit è il *bit di testimone*, ed il suo valore è "0" in un testimone ed "1" nelle altre MAC PDU; quando un terminale con una PDU in attesa di emissione rivela un testimone circolante sull'anello, con priorità inferiore od uguale a quella della PDU in attesa, pone ad "1" il bit di testimone, trasformando così questo nella sequenza di inizio della MAC PDU;
  - il bit M è detto *bit di monitoraggio*, ed è utilizzato dalla stazione di controllo della rete (che è l'unica autorizzata a modificare tale campo) per rimuovere quelle PDU che, per effetto di errori di trasmissione, si trovassero a circolare indefinitamente nell'anello; la stazione di controllo legge il valore del bit *M* di ogni MAC PDU che l'attraversa; se tale bit è uguale a "0", lo sostituisce

- con il valore "1", altrimenti rimuove tale MAC PDU dalla rete;
- gli ultimi tre simboli di tale campo rappresentano i *bit di prenotazione* e permettono ad un terminale di richiedere il livello di priorità con cui deve essere emesso il successivo testimone.
  - *controllo di trama* (FC, Frame Control): i primi due bit permettono di distinguere tra due tipi di PDU:
    - 00 PDU di controllo (contiene unità di dati relative al controllo di accesso);
    - 01 PDU dati (contiene una LLC PDU);nel caso di MAC PDU, i bit 3-8 di tale campo verranno interpretati come bit di controllo, mentre nel caso di PDU di dati essi indicano la priorità dei dati contenuti nella PDU;
  - *indirizzi di destinazione e di sorgente* (DA, Destination Address; SA, Source Address): contengono due indirizzi: quello del terminale di destinazione e quello del terminale di origine; ognuno di questi campi di indirizzo può essere lungo 2 o 6 ottetti, in ogni caso tutte i terminali di una specifica LAN devono avere gli indirizzi di uguale lunghezza;
  - *informazione* (I, Information): può essere lungo 0, 1 o più ottetti e trasporta informazioni relative allo strato MAC, allo strato LLC o alla gestione della rete;
  - *FCS*: è una sequenza di 32 bit ricavata da un polinomio generatore standard di grado 32 e utilizzata per rilevare eventuali errori che abbiano interessato i campi FC, DA, SA ed I della MAC PDU;
  - *delimitatore di fine* (ED, Ending Delimiter): come per il delimitatore di inizio, una PDU valida deve presentare un'opportuna configurazione di simboli nei primi sei sottocampi del campo ED; il sottocampo *I* indica, quando è posto ad "1", che la PDU attuale è una frazione (eventualmente la prima) di una sequenza di PDU; infine se un generico terminale, mediante il controllo di sequenza, rivela un errore sulla PDU, deve porre ad "1" l'ultimo bit del campo ED (bit E).
  - *stato della trama* (FS, Frame Status): i bit *A* e *C* di questo campo sono posti a zero dal terminale sorgente e vengono modificati dal terminale di destinazione per indicare:
    - A = 1 PDU ricevuta correttamente;
    - C = 1 PDU copiata correttamente;gli altri bit di questo campo sono riservati per ulteriori sviluppi del protocollo.



J, K : Simboli di violazione	I/G : Indirizzo individuale o di gruppo
P : Priorità	I/L : Indirizzo universale o locale
T : Token	I : Indicatore di trama intermedia
M : Monitor	E : Indicatore di errore
R : Prenotazione	A : Indicatore di indirizzo riconosciuto
F : Tipo di trama	C : Indicatore di trama copiata
Z : Controllo	X : Bit riservati.

Fig. VI.25 - Formato dei singoli campi della PDU del protocollo a testimone su anello.

Per la rimozione delle PDU dall'anello è, in linea di principio, possibile operare secondo due differenti schemi: *rimozione alla sorgente* e *rimozione a destinazione*.

Nel primo caso, che è quello previsto dallo standard, il nodo sorgente rimuove la PDU dopo che questa ha percorso completamente l'anello; occorre osservare che tale operazione, oltre a consentire un primo scambio di informazioni di riscontro tra sorgente e destinazione, non richiede la decodifica dell'indirizzo di sorgente, e pertanto può avvenire senza introdurre ulteriori ritardi.

Viceversa, nella rimozione operata dal terminale di destinazione, occorre decodificare l'indirizzo di destinazione, e pertanto la ri-emissione della trama dovrà essere ritardata fintantoché non è stato ricevuto il campo DA di questa. Tale schema consente tuttavia una migliore utilizzazione della banda del canale se la lunghezza elettrica dell'anello è maggiore della lunghezza massima di una

PDU.

Lo standard IEEE 802.5 prevede due valori del ritmo binario in linea 1 Mbit/s e 4 Mbit/s. Il numero massimo di terminali connettabili alla rete è uguale a 72 o 260 in relazione al tipo di mezzo trasmissivo utilizzato. Il primo valore si riferisce ad una coppia simmetrica non schermata, mentre il secondo è valido per coppie schermate.

## **VI.9 Lo strato LLC nelle LAN**

Il protocollo di controllo del collegamento logico, definito nello standard IEEE 802.2, ha il compito di:

- realizzare lo scambio delle unità informative tra i terminali connessi alla LAN;
- organizzare il flusso delle LLC PDU;
- gestire e interpretare i comandi e le risposte;
- eseguire le funzioni di rivelazione e di recupero degli errori.

Lo strato LLC è stato definito in modo indipendente rispetto alle diverse alternative di protocolli d'accesso della serie IEEE 802 e può essere quindi applicato a qualsiasi tipo di rete locale.

Per soddisfare i requisiti di un'ampia gamma di applicazioni sono stati definiti due tipi di servizio di trasferimento tra le entità dello strato LLC:

- il servizio *senza connessione* (connectionless);
- il servizio *con connessione* (connection-oriented).

Il servizio senza connessione utilizza un sotto-insieme delle funzionalità del servizio dello strato LLC ed è impiegato quando i protocolli di più alto livello dispongono delle funzionalità necessarie per il controllo degli errori e della sequenzialità delle unità informative scambiate. Questo tipo di servizio non garantisce la consegna delle LLC PDU.

Il servizio con connessione è comparabile con quello offerto da altri standard internazionali come, ad esempio, dall'HDLC o dal protocollo X.25 livello 2 (cfr. par. III.4). Tale servizio garantisce quindi il trasferimento affidabile ed in sequenza delle LLC PDU su di una connessione logica denominata *connessione di strato LLC* (Data Link Connection).

È possibile che un terminale gestisca contemporaneamente varie connessioni logiche con terminali diversi. Ogni connessione logica è instaurata e gestita indipendentemente dalle altre, sia per quanto riguarda la definizione dei suoi parametri, sia dal punto di vista dello scambio informativo.

La possibilità per un terminale di gestire una pluralità di connessioni di

strato LLC con terminali diversi è una diretta conseguenza dell'ambiente multiaccesso tipico delle LAN. Nelle WAN un ramo è, nella maggioranza dei casi, di tipo punto-punto e pone in corrispondenza due nodi della rete; le connessioni di strato di collegamento, se sono multiple, hanno tutte uguale origine e destinazione. Nel caso delle LAN, un unico mezzo trasmissivo collega una molteplicità di terminali e quindi possono coesistere sulla stessa connessione fisica più connessioni logiche distinte con origine e destinazione diverse. È quindi necessario prevedere i mezzi per distinguere tali connessioni.

Nel seguito del paragrafo verranno descritte dapprima (§ VI.9.1) le principali caratteristiche del servizio di strato LLC; successivamente (§ VI.9.2) si parlerà del relativo protocollo di strato nelle sue due modalità di funzionamento.

#### VI.9.1 Servizio di strato LLC.

In accordo con quanto specificato dal modello OSI, nel descrivere il servizio di strato LLC, facciamo riferimento alle interazioni tra le entità dello strato LLC e quello dello strato superiore: le seconde sono indicate come *utenti del servizio LLC* (o più brevemente come utenti, quando tale termine non consenta equivocationi), mentre le prime sono i *fornitori del servizio*.

Nel servizio di strato LLC le interazioni tra utente e fornitore sono descritte attraverso *primitive*, che sono raggruppate in *classi*. Ogni classe descrive un elemento del servizio di strato LLC. Le classi di primitive definite nell'ambito degli standard della famiglia IEEE 802 comprendono (Fig. VI.26):

- la primitiva di *richiesta* (REQUEST), che consente all'utente locale di richiedere un servizio al fornitore;
- la primitiva di *indicazione* (INDICATION), che consentono al fornitore di indicare all'utente un evento, che si è presentato all'interno allo strato e che è significativo per l'utente stesso;
- la primitiva di *conferma* (CONFIRM), che permette al fornitore di confermare richiedente l'esito del servizio offerto.

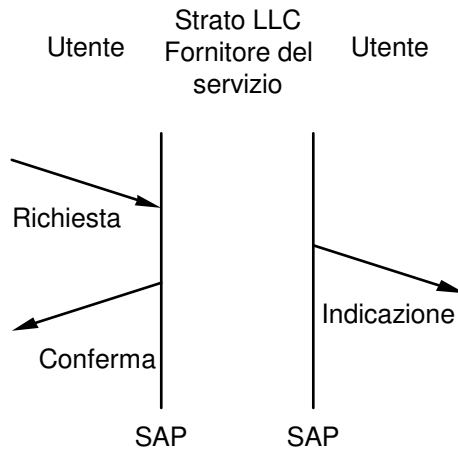


Fig. VI.26 - Primitive di servizio definite per lo strato LLC.

Rispetto alla struttura generale del modello OSI, si può notare la mancanza della primitiva di *risposta* (RESPONSE), utilizzata per consentire all'utente di rispondere alla primitiva di indicazione. La primitiva di risposta non è stata introdotta perché i servizi previsti negli standard IEEE 802 sono strutturati in modo che, attraverso il servizio di strato LLC, un utente non possa ottenere informazioni sull'utente remoto.

Come precedentemente accennato, lo strato LLC fornisce due tipi di servizio di trasferimento e cioè quello senza connessione e quello con connessione. Il servizio senza connessione prevede lo scambio di *unità di dati del servizio di collegamento* (Link Service Data Unit - LSDU) senza la preventiva instaurazione di una connessione di strato LLC. Il servizio con connessione fornisce i mezzi per instaurare, utilizzare, gestire e abbattere una connessione di strato LLC. Nelle Tab. VI.1 e Tab. VI.2 sono riportate, per i due tipi di servizio di trasferimento le relative primitive con i parametri definiti dal colloquio tra due strati omologhi.

<b>Primitiva di servizio</b>	<b>Parametri</b>
L_DATA Request	Indirizzo locale Indirizzo remoto LSDU Classe di servizio
L_DATA Indication	Indirizzo locale Indirizzo remoto LSDU Classe di servizio

Tab. VI.1 - Primitive della classe di servizio senza connessione.

<b>Fase della comunicazione</b>	<b>Primitiva di servizio</b>	<b>Parametri</b>
Instaurazione della Connessione	L_CONNECT Request	Indirizzo locale Indirizzo remoto Classe di servizio
	L_CONNECT Indication	Indirizzo locale Indirizzo remoto Classe di servizio Stato
	L_CONNECT Confirm	Indirizzo locale Indirizzo remoto Stato
Trasferimento dati	L_DATA_CONNECT Request	Indirizzo locale Indirizzo remoto LSDU
	L_DATA_CONNECT Indication	Indirizzo locale Indirizzo remoto LSDU
	L_DATA_CONNECT Confirm	Indirizzo locale Indirizzo remoto Stato
Riinizializzazione della Connessione	L_RESET Request	Indirizzo locale Indirizzo remoto
	L_RESET Indication	Indirizzo locale Indirizzo remoto Motivo
	L_RESET Confirm	Indirizzo locale Indirizzo remoto Stato
Controllo di flusso	L_CONNECTION_FLOW CONTROL Request	Indirizzo locale Indirizzo remoto Quantità di dati permessa

verso lo strato 3	L_CONNECTION_FLOW CONTROL Indication	Indirizzo locale Indirizzo remoto Quantità di dati permessa
Disconnessione della Connessione	L_DISCONNECT Request	Indirizzo locale Indirizzo remoto
	L_DISCONNECT Indication	Indirizzo locale Indirizzo remoto Motivo
	L_DISCONNECT Confirm	Indirizzo locale Indirizzo remoto Stato

Tab. VI.2 - Primitive della classe di servizio con connessione.

Occorre considerare anche le primitive di servizio che descrivono le interazioni tra le entità dello strato LLC e quelle dello strato MAC nella fase di trasferimento dei dati. Queste primitive consentono la definizione del servizio che lo strato MAC deve fornire ad un'entità locale dello strato LLC per consentire, a quest'ultima, di scambiare unità di dati con un'entità remota (*servizio di strato MAC*). Tali primitive sono descritte nella Tab. VI.3.

Nel caso di servizio di trasferimento senza connessione, la emissione di una LLC PDU può avvenire in un qualsiasi istante senza la preventiva instaurazione della connessione logica. L'unica classe di primitive necessarie è quindi quella relativa al trasferimento dei dati (L\_DATA). La primitiva di richiesta L\_DATA è emessa dallo strato superiore per richiedere il trasferimento di una LSDU. La primitiva di indicazione L\_DATA è invece utilizzata dallo strato LLC per avvisare l'utente della ricezione di una LSDU.

<b>Fase della comunicazione</b>	<b>Primitiva di servizio</b>	<b>Parametri</b>
Trasferimento dei dati verso lo strato MAC	MA_DATA Request	Indirizzo di destinazione MAC_SDU
	MA_DATA Indication	Indirizzo di destinazione Indirizzo di sorgente MAC_SDU Stato Classe di servizio richiesta
	MA_DATA Confirm	Stato Classe di servizio fornita

Tab. VI.3- Primitiva di scambio dati con lo strato MAC.

La procedura di instaurazione di una connessione logica da parte di un utente dello strato LLC è effettuata per mezzo della classe di primitive L\_CONNECT. L'utente che vuole instaurare una connessione trasferisce allo strato LLC una primitiva di richiesta L\_CONNECT, indicante l'indirizzo dell'utente remoto con cui si vuole realizzare una connessione. Nella stessa primitiva è anche indicato il valore della priorità che si richiede per lo scambio dei dati su quella connessione. L'entità di strato LLC che riceve tale primitiva inizia la procedura di instaurazione della connessione, mediante il protocollo di strato LLC. L'entità remota, quando riceve la richiesta di instaurazione di una connessione, trasferisce al proprio utente la primitiva di indicazione L\_CONNECT. Al termine della procedura, all'utente che l'ha iniziato viene trasferita la primitiva di risposta L\_CONNECT. Entrambe le primitive di indicazione e di risposta contengono un campo che informa se l'instaurazione della connessione ha avuto successo o, eventualmente, le ragioni della mancata instaurazione.

#### VI.9.2 Protocollo di strato LLC.

Il protocollo LLC definito nello standard IEEE 802.2 è di tipo orientato al bit ed è basato sullo scambio di PDU aventi struttura simile a quella delle trame HDLC (Fig. VI.27).

Indirizzo di destinazione DSAP (8 bit)	Indirizzo di sorgente SSAP (8 bit)	Controllo  (Y bit)	Informazione  (8*M bit)
---	---	--------------------------	-------------------------------

Y = 8 Funzionamento di tipo 1  
 Y = 16 Funzionamento di tipo 2

Fig. VI.27 - Struttura della LLC PDU.

I due campi di *indirizzo* (DSAP e SSAP) hanno ambedue lunghezza uguale a un otetto ed identificano i punti di accesso al servizio (SAP) di destinazione e di sorgente della LLC PDU. La struttura dei due campi è mostrata in Fig. VI.28.

DSAP							SSAP							
I/G	D	D	D	D	D	D	C/R	S	S	S	S	S	S	S

I/G = 0 Indirizzo individuale I/G = 1 Indirizzo di gruppo	C/R = 0 Comando C/R = 1 Risposta
--	-------------------------------------

Fig. VI.28 - Struttura del campo di indirizzo di una LLC PDU.

Il primo bit del campo DSAP indica se l'indirizzo di destinazione è di tipo individuale o di gruppo, mentre il primo bit del campo SSAP stabilisce se la LLC PDU è un comando o una risposta. I restanti sette bit dei due campi sono i veri e propri bit di indirizzo. La configurazione di sette "1" consecutivi del campo DSAP indica un indirizzamento globale; in questo caso la LSDU deve essere consegnata a tutti i SAP attualmente attivi. La configurazione di sette bit "0" rappresenta invece un indirizzamento nullo e non identifica alcun SAP.

Il campo di *controllo* consiste di uno o due ottetti che identificano il tipo ed i numeri di sequenza delle LLC PDU. Lo scambio delle LLC PDU può avvenire secondo due tipi di funzionamento:

- 1) *funzionamento di tipo 1* : le unità di dati sono scambiate senza instaurare una connessione di strato LLC; non sono previsti meccanismi di recupero degli errori o di riscontro e non è effettuato alcun controllo di flusso.
- 2) *funzionamento di tipo 2*: le unità di dati sono trasferite su una connessione di strato LLC, che deve essere opportunamente instaurata, mantenuta ed abbattuta; sono previste procedure di controllo di flusso, di recupero di errori e di riscontro delle trame ricevute correttamente.

Il funzionamento di tipo 2 è applicato nel caso di connessioni bilanciate (cfr. par. III.3). In questo caso le LLC PDU sono numerate e viene attivata una

numerazione indipendente per ogni coppia di SAP di sorgente e di destinazione

Sono definite tre tipi di unità di dati: a) le LLC PDU *informative numerate*; b) le LLC PDU *di supervisione*; c) le LLC PDU *non numerate*. Il funzionamento di tipo 1 prevede solo lo scambio di LLC PDU non numerate, mentre il tipo 2 utilizza anche LLC PDU di supervisione e informative numerate.

Oltre alla classificazione dei due tipi di funzionamento, è prevista anche la classificazione delle entità di strato LLC. Una LLC entità di *classe 1* può fornire solo un funzionamento di tipo 1, con indirizzamento globale, individuale o nullo, senza riscontri e senza controllo di flusso. Una LLC entità di *classe 2* può fornire funzionamenti sia di tipo 1 che di tipo 2.

L'insieme dei comandi e delle risposte supportate dalle entità di classe 1 e 2 è indicato in Tab. VI.4.

Tipo di LLC-entità	Funzionamento	Comandi	Risposte
Classe 1	Tipo 1	UI XID TEST	XID TEST
Classe 2	Tipo 1	XID TEST UI	XID TEST
	Tipo 2	I RR RNR REJ SABME DISC	I RR RNR REJ UA DM DISC FRMR

Tab. VI.4 - LLC PDU gestite dalle due classi di LLC-entità.

#### *FUNZIONAMENTO DI TIPO 1.*

Le LLC PDU utilizzate nel funzionamento di tipo 1 sono quelle necessarie per lo scambio di unità informative non numerate, senza riscontro e senza

controllo di flusso. Queste LLC PDU possono essere di tre tipi:

- *informazione non numerata* (Unnumbered Information - UI) per gestire la trasmissione di informazioni non numerate ad una o più entità di destinazione;
- *prova* (TEST) per verificare lo stato della corrispondenza fra le LLC-entità;
- *identificazione dello scambio* (Exchange Identification - XID) per l'esecuzione di procedure di gestione dello strato; la LLC-entità remota deve rispondere al comando XID con una risposta XID che identifica l'entità stessa e che contiene un campo informativo uguale a quello ricevuto.

La codifica del campo di controllo contenuto nelle PDU definite per le operazioni di tipo 1 è riportata in Fig. VI.29. Il bit *Poll/Final* (P/F) è utilizzato nelle PDU di comando come bit Poll e consente, se posto ad 1, di sollecitare il terminale remoto a emettere una PDU di risposta con il bit Final posto anch'esso ad 1.

1	1	0	0	P	0	0	0	UI
1	1	1	1	P	1	0	1	XID (comando)
1	1	0	0	P	1	1	1	TEST (comando)
1	1	1	1	F	1	0	1	XID (risposta)
1	1	0	0	F	1	1	1	TEST (risposta)

P/F : bit di Poll/Final

Fig. VI.29 - Codifica del campo di controllo delle PDU definite per il funzionamento di tipo 1.

Il bit P/F è utilizzato, ad esempio, nel caso di situazioni di recupero, rese necessarie dal silenzio prolungato del terminale lontano. In questo caso, il terminale che non riceve risposta dal terminale remoto emette una PDU TEST con il bit P=1. Tale PDU richiede all'entità remota di emettere una risposta TEST con il bit F allo stesso valore del bit P ricevuto e con lo stesso campo informativo. Se la procedura non ottiene risposta, il terminale remoto deve essere considerato fuori servizio.

La PDU XID può avere varie applicazioni. Ad esempio, una XID con un

indirizzo nullo sollecita la risposta da parte di tutti i terminali ed è utilizzata per localizzare i terminali stessi; una XID recante un particolare indirizzo di gruppo può essere utilizzata per verificare quali terminali appartengono a quel gruppo.

#### *FUNZIONAMENTO DI TIPO 2.*

Le trame definite nel caso di funzionamento di tipo 2 contengono un campo di controllo analogo a quello definito per le trame del protocollo X.25 livello 2, ma a differenza di queste, l'unica numerazione prevista è modulo 128. Di conseguenza il massimo numero di trame emesse e non riscontrate è uguale a 127.

Le LLC PDU definite nel funzionamento di tipo 2 differiscono fra loro per la struttura del campo di controllo e per la presenza o meno del campo informativo. La struttura del campo di controllo è mostrata in Fig. VI.30. Queste LLC PDU sono di tre tipi:

- a) *Informative (I)*: consentono il trasferimento numerato dei dati; il campo di controllo comprende quattro sottocampi:
  - *identificatore* di LLC PDU di tipo I; il bit è posto a 0;
  - *numero di sequenza in trasmissione*  $N(S)$ ; questo è una variabile modulo 128 che indica il numero d'ordine dell'ultima LLC PDU di tipo I emessa; è aggiornato ad ogni emissione di PDU di tipo I sulla base del valore della *variabile di stato in trasmissione*  $V(S)$ ;
  - bit *Poll/Final (P/F)*; nelle LLC PDU di comando è utilizzato come bit Poll posto ad 1 per sollecitare una risposta da parte dell'entità remota; a sua volta l'entità remota deve rispondere a questa interrogazione con una PDU di risposta che avrà il bit Final posto ad 0;
  - *numero di sequenza in ricezione*  $N(R)$ ; questa una variabile modulo 128 che indica il numero d'ordine della LLC PDU di tipo I attesa; il suo valore deve essere uguale a quello della *variabile di stato in ricezione*  $V(R)$  ed indica che sono state ricevute correttamente tutte le LLC PDU numerate da 0 a  $N(R)-1$ .
- b) *Supervisione (S)*: gestiscono le funzioni di supervisione delle connessioni consentendo: 1) il *riscontro positivo* (comandi e risposte RR, Receive Ready); 2) la *richiesta di ritrasmissione* (comandi e risposte REJ, Reject); 3) la richiesta al terminale lontano di sospendere temporaneamente la emissione di LLC PDU di tipo I (comandi e risposte RNR, Receive Not Ready); il campo di controllo delle LLC PDU di tipo S differisce da quello delle PDU di tipo I per la mancanza del sottocampo  $N(S)$  e contiene due bit che consentono di distinguere le tre PDU di tipo S.

c) *Non Numerate* (U): consentono la emissione di informazioni in operazioni senza riscontro e senza connessione; oltre alle LLC PDU già descritte nel caso di funzionamento di tipo 1 (comandi e risposte XID, comandi e risposte TEST, comando UI), sono previsti: 1) il comando di connessione bilanciata di due LLC-entità (SABME, Set Asynchronous Balanced Mode Extended); 2) il riscontro non numerato (UA, Unnumbered Acknowledgment); 3) il comando di disconnessione (DISC, Disconnect); 4) la risposta DM (Disconnect Mode), impiegata, da parte dell'entità remota, per indicare che non è connessa; 5) la risposta FRMR (Frame Reject), che indica, in operazioni bilanciate, un errore non recuperabile con la ritrasmissione, con conseguente necessità di reinizializzare la connessione di strato LLC.

	1	2	3	4	5	6	7	8	9	10 - 16
PDU Informativa	0	N(S)							P/F	N(R)
PDU di Supervisione	1	0	S	S	X	X	X	X	P/F	N(R)
PDU Non Numerate	1	1	M	M	P/F	M	M	M		

N(S) : Numero di sequenza delle PDU in trasmissione;  
 N(R) : Numero di sequenza delle PDU in ricezione;  
 S : Codice di supervisione;  
 M : Codice operativo;  
 P/F : Bit di Poll/Final;  
 X : Bit riservati;

Fig. VI.30- Codifica del campo di controllo delle PDU definite per il funzionamento di tipo 2.

Le procedure previste per il funzionamento di tipo 2 sono analoghe a quelle del protocollo X.25 livello 2. Per brevità esse non saranno qui commentate, per i dettagli si rimanda il lettore all'unità 7 e allo standard IEEE 802.2.

L'instaurazione di una connessione può essere iniziata da qualsiasi entità di strato LLC. L'entità iniziante emette una SABME PDU e fa partire un temporizzatore. Alla ricezione della UA PDU di risposta, l'entità porrà a zero le variabili V(S) e V(R), fermerà il temporizzatore ed entrerà nella fase dati. Se in risposta ad una SABME PDU, l'entità riceve una DM PDU, l'instaurazione della connessione non è andata a buon fine e l'entità segnalerà l'evento allo strato superiore. Se il temporizzatore si esaurisce prima della ricezione di una risposta, l'entità emette nuovamente una SABME PDU e fa ripartire il temporizzatore. In mancanza di risposte la procedura viene ripetuta per un numero fissato di volte,

dopo di che l'instaurazione è considerata fallita.

## **VI.10 La MAN DQDB**

Lo standard IEEE 802.6, a differenza dei precedenti che trattano esclusivamente protocolli per LAN, è orientato alla definizione dell'architettura di una MAN.

Nell'accezione dello standard, una MAN è una rete in grado di fornire in modo integrato un'ampia gamma di servizi, che vanno da quelli per dati a quelli vocali e video, su di un'area geografica sufficientemente estesa. In generale, una MAN è vista come composta da un insieme di sottoreti connesse tra loro da opportuni dispositivi di interconnessione e di instradamento (bridge, router). Lo standard IEEE 802.6 descrive le funzionalità di una sottorete, denominata Distributed Queue Dual Bus (DQDB), che può essere utilizzata come parte componente di una MAN.

Una rete DQDB è di tipo multiaccesso ed è in grado di fornire un ambiente di comunicazioni integrate. Essa infatti può essere supporto sia di comunicazioni di dati, con trasferimento senza connessione e/o con connessione, sia di comunicazioni di tipo isocrono, come ad esempio quelle vocali. Queste funzionalità sono riassunte nella Fig. VI.31, che mostra l'architettura di una rete DQDB. Come si può notare, lo strato DQDB deve comprendere sia le funzionalità necessarie a regolare l'accesso al canale trasmissivo, sia quelle necessarie all'adattamento dei servizi al tipo di trasferimento adottato in rete. Nel seguito ci occuperemo esclusivamente degli aspetti riguardanti le procedure d'accesso al canale trasmissivo.

La particolare struttura protocollare della rete DQDB nasce allo scopo di fornire un servizio di trasferimento integrato per diverse tipologie di servizi. Nel caso di interconnessione tra LAN, lo strato DQDB accetta in ingresso LLC PDU che provvede a trasferire tra le due LAN a colloquio. In questo caso le funzionalità richieste allo strato DQDB sono esclusivamente quelle di strato MAC. Invece, nel caso di trasferimento di servizi isocroni, lo strato deve anche eseguire tutte quelle funzioni che rendano possibile la ricostruzione del messaggio originale a destinazione.

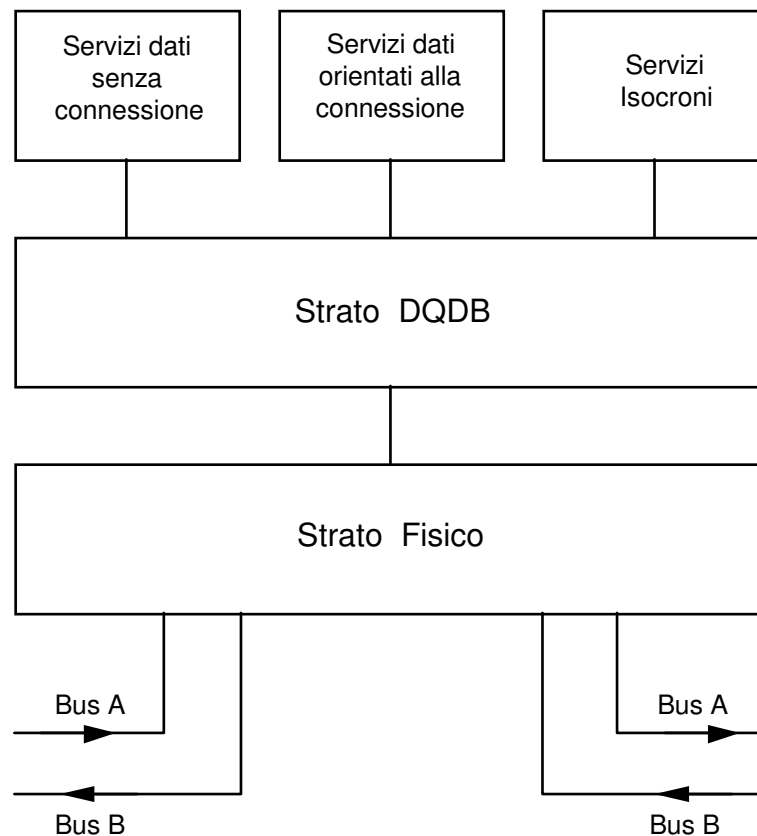


Fig. VI.31 - Architettura logica dei protocolli IEEE 802.6.

La rete DQDB utilizza una topologia a doppio bus unidirezionale. IN questo caso, per aderenza alla terminologia dello standard, i terminali sono chiamati *stazioni di rete*.

La capacità di ogni bus è suddivisa in intervalli temporali (IT) di lunghezza fissa che sono generati dalla *stazione di testa* di ogni bus (head bus). Tali IT possono essere usati per la emissione, da parte stazioni connesse alla rete, in accordo alle regole del protocollo d'accesso. Il flusso di dati è terminato da un opportuno dispositivo posto alla fine di ogni bus.

La struttura di una stazione DQDB è mostrata in Fig. VI.32. Ognuna di queste è costituita da un'*unità d'accesso*, che svolge le funzioni del protocollo DQDB, e da due *interfacce* identiche, una per ogni bus, che eseguono le operazioni di lettura e di scrittura. L'operazione di lettura su un bus precede quella di scrittura ed è completamente passiva, mentre l'operazione di scrittura è ovviamente attiva ed è eseguita semplicemente tramite una funzione "OR" tra il flusso di bit del bus e quello emesso dalla stazione.

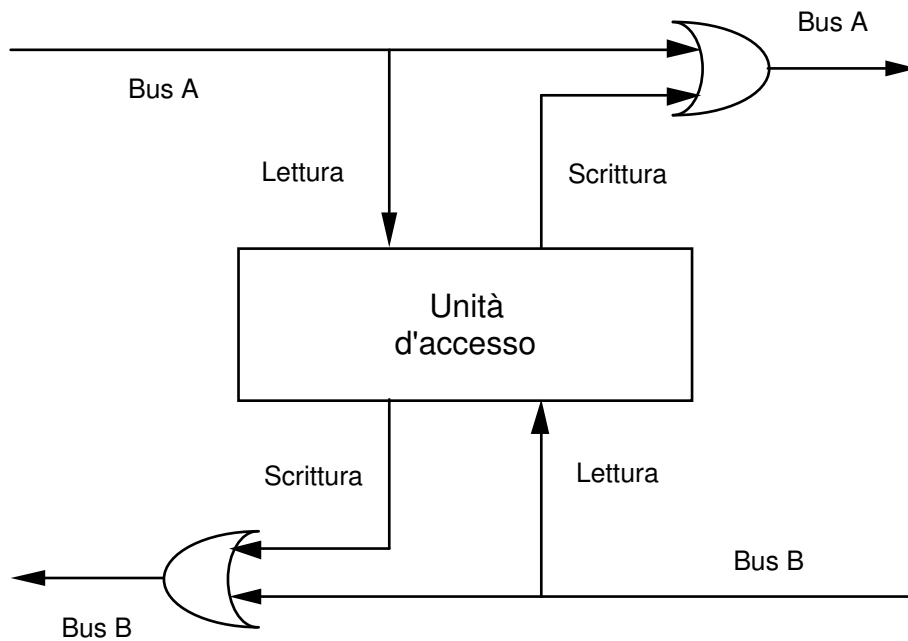


Fig. VI.32 - Struttura di una stazione DQDB.

Il protocollo DQDB prevede due modalità di accesso al bus: 1) un accesso con *arbitraggio a coda* (Queued Arbitrated - QA); 2) un accesso con *pre-arbitraggio* (Pre Arbitrated - PA).

Ognuna di queste modalità utilizza IT di tipo diverso, chiamati rispettivamente *QA* e *PA*. Ogni IT (Fig. VI.33), sia di tipo QA che PA, ha lunghezza fissa, uguale a 53 ottetti, ed comprende: 1) *campo di controllo accesso* (Access Control Field - ACF), di lunghezza uguale ad un ottetto, che contiene le informazioni per l'utilizzazione della sua capacità; 2) un *segmento informativo* (segment), che contiene le informazioni da trasferire.

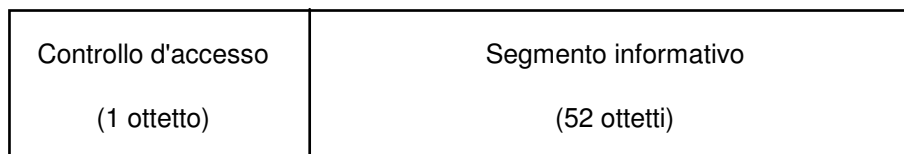


Fig. VI.33 - Struttura di uno slot nel protocollo DQDB.

La struttura del campo ACF è mostrata in Fig. VI.34; i bit che lo compongono hanno il seguente significato:

- *bit di occupato* (B): indica se l'IT è attualmente occupato;
- *bit di tipo* (SL\_TYPE): indica se l'IT è di tipo QA o PA;
- *bit di ricezione del segmento precedente* (PSR): indica se il segmento precedente è stato ricevuto oppure no;
- *quattro bit di richiesta* (R): indicano che un segmento informativo è stato

inserito nella coda distribuita in accordo alla sua priorità (nel seguito supporremo, per semplicità che sia prevista una sola classe di priorità e quindi sia sufficiente un solo bit R);

- l'ulteriore bit del campo ACF è riservato per usi futuri.

Occupato (B) (1 bit)	Tipo (1 bit)	Bit riservati (1 bit)	Ricezione segmento precedente (1 bit)	Richiesta (R) (8*M bit)
----------------------------	-----------------	-----------------------------	--	-------------------------------

Fig. VI.34 - Struttura del campo ACF.

L'accesso con arbitraggio a coda è regolato dal *protocollo a coda distribuita* ed è orientato alle comunicazioni di dati, mentre l'accesso pre-arbitrato è usato per le comunicazioni isocrone e riproduce uno schema di trasferimento dell'informazione analogo a quello a circuito.

Nel seguito saranno analizzate entrambe le procedure di accesso.

Il protocollo a coda distribuita è ad accesso controllato. Esso ha lo scopo di riprodurre, in modo distribuito, il funzionamento di un sistema a coda con disciplina primo arrivato-primo servito. In altre parole, il protocollo, tramite un opportuno scambio di informazioni, assegna il diritto di emissione alle stazioni rispettando l'ordine temporale di arrivo dei segmenti informativi a queste.

Il protocollo è basato sull'utilizzazione di due bit del campo ACF: il *bit di occupato B* (busy) che indica se l'IT è attualmente occupato e il *bit di richiesta R* (request) che indica che un segmento informativo è stato inserito nella coda distribuita.

Ogni nodo, contando il numero di richieste che riceve ed il numero di IT vuoti che passano, è in grado di determinare, istante per istante, il numero totale dei segmenti accodati nelle stazioni a valle su ogni bus. Disponendo di tale informazione, una stazione, quando deve trasmettere un segmento, ha la possibilità di determinare immediatamente la sua posizione nella coda distribuita. Se non ci sono segmenti in coda l'accesso può essere immediato, altrimenti occorrerà attendere che tutti i segmenti in coda siano trasmessi.

È il caso di sottolineare che, rispetto ai protocolli MAC sino ad ora descritti, nel protocollo a coda distribuita, le stazioni memorizzano le informazioni relative allo stato della rete che sono utilizzate al momento dell'accesso. Al contrario, nei protocolli precedenti, tutta l'informazione relativa allo stato della rete era contenuta nella rete stessa ed i terminali avevano il compito di attendere il verificarsi di opportuni eventi (canale libero, ricezione

del testimone, ecc.).

La Fig. VI.35 descrive in dettaglio l'algoritmo di base del protocollo a coda distribuita per l'accesso ad uno dei due bus, ad esempio il bus A. Un algoritmo identico presiede l'accesso al bus B. Nel caso in esame, l'algoritmo si basa sull'elaborazione del bit B, trasportato dagli IT in transito sul bus A, e del bit R, trasportato in senso opposto dagli IT in transito sul bus B.

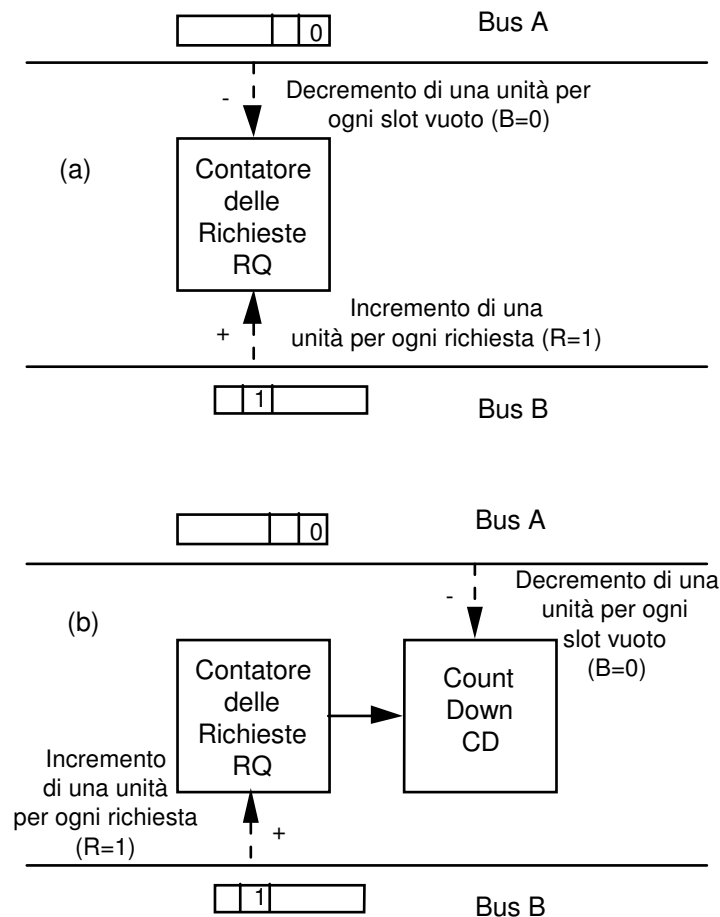


Fig. VI.35 - Algoritmo di coda distribuita: a) aggiornamento del contatore RQ; b) aggiornamento del contatore CD.

Ogni stazione tiene memoria del numero di segmenti collocati nella coda distribuita da parte delle stazioni a valle sul bus A, aggiornando il proprio *contatore RQ* (Request Counter). Tale contatore viene incrementato ogniqualvolta la stazione rivela, sul bus B, il transito di un bit R=1 e viene decrementato ogniqualvolta, sul bus A, transita un IT vuoto (bit B=0).

Quando una stazione deve emettere un segmento informativo, avverte le stazioni a monte sul bus A ponendo, alla prima opportunità, un bit R=1 su di un IT del bus B. Contemporaneamente salva il contenuto corrente del contatore RQ nel *contatore CD* (Count Down). Con tale operazione il segmento è collocato

nella coda distribuita. Una stazione può immettere nella coda distribuita solo un segmento alla volta; il segmento successivo può essere inserito solo quando la emissione del precedente è stata completata.

Il valore del contatore CD indica la posizione del segmento nella coda distribuita. La stazione guadagnerà quindi l'accesso al canale solo dopo che le stazioni a valle avranno esaurito le emissioni precedentemente richieste. Il contatore CD viene decrementato nello stesso modo del contatore RQ, in tal modo la stazione ha il diritto alla emissione nel primo IT vuoto successivo a quello che causa l'azzeramento del contatore CD.

Ovviamente durante il tempo in cui una stazione è in attesa di una emissione, deve comunque continuare l'aggiornamento del contatore RQ.

È evidente che queste operazioni garantiscono normalmente la emissione dei segmenti rispettando l'ordine di presentazione delle richieste.

Un esempio della procedura di accodamento del protocollo d'accesso a coda distribuita è mostrato nella Fig. VI.36 nel caso di rete composta da cinque stazioni. Inizialmente supponiamo che il valore di tutti i contatori RQ sia posto a zero e tutti gli IT che passano sul bus A siano occupati in modo da non consentire emissioni e comportare decrementi dei contatori RQ. Supponiamo inoltre che le stazioni 5, 2 e 3 presentino nell'ordine una richiesta. Al termine delle operazioni, il contatore RQ della stazione 1 avrà valore 3 avendo visto passare tre richieste sul bus B. Analogamente il valore del contatore RQ della stazione 4 avrà valore 1, poiché l'unica stazione a valle che ha effettuato una richiesta è stata la stazione 5. La stazione 5 avrà i contatori CD e RQ a 0 poiché è stata la prima ad effettuare la richiesta e non ha nessun'altra stazione a valle. La stazione 3 avrà il contatore CD ad 1 ed il contatore RQ=0. Il primo indica che la stazione 3 dovrà attendere la emissione di una stazione posta a valle, e cioè la stazione 5, prima di poter effettuare la propria. Il secondo indica che, dopo la richiesta della stazione 3 non è stata presentata nessun'altra richiesta da parte delle stazioni a valle. Per ragioni analoghe la stazione 2 è caratterizzata da CD=1 e RQ=1.

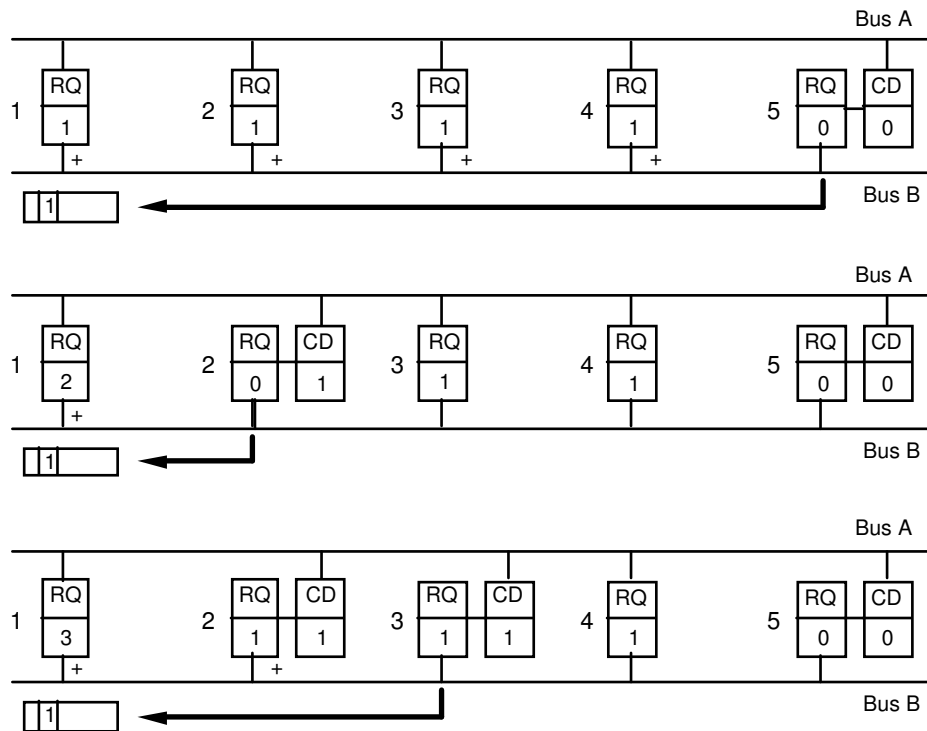


Fig. VI.36 - Inserimento dei segmenti nella coda distribuita.

Un esempio della procedura di trasmissione del protocollo d'accesso a coda distribuita a partire dallo stato ora è mostrato nella Fig. VI.37. Supponiamo che nessuna altra richiesta venga presentata dalle stazioni e che gli IT che transitano sul bus A siano tutti vuoti. In questo caso, il primo IT vuoto verrà utilizzato dalla stazione 5 che eliminerà quindi il contatore CD e sarà disponibile, eventualmente, a richiedere un'altra emissione. Al transito dello slot vuoto le altre stazioni hanno provveduto a decrementare i propri contatori (CD e RQ). Il secondo IT vuoto verrà utilizzato dalla stazione 2 perché è la prima stazione attraversata che ha il contatore CD posto a 0. Il terzo IT vuoto sarà utilizzato infine dalla stazione 3. A questo punto lo stato della rete è tornato quello iniziale con tutti i contatori posti a 0.

È il caso di sottolineare che, a causa dei tempi di propagazione finiti, può accadere che l'ordine di presentazione delle richieste non corrisponda esattamente a quello di emissione. Ciò può accadere se l'intervallo di tempo tra le due richieste è inferiore a quello di propagazione tra le due stazioni. Conseguenza del fenomeno precedente e che non è più rispettato il principio di equità nell'assegnazione del diritto d'accesso alla rete. Infatti, specialmente se la rete è molto estesa, alcune stazioni, quelle più vicine alla stazione di testa, possono essere avvantaggiate nel guadagnare il diritto di emissione.

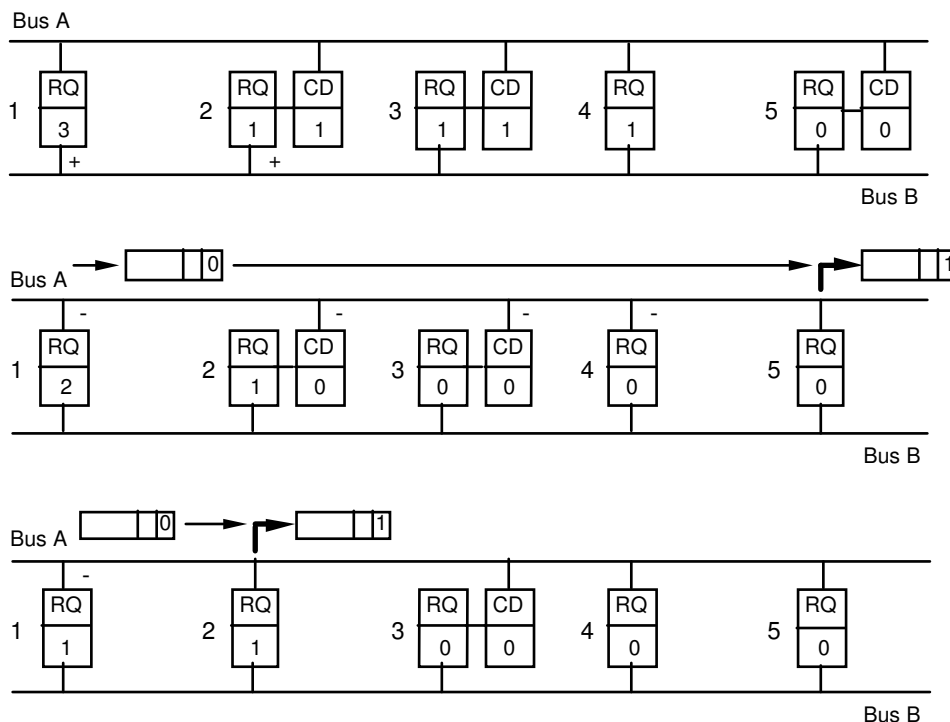


Fig. VI.37 - Trasmissione dei segmenti in coda distribuita.

Per ovviare a ciò, lo standard IEEE 802.6 prevede l'applicazione di un ulteriore meccanismo, detto di *bilanciamento di banda* (bandwidth balancing). Questo modifica parzialmente le modalità di utilizzazione degli IT vuoti, costringendo una stazione a non utilizzarne alcuni, anche se ha segmenti informativi in coda.

Lo standard definisce un parametro di sistema, detto *modulo di bilanciamento di banda* (bandwidth balancing module - BWB\_MOD), ed un *contatore di emissioni* (bandwidth balancing counter). Tale contatore ha modulo uguale BWB\_MOD ed è incrementato quando la stazione effettua una emissione. Dopo che la stazione ha effettuato BWB\_MOD emissioni consecutive, il contatore è azzerato, ciò indica che l'IT vuoto successivo, che dovrebbe essere utilizzato per la trasmissione, deve essere lasciato transitare inalterato.

In questo modo alcuni IT vuoti sono lasciati a disposizione delle stazioni a valle. È possibile dimostrare che ciò è sufficiente ad eliminare qualsiasi fenomeno di diseguità nell'accesso al canale trasmissivo.

La seconda modalità d'accesso prevista nello standard IEEE 802.6 è quella relativa agli IT PA. Tale tecnica di accesso è riservata alle comunicazioni isocrone ed è realizzata con modalità analoghe al trasferimento a circuito.

Quando una stazione deve eseguire una comunicazione isocrona, la

stazione di testa del bus provvederà ad emettere, al ritmo sufficiente per il servizio previsto, gli IT del tipo PA. Tali IT potranno essere utilizzati solo dalla stazione che ne ha fatto richiesta. Il riconoscimento di tali IT avviene mediante la lettura di un apposito campo di indirizzo, denominato *identificatore di canale virtuale* (Virtual Channel Identifier - VCI). Questo identifica univocamente una chiamata.

## **VII Internet**



## VIII La rete ISDN

### VIII.1 Architettura di una ISDN

#### VIII.1.1 Interfacce di accesso

In Fig. VIII.1 sono mostrati alcuni esempi di interfacce utente-rete in una ISDN. Si considerano gli accessi di:

- 1) un singolo apparecchio terminale ISDN (nel seguito indicato come *terminale ISDN*) che si suppone abbia potenzialità multiservizio;
- 2) un insieme di terminali ISDN collocati nella stessa postazione di utente con configurazione punto-multipunto;
- 3) reti private, quali LAN o reti basate su PABX;
- 4) centri specializzati per l'immagazzinamento e per l'elaborazione dell'informazione;
- 5) altri tipi di reti, quali quelle dedicate ad un servizio o altre ISDN.

Le interfacce utente-rete dovrebbero consentire:

- l'uso della stessa interfaccia da parte di differenti tipi di terminali;
- la portabilità dei terminali da un sito ad un altro in ambito sia nazionale che internazionale;
- l'evoluzione separata dei terminali e delle apparecchiature di rete, in termini sia di tecnologie impiegate che di scelte sistemistiche adottate;
- la connessione efficiente con i centri specializzati di cui in 4) o con le altre reti di cui in 5).

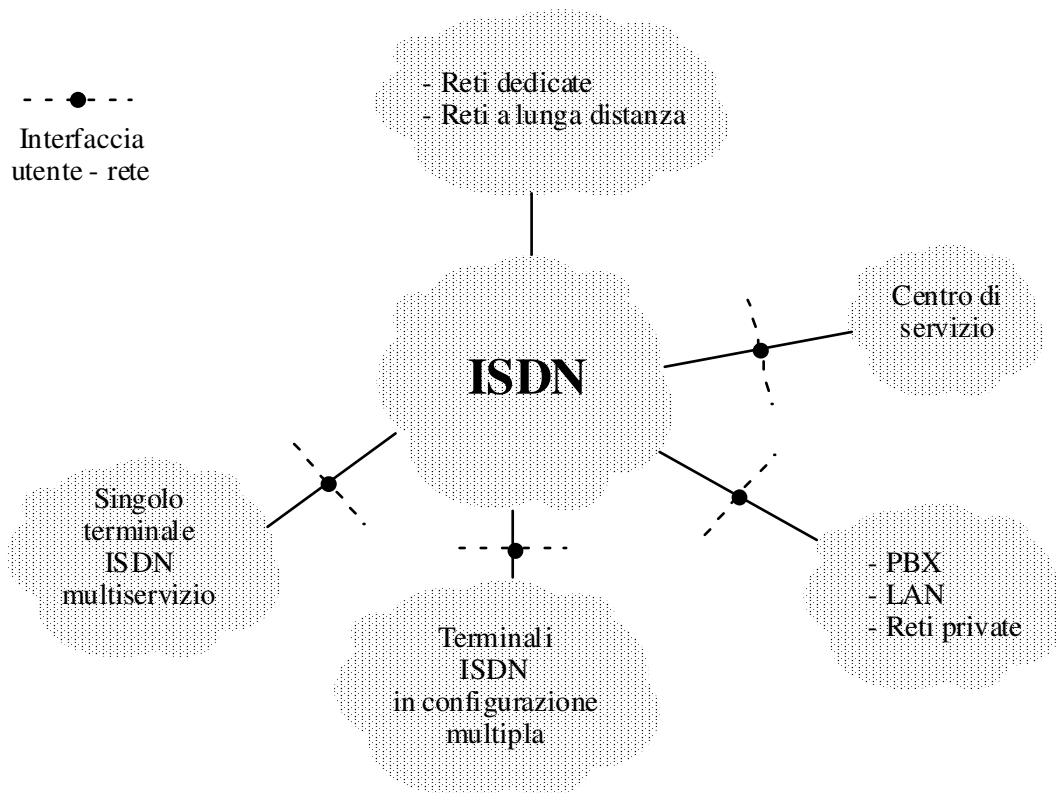


Fig. VIII.1 - Interfacce di accesso a una ISDN.

Nell'architettura funzionale dell'accesso utente-rete possono in generale essere distinti due elementi fondamentali, e cioè i *gruppi funzionali* e i *punti di riferimento*, che sono utilizzati per definire le *configurazioni di riferimento* e le *interfacce di accesso*.

I gruppi funzionali sono insiemi di funzioni che possono essere necessarie in un accesso a una ISDN; il loro significato è essenzialmente logico. Ad essi possono quindi non corrispondere direttamente apparecchiature fisiche di rete. Inoltre, in particolari configurazioni di accesso, possono essere presenti o meno.

Un punto di riferimento rappresenta la separazione concettuale tra due gruppi funzionali e, quindi, il punto di visibilità logica delle funzioni espletate dai gruppi funzionali che esso separa. In particolari configurazioni di accesso, tali punti possono essere interni a specifici gruppi funzionali e quindi possono non corrispondere a un reale punto di accesso fisico; in altre configurazioni esiste invece corrispondenza tra punti di riferimento e interfacce fisiche.

Le interfacce costituiscono l'insieme delle caratteristiche fisiche e delle procedure logiche che debbono essere garantite nei punti di riferimento fisicamente individuati.

Le configurazioni di riferimento sono strutture concettuali, comprendenti gruppi funzionali e punti di riferimento. Esse sono utilizzate per identificare le

varie possibilità con cui un utente può accedere a una ISDN.

La struttura delle interfacce di accesso a una ISDN è specificata dalle raccomandazioni dell'ITU-T nella serie I.400. In queste sono descritte i gruppi funzionali e le modalità di interazione tra di essi e con altre entità funzionali descritte in altre raccomandazioni. È opportuno precisare che, in tutti questi documenti di normalizzazione, sono tracciate solo le linee entro cui incanalare lo sviluppo delle unità funzionali ivi considerate, lasciando, alla completa discrezionalità dei gestori e dei costruttori, la definizione delle relative modalità realizzative.

Per comprendere le scelte che verranno illustrate nel seguito di questo paragrafo, occorre tenere presente che, come già sottolineato, ad ogni chiamata corrispondono due flussi informativi distinti, quelli dell'informazione di utente e dell'informazione di segnalazione. Le caratteristiche del secondo di questi flussi e le relative procedure per il controllo della chiamata devono essere tali da consentire di operare, in modo efficace, nel trattamento di vari tipi di chiamata, anche se i requisiti prestazionali dei relativi servizi di telecomunicazione sono molto diversi.

Di conseguenza occorre garantire la completa indipendenza dei due flussi. Ciò implica che le informazioni di utente e quelle di segnalazione debbono essere trasferite su canali separati sia nella sezione di accesso che in quella interna. Per quest'ultima provvede la rete di segnalazione a canale comune. Nella sezione di accesso questo obiettivo è conseguito affidando a uno specifico canale il compito di trasferire l'informazione di segnalazione attraverso l'interfaccia utente-rete.

La trattazione seguente affronta dapprima la definizione degli elementi strutturali della sezione di accesso di una ISDN; successivamente vengono fornite le caratteristiche dei tipi di canali di accesso che sono stati normalizzati per una N-ISDN; infine si conclude con la descrizione delle strutture di interfaccia, risultanti dalla composizione, in un opportuno schema di multiplazione, dei detti canali di accesso.

In Fig. VIII.2 sono mostrate le configurazioni di riferimento. In queste si distinguono cinque tipi di gruppi funzionali e tre tipi di punti di riferimento.

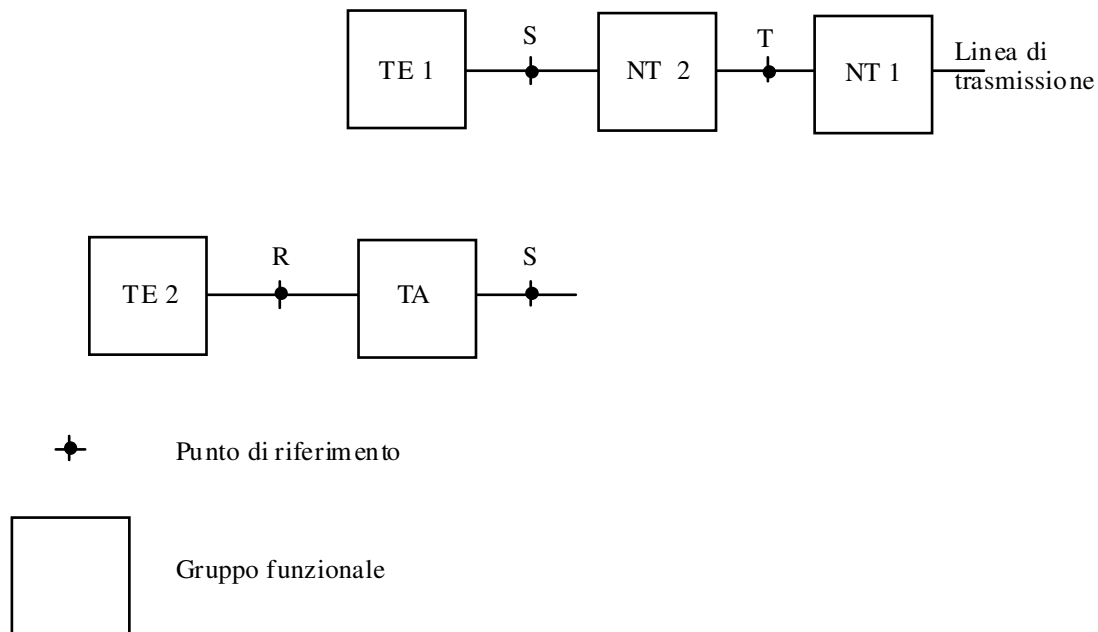


Fig. VIII.2 - Configurazioni di riferimento per l'interfaccia utente-rete in una ISDN.

Tra i gruppi funzionali, vengono distinte due terminazioni di rete, in luogo di una singola unità di interfaccia. Ciò si spiega con l'esigenza di tenere conto delle diversità presenti, a livello internazionale, nella regolamentazione della fornitura dei servizi di telecomunicazione. In particolare, nei Paesi in cui tale fornitura avviene in regime di concorrenza tra varie Compagnie di esercizio (ad esempio, in Nord America), è necessario distinguere due sezioni della terminazione di rete: una di proprietà del gestore, l'altra del cliente. Tali due sezioni sono indicate con i numeri 1 e 2, rispettivamente.

La *Terminazione di rete 1* (NT1) è un gruppo funzionale che racchiude funzioni orientativamente appartenenti a quelle dello strato 1 del modello OSI. Queste funzioni hanno come obiettivo una adatta terminazione della rete da un punto di vista fisico.

La *Terminazione di rete 2* (NT2) è un gruppo funzionale che comprende funzioni orientativamente appartenenti allo strato 1 e, talvolta, anche agli strati 2 e 3 del modello OSI. Queste funzioni hanno lo scopo di controllare l'accesso a una particolare configurazione degli apparecchi terminali nella postazione di utente, quando la molteplicità di questi è maggiore dell'unità. Controllori di terminali, LAN e PABX sono esempi di apparecchiature o di combinazioni di apparecchiature che svolgono funzioni NT2.

L'*Apparecchio terminale* (TE) è un gruppo funzionale che comprende funzioni, orientativamente e in generale, appartenenti a tutti i sette strati del modello OSI. Esempi di apparecchiature o di combinazioni di apparecchiature

che forniscono queste funzioni sono offerti da apparecchi telefonici numerici, da DTE e da stazioni di lavoro con potenzialità multiservizio.

Vengono distinti due tipi di apparecchi terminali: l'*apparecchio terminale 1* (TE1) e l'*apparecchio terminale 2* (TE2). Il TE1 ha un'interfaccia che risponde alla normativa ISDN ed è quindi ISDN-compatibile, mentre il T2 non presenta analoghe condizioni di compatibilità. Esempi di TE2 sono i DTE specificati nelle raccomandazioni ITU-T delle serie V ed X.

Per la connessione di un TE2 a una ISDN è allora necessario un ulteriore gruppo funzionale, chiamato *Adattatore di terminale* (TA). Questo include funzioni orientativamente appartenenti a quegli strati del modello OSI che permettono a un TE di essere servito da una interfaccia di accesso ISDN.

Sempre con riferimento alla Fig. VIII.2, precisiamo che:

- il *punto di riferimento T* è la delimitazione tra i gruppi funzionali NT2 ed NT1;
- il *punto di riferimento S* è la delimitazione tra i gruppi TE1 e NT2 o tra i gruppi TA e NT2;
- il *punto di riferimento R* è la delimitazione tra i gruppi funzionali TE2 e TA.

In Fig. VIII.3 sono mostrati alcuni esempi di configurazioni di accesso, che possono essere ipotizzate, utilizzando quelle di riferimento presentate nella Fig. VIII.3 e combinando in modi opportuni i differenti gruppi funzionali precedentemente descritti. In particolare, la Fig. VIII.3a è quella di riferimento, in cui le interfacce fisiche di accesso a una ISDN sono localizzate nei punti di riferimento S e T. Altri due esempi riguardano i casi in cui l'interfaccia fisica è nel solo punto di riferimento S (Fig. VIII.3b) o in quello T (Fig. VIII.3c). Infine la Fig. VIII.3d riguarda una configurazione in cui l'interfaccia fisica di accesso è individuata da entrambi i punti di riferimento S e T, che coincidono.

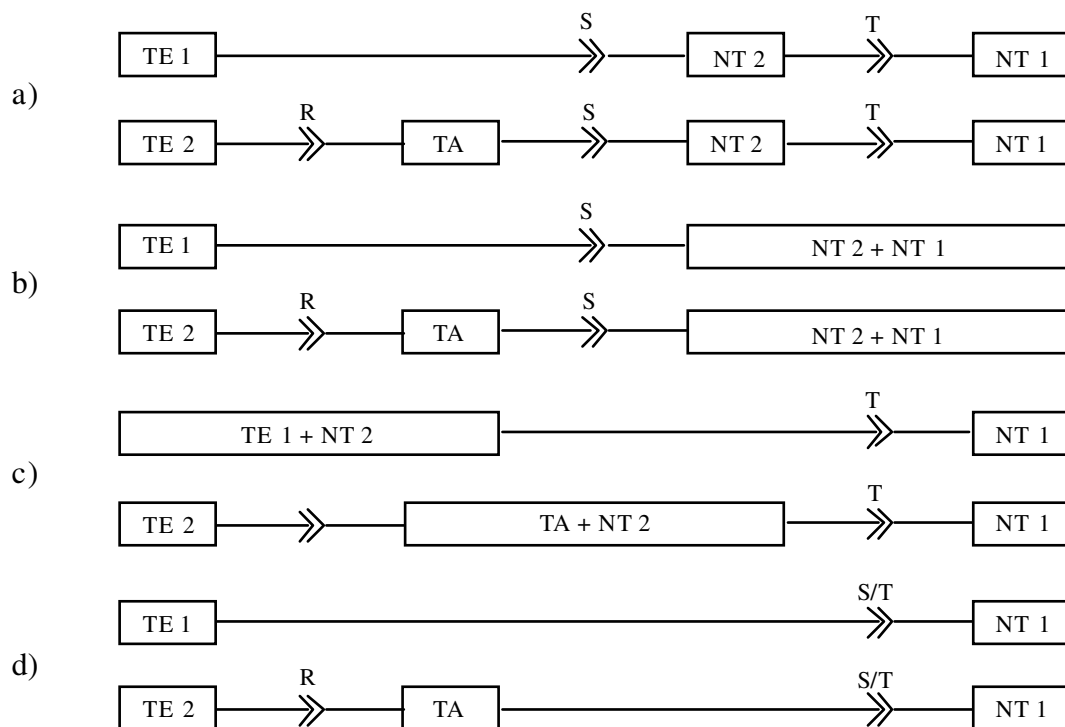


Fig. VIII.3 - Esempi di configurazioni di accesso.

Sono qui di seguito descritti i tre tipi di canale di accesso, le cui caratteristiche sono trattate nella Racc. I.412. Essi sono il *canale B*, il *canale D* e i *canali H*.

Il canale B ha capacità di trasferimento uguale a 64 kbit/s con relativa temporizzazione. Su di esso possono essere trasferiti vari tipi di informazioni d'utente. Non viene invece trasportata informazione di segnalazione per il controllo delle connessioni commutate a circuito entro la ISDN. I flussi di informazione d'utente, che possono essere convogliati su un canale B, hanno ritmi binari minori o uguali a 64 kbit/s e possono essere orientativamente dei seguenti tipi:

- voce di qualità telefonica codificata a 64 kbit/s, secondo la Racc. G.711;
- informazione di dati, corrispondenti alle classi di servizio della Racc. X.1 per trattamento a circuito o a pacchetto;
- voce di qualità migliorata codificata a 64 kbit/s, secondo la Racc. G.722;
- voce codificata a sotto-ritmo, sia da sola, che in unione con altri flussi di informazione numerica.

Per un canale B è possibile selezionare, in modo alternativo, uno dei quattro impieghi ora precisati: ciò può avvenire sia all'atto dell'instaurazione della chiamata, sia con modifica della scelta durante una chiamata già instaurata.

Il canale B può essere utilizzato come accesso per una varietà di modi di trasferimento con servizio orientato alla connessione. Le connessioni possono essere:

- commutate a circuito;
- commutate a pacchetto per DTE capaci di pacchettizzare l'informazione (PM-DTE - Packet Mode DTE);
- connessioni semi-permanenti con nodi a circuito o a pacchetto.

Nel caso in cui il canale B venga utilizzato per convogliare un singolo flusso con ritmo binario inferiore a 64 kbit/s, occorre adattare questo flusso alla capacità di trasferimento del canale, secondo quanto indicato nella Racc. I.460. Infine, quando flussi multipli sono multiplati staticamente su un canale B nell'ambito di una connessione commutata a circuito, la destinazione di questi flussi deve essere la stessa.

Il canale D può avere, in relazione alla struttura di interfaccia, capacità di trasferimento uguali a 16 kbit/s o a 64 kbit/s. Il suo scopo principale è trasportare l'informazione di segnalazione preposta al controllo delle connessioni commutate a circuito. In aggiunta all'informazione di segnalazione un canale D può anche essere usato come supporto di informazione di telemetria e di dati trattati a pacchetto, avente origini da sorgenti con alto grado di intermittenza.

I canali H hanno capacità di trasferimento varie, tutte accompagnate dalla relativa temporizzazione. Sono stati normalizzati, per ora, due tipi di canali H: il canale  $H_0$  con capacità di 384 kbit/s (equivalente a 6 x 64 kbit/s) e il canale  $H_1$  che può avere due varianti; quella corrispondente alla gerarchia numerica europea ( $H_{12}$ ) ha capacità di 1920 kbit/s (30 x 64 kbit/s), mentre quella relativa alla gerarchia numerica nord-americana ( $H_{11}$ ) ha capacità di 1536 kbit/s (24 x 64 kbit/s).

Anche i canali H sono riservati al trasporto dell'informazione d'utente, con modalità di impiego analoghe a quelle del canale B. I tipi di flussi informativi, che possono interessare i canali H sono, ad esempio:

- facsimile veloce;
- video per teleconferenza;
- dati ad alta velocità;
- materiale da programmi audio e musicali ad alta fedeltà;
- informazioni numeriche con ritmo binario inferiore alla capacità di trasferimento di un canale H e opportunamente multiplate tra di loro;
- informazione trattata a pacchetto.

In Tab. VIII.1 sono riassunte le capacità di trasferimento dei canali B,  $H_0$  e

H<sub>1</sub> per i due tipi di gerarchia numerica; sono anche indicate le capacità dei canali H<sub>2</sub> e H<sub>4</sub>, che sono attualmente in corso di normalizzazione.

Canale	Gerarchia numerica basata sul ritmo primario di	
	2.048 kbit/s	1.544 kbit/s
B	64	
H <sub>0</sub>	384	
H <sub>1</sub>	1920 (30x64)	1472 (23x64)
H <sub>2</sub>	32.768	43.000 ÷ 45.000
H <sub>4</sub>	132.032 ÷ 138.240	

Tab. VIII.1 - Capacità di trasferimento dei canali B e H (in kbit/s)

Le interfacce fisiche utente-rette nella sezione di accesso di una ISDN in corrispondenza dei punti di riferimento S e T debbono essere conformi con una delle strutture di interfaccia, che sono descritte nelle Tab. VIII.2 e Tab. VIII.3 e che si distinguono in *strutture di interfaccia con il canale B* e in quelle *con i canali H*.

In tutte queste strutture, la banda disponibile è utilizzata con una moltiplicazione di tipo ibrido, in cui la trama (depurata della parte dedicata al trasferimento di informazioni aggiuntive rispetto a quelle di utente e di segnalazione) è logicamente partizionata in due regioni: una suddivisa in IT e l'altra indivisa. La frontiera tra le due parti della trama è fissa. La regione suddivisa in IT è la sede dei canali B o H; quella indivisa è invece a disposizione del canale D.

Pertanto, nell'ambito dell'interfaccia utente-rette, si ha una *moltiplicazione di strato 1*, che, in presenza di un solo apparecchio terminale nella postazione di utente, è gestita dai due gruppi funzionali TE e NT1.

Tipo	Composizione	Capacità del canale D (kbit/s)	Ritmo binario netto di interfaccia (kbit/s)
Struttura di base	2B+D	16	144
Struttura a ritmo primario (europea)	30B+D	64	2.048
Struttura a ritmo primario (USA)	23B+D	64	1.544

Tab. VIII.2 - Strutture di interfaccia con il canale B

Tipo	Composizione	Capacità del canale D (kbit/s)	Ritmo binario netto di interfaccia (kbit/s)
Struttura a ritmo binario con i canali $H_0$	$5 H_0 + D$	64	2.048
	$3 H_0 + D$	64	1.544
	$4 H_0$	-	
Struttura a ritmo binario con i canali $H_1$	$H_{12} + D$	64	2.048
	$H_{11}$	-	1.544

Tab. VIII.3 - Strutture di interfaccia con i canali H

Per ciò che riguarda poi la strategia di assegnazione della banda disponibile quando nella postazione di utente sono presenti due o più apparecchi terminali, i canali B e H sono pre-assegnati individualmente su base chiamata all'apparecchio terminale che ne fa richiesta con l'apposita procedura di segnalazione di utente. Invece il canale D è, a seconda dei casi assegnato a domanda o pre-assegnato collettivamente.

Per l'accesso al canale D si presenta allora una situazione di accesso multiplo, che deve essere risolta con un opportuno controllo di accesso al mezzo: la gestione del relativo protocollo MAC è ancora compito del gruppo funzionale NT1.

Come appare in Tab. VIII.2, alle strutture d'interfaccia con il canale B appartengono la *struttura di base* e quella *a ritmo primario*. La struttura di base è composta di due canali B e di un canale D, quest'ultimo con capacità di

trasferimento di 16 kbit/s. I canali B possono essere utilizzati in modo indipendente e cioè in differenti connessioni contemporanee. Inoltre, mentre la presenza del canale D è in ogni caso essenziale, uno o entrambi tra i canali B possono non essere attivati in installazioni specifiche.

La struttura a ritmo binario con i canali B comprende (almeno nella versione corrispondente alla gerarchia numerica europea) 30 canali B e un canale D, quest'ultimo con capacità di trasferimento di 64 kbit/s. Circa l'attivazione dei canali previsti in questa interfaccia, vale quanto detto riguardo alla struttura di base: cioè, uno o più tra i canali B può non essere attivo.

Le strutture di interfaccia con i canali H (Tab. VIII.3) sono tutte a ritmo primario, almeno per ciò che riguarda quelle finora normalizzate. Tra queste ultime citiamo la struttura con canali  $H_0$ , che è costituita, sempre con riferimento alla gerarchia numerica europea, da 5 canali  $H_0$  e da un canale D con capacità di trasferimento di 64 kbit/s. Come si può osservare dalla Tab. VIII.3, il canale D può mancare in alcune di queste strutture; quando ciò si verifica si utilizza il canale D presente in un'altra interfaccia.

#### VIII.1.2 Funzioni di trasporto

La descrizione delle possibilità offerte da una ISDN a banda stretta per lo scambio di informazione, al proprio interno o tra due suoi punti di accesso, richiede l'identificazione di opportuni *gruppi funzionali* e delle loro modalità di cooperazione ai fini della fornitura di uno o più tra i servizi ISDN.

Tali gruppi funzionali sono preposti a svolgere:

- funzioni riguardanti l'instaurazione e il controllo di connessioni al livello locale; queste funzioni (LFC - Local Functional Capabilities) sono fornite principalmente dagli autocommutatori locali e risiedono nelle terminazioni di commutatore, nelle entità di commutazione, nonché in quelle di controllo e di gestione di rete;
- funzioni di trasferimento a circuito, di tipo sia commutato che semi-permanente, con capacità di base uguale a 64 kbit/s, ma con possibilità di trasferimento anche per ritmi inferiori o superiori a 64 kbit/s;
- funzioni di trasferimento a pacchetto;
- funzioni di trasferimento della segnalazione inter-nodale a canale comune, per esempio conformi all'SS no.7, almeno per ciò che riguarda le connessioni in ambito internazionale.

Tutti questi gruppi funzionali corrispondono a potenzialità di strato basso, sia fondamentali che addizionali. Sono anche disponibili funzionalità di strato alto (fondamentali e addizionali), a cui si può accedere tramite l'ISDN per la

fornitura di particolari tele-servizi.

L'interazione dei gruppi funzionali sopra citati, tra di loro o con le entità presenti nella sezione di accesso, avviene con l'utilizzazione di protocolli di comunicazione.

All'approfondimento di queste tematiche è dedicato il resto di questo paragrafo che tratta l'architettura generale di una ISDN e il relativo modello di riferimento dei protocolli.

In Fig. VIII.4 è mostrato il modello architetturale di base per una N-ISDN. In questo sono comprese potenzialità di trasferimento sia a circuito che a pacchetto, oltre alle funzionalità preposte alla segnalazione.

Riguardo a queste ultime, si è già sottolineato come una delle scelte architetturali fondamentali in una ISDN sia il trasferimento delle informazioni di utente e di segnalazione su canali separati. In Fig. VIII.4 si distinguono tre possibilità di segnalazione:

- *tra l'utente e la ISDN*, per la richiesta/offerta dei servizi di supporto e dei tele-servizi, con i relativi servizi supplementari;
- *all'interno della ISDN*, per, ad esempio, instradare le chiamate;
- *da utente a utente* per la fornitura dell'omonimo servizio supplementare.

Sempre in Fig. VIII.4 viene rimarcata la disponibilità di interfacce utente-rete di tipo normalizzato, che sono individuate dai punti di riferimento S e/o T e che hanno caratteristiche di accesso universali per l'intera gamma di potenzialità offerte dall'ISDN, nell'ambito di una relazione di comunicazione tra due terminali ISDN.

Le funzioni di strato alto, necessarie per la fornitura dei tele-servizi, sono disponibili nei terminali ISDN, all'interno della rete e in centri di servizio, che sono connessi all'ISDN con un'interfaccia S/T ovvero sono presenti all'interno di una rete dedicata e accessibili tramite l'ISDN. Tali funzioni possono essere attivate nei soli terminali, quando questi utilizzino tele-servizi che non richiedano l'impiego di ulteriori funzioni di strato alto. Possono però essere richieste anche alla ISDN e ai centri di servizio, quando è necessaria la fornitura di prestazioni a valore aggiunto, quali la conversione di protocolli e l'archiviazione di informazione.

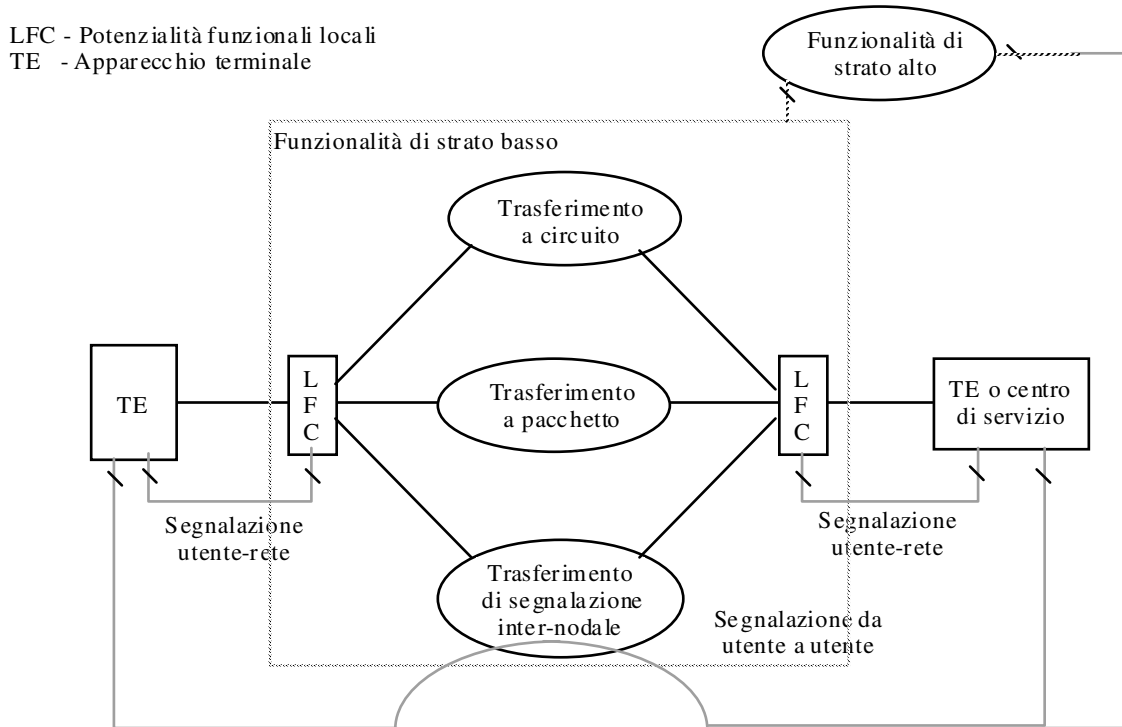


Fig. VIII.4 - Modello architetturale di base di una ISDN.

La Fig. VIII.5 fornisce una rappresentazione più dettagliata del modello di Fig. VIII.4, mettendo in evidenza, oltre ai gruppi funzionali TE1, TE2, TA, NT2, NT1 e ai punti di riferimento R, S, T, anche ulteriori gruppi funzionali delle sezioni di accesso e interna di una ISDN. In particolare, sono ivi considerate le funzionalità di tipo trasmissivo e quelle preposte all'accesso dell'autocommutatore locale.

Le funzionalità di tipo trasmissivo riguardano la utilizzazione della linea d'utente numerica e prevedono due entità terminali: quella dal lato della postazione d'utente è il gruppo funzionale NT1, mentre quella dal lato dell'autocommutatore locale è la *Terminazione di linea* (LT - Line Termination). Oltre a ricoprire il ruolo primario di terminali trasmittente e ricevente alle estremità della linea di utente, le entità NT1 e LT svolgono funzioni aggiuntive, quali il monitoraggio della qualità trasmissiva, la gestione della moltiplicazione di strato 1, l'alimentazione degli apparati di trasmissione e la protezione della linea. Tra queste due entità è prevista l'*interfaccia U*.

Il gruppo funzionale *Terminazione di commutatore* (ET - Exchange Termination) ha il compito di svolgere, dal lato rete, le funzioni appartenenti agli strati 2 e 3, che il gruppo funzionale NT2 svolge dal lato postazione di utente. La delimitazione tra le entità ET e LT è il *punto di riferimento V*. Quest'ultimo usualmente coincide con l'*interfaccia omonima*, che appartiene

alla sezione interna della ISDN e che ha realizzazioni diverse a seconda delle strutture di interfaccia utente-rete.

Infine, sempre in Fig. VIII.5, viene messo in evidenza che la ISDN fornisce accesso alle esistenti reti dedicate in servizio pubblico e, attraverso queste, agli utenti e ai centri di servizio che ad esse fanno capo. Tali strutture dedicate sono, ad esempio, la rete telefonica analogica e la rete per dati a pacchetto.

VIII.1.3 Potenzialità di trasferimento

Per ciò che riguarda poi le potenzialità di trasferimento offerte da una ISDN, ci riferiamo dapprima al *trasferimento a circuito*. Questo può avvenire, su connessioni commutate o semi-permanenti, con ritmi di trasferimento fino a 64 kbit/s. All'interfaccia utente-rete, ciascuna di tali connessioni è supportata da un canale B, mentre, nella sezione interna, lo è da entità funzionali operanti a commutazione di circuito.

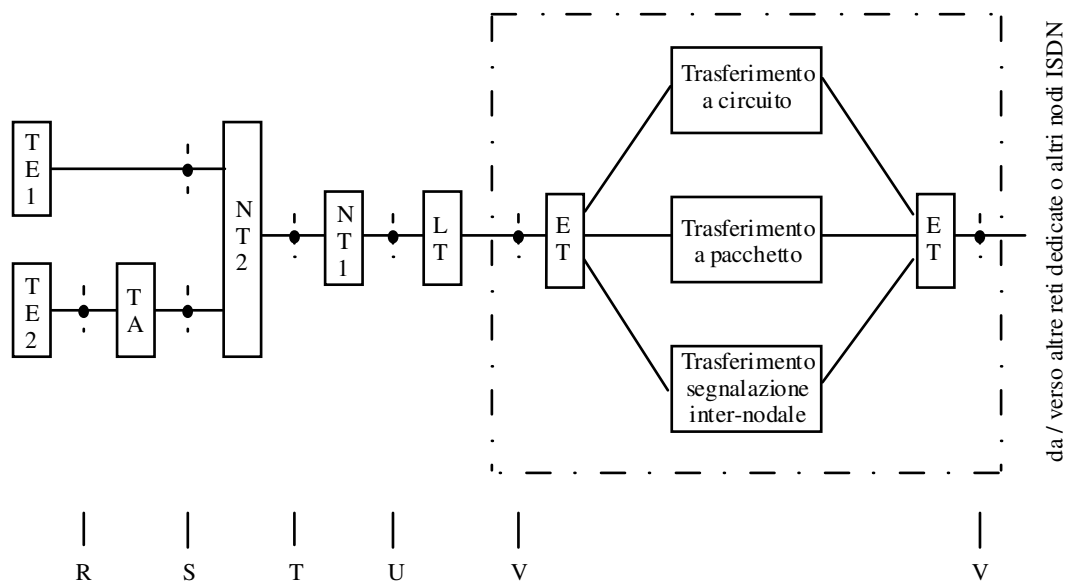


Fig. VIII.5 - Gruppi funzionali e punti di riferimento per l'accesso a una ISDN.

La segnalazione scambiata per il controllo di tali connessioni è trasportata dal canale D nella sezione di accesso (segnalazione utente-rete) e dalle funzionalità della rete di segnalazione a canale comune nella sezione interna (segnalazione inter-nodale).

È anche prevista attualmente la disponibilità di connessioni semi-permanenti con capacità di trasferimento che sono un multiplo di 64 kbit/s; connessioni commutate di questo tipo potranno in futuro essere offerte da entità funzionali di commutazione di circuito a larga banda.

Riguardo poi alla fornitura di connessioni tra apparecchi terminali operanti a ritmo binario inferiore a 64 kbit/s (ad esempio, tra terminali per dati operanti con le classi di servizio previste nella Racc. X.1), queste possono essere ottenute con un adattamento dei relativi ritmi binari. Le modalità di questo adattamento corrispondono a quanto previsto nella Racc. X.30.

Il *trasferimento a pacchetto* tra apparecchi terminali operanti a pacchetto (*PM-TE* - Packet Mode TE) connessi a una ISDN è consentito nell'ambito di varie soluzioni, che possono utilizzare due tipi di gruppi funzionali:

- quelli preposti al *trattamento del pacchetto* (PH - Packet Handling), e cioè a una funzione che sia in grado di trattare in modo completo una chiamata a pacchetto, ad esempio del tipo X.25;
- quelli addetti alla *cooperazione* (IW - Interworking) tra l'ISDN e una rete per dati a pacchetto.

Mediante questi gruppi funzionali, le soluzioni in esame prevedono, per l'accesso, l'utilizzazione del canale B o del canale D, come vettore di trasporto dell'informazione emessa o ricevuta dal PM - TE. Quest'ultimo può essere indifferentemente un TE1 ovvero un PM - DTE con il suo adattatore.

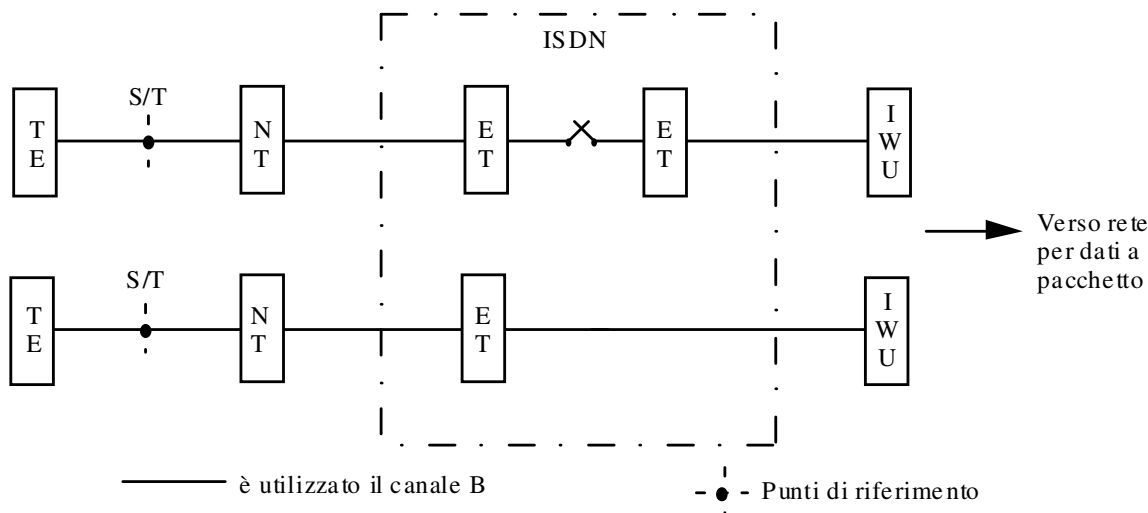
Le *soluzioni attraverso il canale B* possono essere:

- a1) accesso a circuito (commutato o semi-permanente), attraverso l'ISDN, a un gruppo funzionale IW che costituisca porta verso una rete per dati a pacchetto;
- a2) accesso a circuito (commutato o semi-permanente) a un gruppo funzionale PH all'interno della ISDN.

Le soluzioni attraverso il canale D contengono invece i seguenti casi:

- b1) funzioni PH e IW entro la ISDN;
- b2) funzioni PH (senza funzioni IW) entro la ISDN.

Nella soluzione a1), che è descritta in Fig. VIII.6, è utilizzata una connessione a circuito di tipo commutato (rappresentata nella parte alta della figura) o semi-permanente (illustrata nella parte bassa). Per la connessione a circuito, tra un singolo PM - TE e il gruppo funzionale IW di accesso alla rete per dati, viene utilizzato il canale B. Pertanto, nel caso di accesso con connessione di tipo commutato, l'utente deve richiedere preliminarmente, tramite la procedura di segnalazione sul canale D, l'instaurazione della connessione. Successivamente può iniziare la procedura di chiamata prevista negli strati 2 e 3 della Racc. X.25.



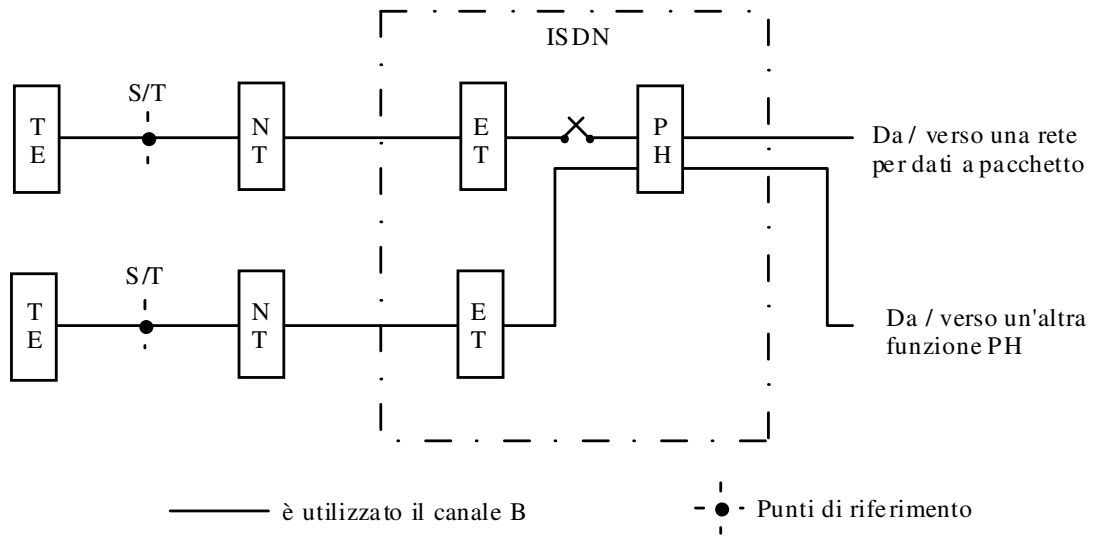
- NT - Terminazione di rete 1 e/o 2
- ET - Terminazione di commutatore
- IWU - Unità di cooperazione tra ISDN e rete per dati a pacchetto
- TE - Apparecchio terminale 1 o PM-DTE con suo adattatore

Fig. VIII.6 - Accesso, tramite il canale B, a una rete per dati a pacchetto.

In Fig. VIII.7 è descritta la soluzione a2). In questo caso viene resa disponibile una connessione fisica tra il terminale e il modulo che effettua la funzione PH. Questo può risiedere indifferentemente nell'autocommutatore locale o in altro sito entro la ISDN. La connessione può essere commutata (parte alta della figura) o semi-permanente (parte bassa della figura). Nel primo caso la instaurazione della connessione richiede la procedura di segnalazione sul canale D. La connessione fisica, una volta resa disponibile, può essere utilizzata per supportare un trasferimento a pacchetto in accordo agli strati 2 e 3 della Racc. X.25.

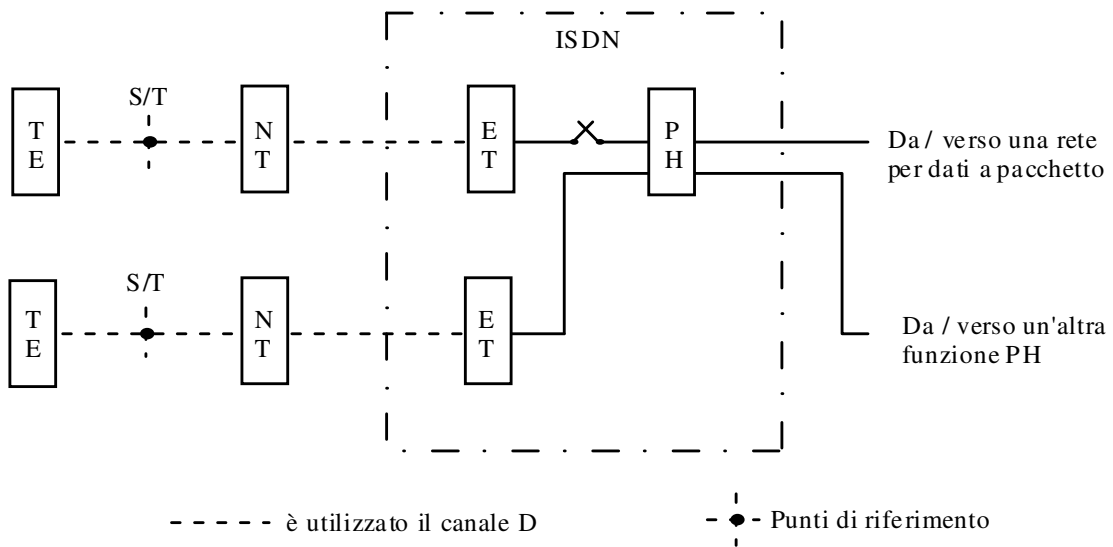
Si noti la differenza fra le due soluzioni rappresentate nelle Fig. VIII.7 e Fig. VIII.8. Nel primo caso il trasferimento dell'informazione (strutturata in pacchetti), attraverso l'ISDN è a circuito e quindi di tipo trasparente; nel secondo caso il trasferimento è invece a pacchetto.

Passando poi alla soluzione b1), che utilizza il canale D (Fig. VIII.8), il PM- TE accede a una funzione PH tramite l'instaurazione di una connessione di strato 2, attuata secondo le modalità del protocollo LAPD (cfr. par. VIII.4). Tale connessione logica può successivamente essere utilizzata per supportare le procedure di chiamata previste nello strato 3 della Racc. X.25.



NT - Terminazione di rete 1 e/o 2  
ET - Terminazione di commutatore  
PH - Funzione di trattamento di pacchetto  
TE - Apparecchio terminale 1 o PM-DTE con suo adattatore

Fig. VIII.7 - Accesso, tramite il canale B, per la fornitura del servizio di circuito virtuale in una ISDN.



NT - Terminazione di rete 1 e/o 2  
ET - Terminazione di commutatore  
PH - Funzione di trattamento di pacchetto  
TE - Apparecchio terminale 1 o PM-DTE con suo adattatore

Fig. VIII.8 - Accesso, tramite il canale D, per la fornitura del servizio di circuito virtuale in una ISDN.

Si può osservare che, in presenza di una molteplicità di terminali nella postazione di utente, l'accesso tramite il canale B, in ambedue le soluzioni di

Fig. VIII.6 e Fig. VIII.7, consente di far operare un solo terminale alla volta. Invece, con l'accesso tramite il canale D, possono operare simultaneamente più terminali. Tale circostanza è da mettere in relazione alla possibilità, offerta da questo canale, di moltiplicare più connessioni di strato 2 su un'unica connessione di strato 1 (cfr. par. VIII.4).

#### VIII.1.4 Il blocco dei protocolli ISDN

Per descrivere, in termini modellistici, gli scambi di informazione verso, attraverso o all'interno di una ISDN è stato definito un *modello di riferimento per i protocolli di una ISDN* (ISDN-PRM - Protocol Reference Model).

Le entità in comunicazione possono essere: 1) utenti della ISDN; 2) un utente della ISDN e una entità funzionale entro una ISDN, come, ad esempio, una risorsa di controllo di rete; 3) un utente della ISDN e una entità funzionale all'interno o all'esterno di una ISDN, come, ad esempio, una risorsa di archiviazione o di elaborazione o di trattamento di messaggi; 4) due o più entità funzionali entro una ISDN, come, ad esempio, una risorsa di gestione di rete e un'altra di commutazione; 5) una entità funzionale della ISDN e una entità appartenente o facente capo a una rete dedicata a un servizio.

Lo scopo delle comunicazioni tra queste entità funzionali è quello di supportare, con potenzialità ISDN, i servizi portanti e i teleservizi. Le potenzialità in questione, nell'ambito di una N-ISDN, sono, ad esempio: a) connessioni commutate a circuito, sotto il controllo della segnalazione a canale comune; b) connessioni commutate a pacchetto; c) segnalazione tra utenti e centri di servizio presenti in rete; d) una combinazione delle potenzialità sopra elencate, come nel caso di comunicazioni multimediali in cui simultaneamente più modi di trasferimento possono aver luogo sotto il controllo della segnalazione a canale comune. Quest'ultimo può essere il caso, ad esempio, della comunicazione di un utente con un centro di servizio; tale comunicazione può utilizzare una connessione a pacchetto per il verso utente-centro di servizio sulla quale viaggiano solo comandi interattivi, e una connessione a circuito per il verso contrario, sulla quale può verificarsi il trasferimento di una grossa mole di dati.

In questo ambito possono aversi vari tipi di flussi informativi quali, ad esempio, quelli legati a: A) processi di controllo di chiamata, gestiti fuori banda; B) scelta delle caratteristiche di una connessione all'atto dell'instaurazione di chiamata e ri-negoziazioni delle medesime durante lo svolgimento di questa; C) sospensioni di connessioni; D) chiamate multimediali; E) comunicazioni bidirezionali asimmetriche; F) funzioni di

gestione di rete, ivi compresi gli aspetti legati all'esercizio e alla manutenzione; G) attivazioni e disattivazioni delle sorgenti di energia di alimentazione degli apparecchi terminali nelle postazioni di utente.

Per un'adeguata descrizione di un ambiente di comunicazione così variegato si è riconosciuta la necessità di un modello di riferimento, che, pur adottando i principi base del modello OSI, ne costituisca una opportuna generalizzazione.

Il principio-base del modello OSI è quello della *stratificazione*, e cioè un principio che è stato dimostrato essere il più idoneo per la definizione dell'architettura di processi di comunicazione. L'ISDN-PRM adotta quindi i concetti di base del modello OSI e la terminologia ad essi associata, con riferimento specifico ai concetti di strato e di servizio di strato, nonché alle nozioni di primitiva di servizio e di protocollo di strato. Tuttavia, accanto all'applicazione di questo principio-base e rispetto al modello OSI, la definizione dell'ISDN-PRM ha richiesto una estensione dei significati di *sistema terminale* e di *sistema di rilegamento*: ciò allo scopo di poter rappresentare elementi specifici della ISDN non previsti nelle attuali reti per dati.

Circa l'identificazione degli strati, nell'ISDN-PRM proposto per una N-ISDN, gli strati utilizzati sono in numero di 7 come nel modello OSI, ma, rispetto a questo, sono identificati con numeri invece che con titoli (ad esempio, "strato 3" invece di "strato di rete"). I titoli potrebbero infatti ingenerare confusione se usati in un contesto ISDN. Per ciò che riguarda poi l'ISDN-PRM per una B-ISDN, una sostanziale modifica del modo di trasferimento adottato rispetto ai modi tradizionali previsti nell'ambito della N-ISDN, ha richiesto una completa ri-definizione degli strati più bassi. Nel seguito, per semplicità e per brevità espositive, ci riferiremo esclusivamente al caso del modello di riferimento per una N-ISDN.

Circa poi l'estensione del significato di sistema terminale o di rilegamento, questa è legata all'esigenza di distinguere, in un ambiente ISDN, tre principali tipi di flussi informativi, e cioè quelli relativi alle informazioni di utente, di controllo e di gestione.

L'informazione di utente (*U-informazione*) è già stata definita (cfr.ad es; [1]). Essa può essere trasferita tra due o più utenti o tra utenti e centri di servizio e, nel trasferimento, può essere trattata dalla rete in modo trasparente oppure può essere elaborata, come accade nei processi di archiviazione e in quelli di codifica crittografica svolti all'interno della rete stessa.

L'informazione di controllo (*C-informazione*) è quella di segnalazione già

più volte considerata in precedenza. È necessaria affinché possa avvenire il trasferimento della U-informazione. Essa è quindi costituita dall'informazione preposta a: 1) controllare una connessione di rete, per instaurarla o per abbatterla; 2) controllare l'uso di una connessione di rete già instaurata, per modificarne, ad esempio, le caratteristiche durante una chiamata; 3) controllare la fornitura di servizi supplementari.

L'informazione di gestione (*M-informazione*) ha sostanzialmente due finalità principali. Da un lato, riguarda tutti gli aspetti di *gestione locale* relativi al trasferimento dei precedenti due tipi di informazione e al coordinamento tra le funzioni che vengono attivate a questo scopo. Dall'altra, ha significatività in un contesto di *gestione globale*, che riguarda la gestione delle risorse all'interno della rete e lo scambio di flussi informativi preposti a compiti di esercizio e di manutenzione.

Lo scambio di ognuno di questi tre tipi di informazione attraverso la rete o nell'ambito di ogni elemento funzionale di una ISDN è descritto in un apposito piano. Si parla allora di *piano d'utente* (piano U), di *piano di controllo* (piano C) e di *piano di gestione* (piano M).

Mentre i primi due hanno una struttura a strati, il terzo è a sua volta suddiviso in due parti corrispondenti alle due finalità della M-informazione a cui si è accennato in precedenza. La prima di queste parti, che comprende le *funzioni di gestione di piano*, riguarda il coordinamento locale tra tutti i piani e non è stratificata. La seconda parte, che include le *funzioni di gestione di strato*, ha rilevanza attraverso tutti gli elementi funzionali della rete ed è di tipo stratificato. L'applicazione dei concetti di strato e di piano ha condotto alla definizione dell'elemento di base di un ISDN-PRM, e cioè del *blocco di protocolli* (PB - Protocol Block).

Un PB, che può essere considerato una generalizzazione del concetto di sistema utilizzato in un ambiente OSI, può descrivere vari gruppi funzionali nelle sezioni di accesso e interna di una ISDN. Alla prima categoria appartengono, ad esempio, un apparecchio terminale o una terminazione di rete; nella seconda sono invece identificabili una terminazione di commutatore, un punto di segnalazione o un punto di trasferimento della segnalazione.

In Fig. VIII.9 viene mostrata la struttura semplificata di un PB, quale è oggi definita per una N-ISDN. Si distinguono i tre piani, U, C e M e ognuno dei primi due è potenzialmente strutturato in 7 strati. Nel piano M non vengono distinte le funzioni di gestione di piano e di strato; si tiene conto, invece, della suddivisione del piano C in due parti, chiamate *piano di controllo locale* e *piano di controllo globale*. Il primo di questi riguarda scambi di informazione

di controllo tra entità adiacenti, mentre il secondo è coinvolto quando tali scambi avvengono tra entità remote.

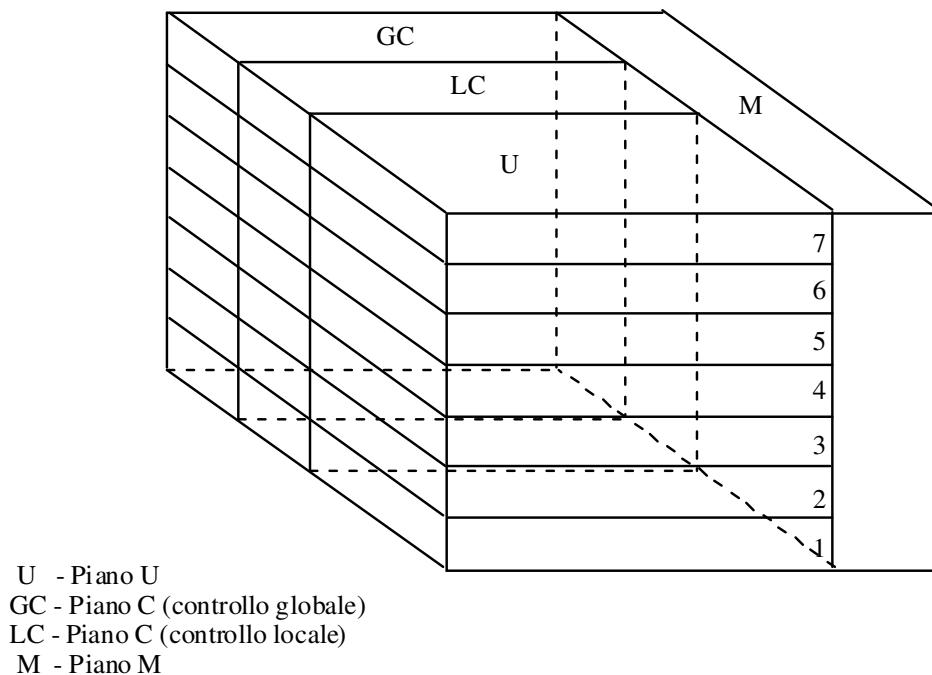


Fig. VIII.9 - Struttura semplificata di un PB per una N-ISDN.

In questa struttura, alcuni strati possono essere vuoti: ciò accade, ad esempio, nei PB che descrivono elementi della sezione interna di rete, ove, nel piano U, non sono normalmente presenti gli strati superiori al terzo.

All'interno di un PB sono possibili varie interazioni tra i raggruppamenti funzionali residenti in ognuno dei tre piani. Tra queste si possono citare le *interazioni tra i piani U e C*, che sono necessarie per sincronizzare le attività di questi piani. I piani U e C non interagiscono direttamente e possono comunicare solo con le entità del piano M, utilizzando apposite *primitive di gestione*. Tali entità, che svolgono funzioni di gestione di piano, provvedono così a coordinare le attività nei piani U e C. Nell'ambito poi di ognuno di questi, le *interazioni tra strati adiacenti* avvengono con primitive di servizio, secondo le usuali modalità previste nel modello OSI.

Se poi si considerano, come in Fig. VIII.10, le interazioni di un PB con l'esterno, si possono distinguere: 1) le *interazioni di un PB con i mezzi fisici*, e cioè quelle che debbono assicurare il trasferimento fisico delle informazioni pertinenti a ogni piano tra due PB; 2) le *interazioni in corrispondenza della faccia superiore di un PB*, e cioè quelle che riguardano la comunicazione tra ogni piano e vari processi applicativi; 3) le *interazioni tra entità paritetiche*, e cioè tra entità di un dato piano appartenenti allo stesso strato e residenti in PB

diversi.

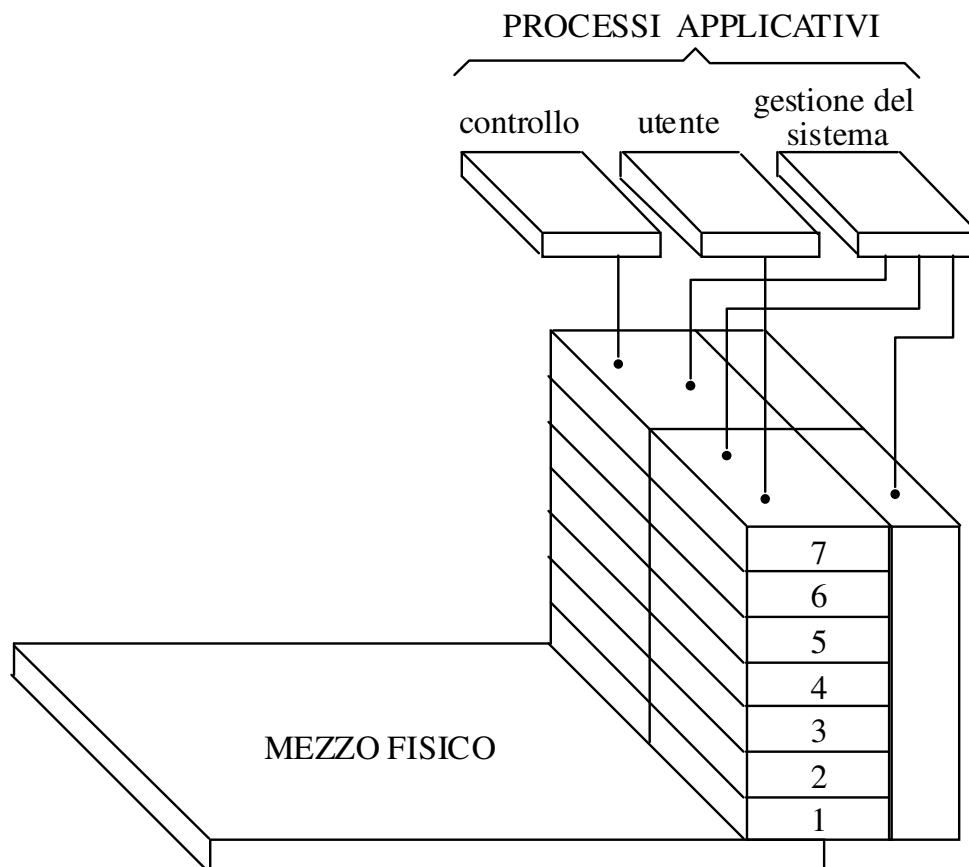


Fig. VIII.10 - Interazioni di un PB con l'esterno.

Circa le interazioni di cui in 1), il trasferimento fisico dei flussi informativi di piano può avere luogo, in alternativa, su mezzi fisici separati connessi ai singoli piani ovvero su un unico mezzo fisico condiviso. Questo secondo caso si verifica, ad esempio, per i piani U e C quando si impiega una segnalazione in banda ovvero quando la U-informazione è trasferita sul canale D.

Relativamente poi alle interazioni di cui in 2), queste comprendono il trasferimento di informazione da/verso le applicazioni d'utente, quelle di controllo e quelle di gestione di sistema.

Infine le interazioni tra PB diversi di cui in 3) si svolgono secondo gli usuali principi dell'architettura stratificata e non vengono considerati in Fig. VIII.10 per motivi di semplicità grafica.

Per l'utilizzazione dei PB come elementi componenti di un ISDN-PRM per rappresentare sistemi terminali o di rilegamento, è opportuno distinguere questi due casi. Per rappresentare un sistema terminale è sufficiente un singolo PB, mentre la rappresentazione di un sistema di rilegamento richiede l'impiego di due PB, di cui ognuno è la versione speculare dell'altro.

Ai fini dell'interpretazione di un ISDN-PRM, un concetto di base è quello di *contesto di comunicazione*, e cioè di scambio informativo tra due o più PB.

Da quanto sopra risulta evidente che i PB coinvolti in un contesto di comunicazione, che sono chiamati *PB terminali*, non corrispondono ai sistemi terminali del modello OSI, ma possono risiedere nell'apparecchio terminale, nei nodi di commutazione della rete o in altri apparati dell'ISDN. Un qualsiasi elemento dell'ISDN può avere molti contesti di comunicazione operanti in concorrenza o in sequenza. L'informazione passa da un contesto all'altro attraverso un processo applicativo o una entità del piano M.

A titolo di esempio consideriamo la Fig. VIII.11, che è relativa ad una connessione commutata a circuito attraverso il canale B. Sono identificabili quattro contesti di comunicazione, indicati con A, B, C e D.

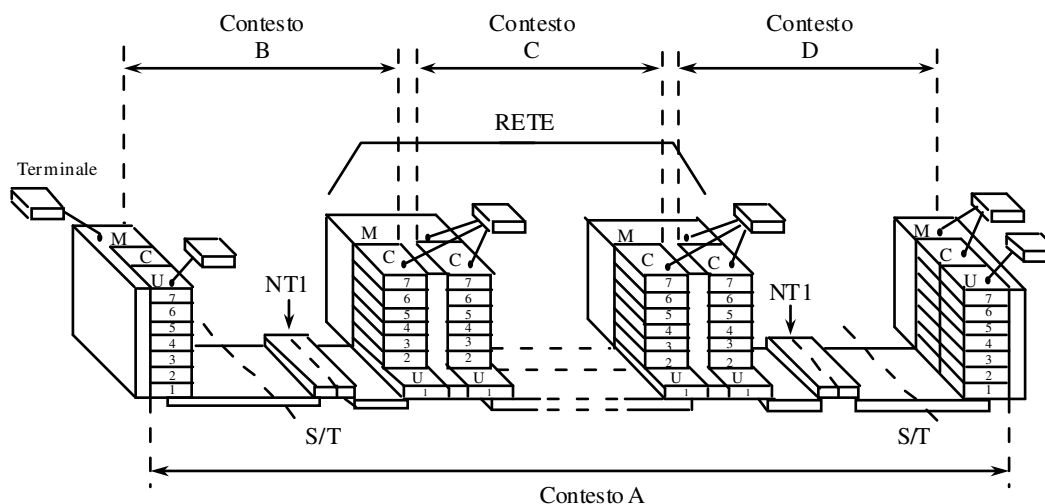


Fig. VIII.11 - Contesti di comunicazione in una connessione commutata a circuito attraverso il canale B.

Il contesto A rappresenta lo scambio informativo da utente a utente attraverso la connessione commutata a circuito. I contesti B e D sono nel dominio della segnalazione di utente tra un apparecchio terminale e la ISDN. Il contesto C è nel dominio della segnalazione inter-nodale, che si assume essere a canale comune.

Nel contesto A, i due apparecchi terminali contengono ognuno un PB terminale per tale contesto. Ogni NT1 effettua le funzioni di strato 1 necessarie per trasferire le cifre binarie sul canale di accesso tra il punto di riferimento T e la linea d'utente. I due nodi di commutazione fanno passare il flusso di cifre binarie relativo a questa chiamata attraverso le loro reti di connessione, controllate dalle entità di gestione dei nodi di rete. In ogni apparecchio terminale è presente un processo applicativo che genera e riceve l'informazione

trasferita sulla connessione commutata a circuito.

Passando poi ai contesti B e D è necessaria, per la segnalazione di utente, la presenza di un PB terminale sia nell'apparecchio terminale che nella terminazione dell'autocommutatore locale. All'interno dell'apparecchio terminale un processo avvia l'emissione di informazione di segnalazione verso la terminazione dell'autocommutatore. Presso questa terminazione, tale informazione di segnalazione è trasferita a un processo di gestione della chiamata, il quale può controllare la rete di connessione nel modo necessario a stabilire un circuito fisico all'interno della terminazione.

Per entrambi i contesti B e D va osservato che gli strati dei protocolli di segnalazione sono indipendenti dagli strati dei protocolli impiegati per il trasferimento dell'informazione da utente a utente, con l'eccezione dello strato 1, in quanto il canale D, preposto al trasporto della segnalazione, è multiplexato staticamente con il canale B per il trasporto dell'informazione d'utente.

Per quanto riguarda poi il contesto C, basta aggiungere che ognuna delle terminazioni di autocommutatore deve contenere un PB terminale per la segnalazione inter-nodale. Questa è avviata dai processi di gestione delle chiamate in ogni terminazione di autocommutatore, in accordo con quanto richiesto. Nella situazione prospettata in figura i protocolli di segnalazione sono indipendenti da quelli del contesto A per tutti i sette strati: infatti la segnalazione inter-nodale viaggia su un portante fisico indipendente.

## VIII.2 Lo strato 1 ISDN

Consideriamo lo strato fisico dell'interfaccia utente-rete con riferimento all'accesso di base in una ISDN a banda stretta. Per ciò che riguarda la composizione impiantistica della postazione di utente, sono state normalizzate la *configurazione punto-punto* e quella *punto-multipunto*, che sono illustrate in Fig. VIII.12. La seconda è un *bus passivo*, in cui i TE, in numero non superiore a 8 unità, possono comunicare solo con la rete e non tra di loro. È previsto anche un *bus esteso*, che è una combinazione delle due configurazioni precedenti.

L'interfaccia fisica utente-rete è a 8 fili con accoppiamento a trasformatore. Tra questi fili, due coppie sono dedicate al trasferimento dei segnali nelle due direzioni di trasmissione dal TE all'NT e viceversa. Le stesse due coppie di fili possono essere utilizzate per trasferire l'energia di alimentazione dall'NT al TE. Questa energia può essere fornita dall'autocommutatore locale o da altra sorgente, localizzata, ad esempio, nello stesso NT. La sorgente di alimentazione

si può rendere necessaria, dato che un TE in ambiente ISDN è un'apparecchiatura piuttosto complessa che può richiedere un'elevata potenza di alimentazione ed esistono strette limitazioni sulle possibilità di trasferimento di energia elettrica sui legamenti di utente realizzati con mezzi trasmissivi in rame o in fibra ottica.

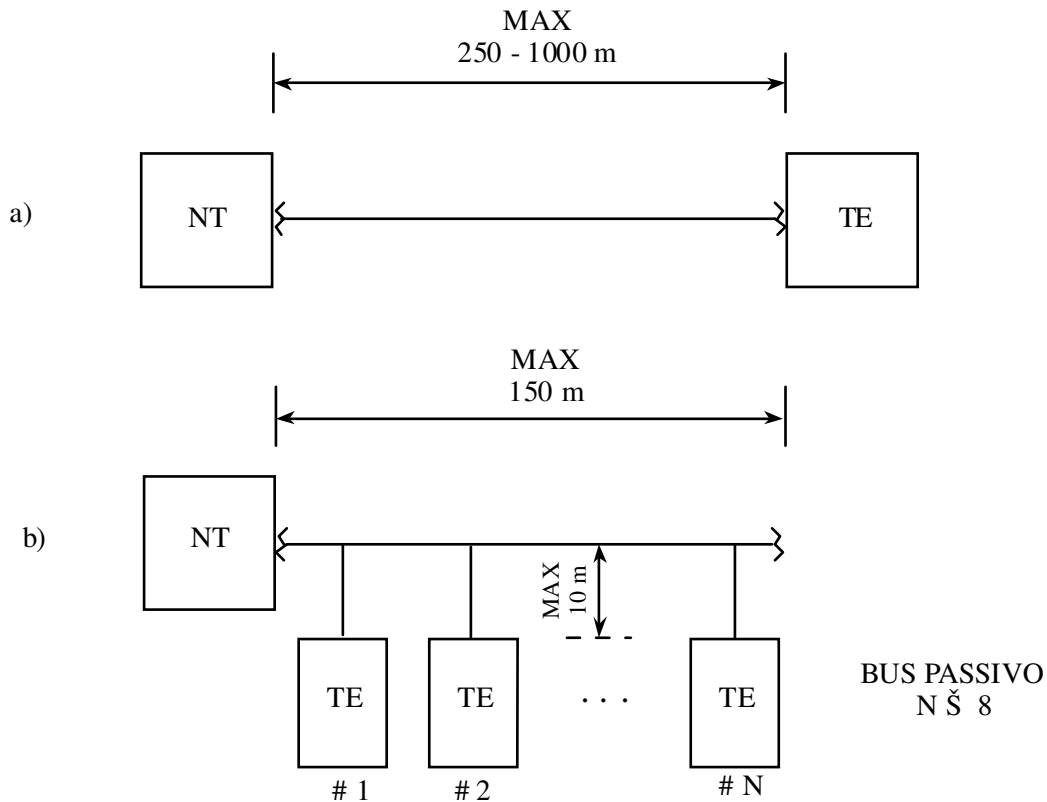
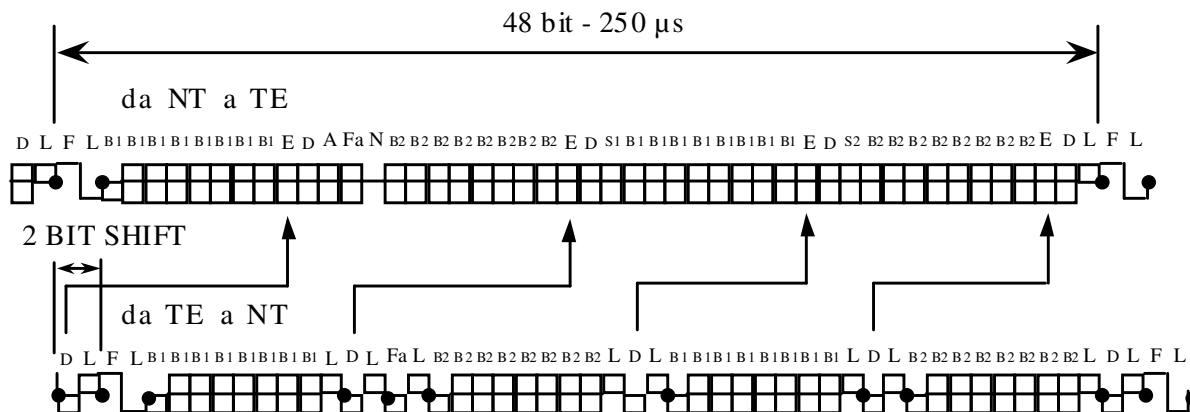


Fig. VIII.12 - Configurazioni di base per l'accesso a una ISDN: configurazioni punto-punto (a) e punto-multipunto (b).

L'aspetto più rilevante del protocollo di strato 1 è rappresentato dalla funzione MAC per l'accesso al canale D, che è necessaria nel caso di configurazione punto-multipunto, ove si presenta una condizione di accesso multiplo. Tale funzione è di attuazione particolarmente semplice, dato che sarebbe ingiustificato imporre un appesantimento non necessario alle configurazioni punto-punto, che non richiedono risoluzione di contese di accesso.

In Fig. VIII.13 è mostrato il particolare formato di trama di strato 1, che è utilizzato per moltiplicare staticamente i canali B e D nell'accesso di base e che si differenzia per le due direzioni di trasmissione.

In ambedue le direzioni, la trama è composta da 48 cifre binarie con una durata di 250  $\mu$ s. La capacità complessiva di trasferimento è pertanto di 192 kbit/s ( $48/(250 \cdot 10^{-6})$  bit/s). Nell'ambito di questa, la capacità netta è di 144 kbit/s ed è utilizzata come supporto di due canali B e di un canale D. La capacità residua, che è di 48 kbit/s, è invece utilizzata per il trasferimento di informazione aggiuntiva, i cui scopi saranno, almeno parzialmente, illustrati qui di seguito.



- |  |                                 |
|--|---------------------------------|
| F - Bit di sincronizzazione di trama   | B1 - Bit entro il canale B n° 1 |
| L - Bit di bilanciamento               | B2 - Bit entro il canale B n° 2 |
| D - Bit del canale D                   | S1-S2 - Bit non utilizzati      |
| E - Bit del canale d'eco               | N - Bit posto a $\bar{F}_A$     |
| Fa - Bit ausiliario di sincr. di trama | A - Bit di attivazione          |

Fig. VIII.13 - Struttura della trama di strato 1 nei punti di riferimento S e T dell'interfaccia utente-rete per l'accesso a una ISDN.

Circa l'utilizzazione della capacità netta, in ogni trama sono disponibili 2 ottetti di cifre binarie per ciascuno dei due canali B e 4 cifre binarie per il canale D. La capacità residua corrisponde invece a 12 cifre binarie.

Per ambedue le direzioni di trasmissione, il codice di linea impiegato è di tipo *pseudo-ternario*, con durata relativa percentuale di impulso uguale al 100%. Nella codifica, un UNO binario è rappresentato da assenza di segnale, mentre gli ZERO binari sono rappresentati, di regola, con impulsi alternativamente positivi e negativi.

Ciò non è però sufficiente ad assicurare che la componente continua del segnale sia soppressa trama per trama. A tale scopo sarebbe necessario che una trama, nella sua interezza o nei suoi vari gruppi componenti, contenesse un

numero pari di ZERO binari. Per conseguire questo risultato vengono utilizzati i *bit di bilanciamento L* (bit L). Ognuno di questi è uno ZERO, se il numero di ZERO che seguono il precedente bit L è dispari; è invece un UNO quando questo numero è pari.

Ogni trama si apre con il *bit di sincronizzazione di trama F* (bit F), seguito da un bit L. Il bit F è sempre uno ZERO ed è trasmesso con polarità positiva, violando la regola di bipolarità. Il bit L, associato al bit F, è sempre uno ZERO. La sincronizzazione di trama è ottenuta tramite la sequenza di queste due cifre, a cui segue una violazione di polarità: cioè il primo ZERO, dopo la coppia dei bit F e L, ha la stessa polarità del bit L (cioè polarità negativa).

La rappresentazione della trama in Fig. VIII.13 definisce l'organizzazione dei differenti canali e la polarità che può essere assunta da ogni cifra binaria.

Le trame trasmesse nella direzione da NT a TE supportano, in aggiunta ai canali sulle trame nel verso contrario, anche un *canale d'eco* (canale E), che è utilizzato dal meccanismo di controllo di accesso al canale D. Sul canale E l'NT ritrasmette verso i TE le cifre binarie ricevute sul canale D. L'ultimo bit della trama è il bit L, che sopprime la componente continua dell'intera trama.

Le trame trasmesse nella direzione da TE a NT sono costituite da 10 gruppi di cifre binarie. Tra questi, 6 gruppi comprendono 2 sole cifre binarie, mentre i rimanenti 4 sono costituiti da 1 ottetto. Ogni gruppo è individualmente bilanciato da un bit L, come necessario dato che ogni trama di questo tipo è il risultato della trasmissione dei suoi gruppi componenti da parte di differenti terminali e non sarebbe quindi possibile operare un bilanciamento utilizzando un singolo bit L. Inoltre ogni primo ZERO di un gruppo è trasmesso con una polarità negativa; ciò garantisce l'identità delle polarità di trasmissione degli ZERO nei canali su cui può avvenire collisione.

Per chiarire l'importanza di tale ultimo accorgimento, osserviamo che il segnale risultante sul bus, quando avviene trasmissione simultanea da parte di due o più terminali, è la somma algebrica dei segnali trasmessi dai singoli terminali. Conseguentemente, se si tiene conto della regola di codifica di linea e della soglia di decisione del ricevitore nell'NT, la trasmissione simultanea di due o più UNO viene letta in ricezione come un UNO, mentre quella di un UNO e di uno ZERO viene letta come uno ZERO. Invece, se vengono trasmessi simultaneamente due o più ZERO senza l'accorgimento sopra descritto, si potrebbero verificare due casi: se gli ZERO fossero trasmessi con la stessa polarità, in ricezione verrebbe letto uno ZERO; invece, se gli ZERO fossero trasmessi con polarità opposta, il ricevitore dell'NT rivelerebbe assenza di segnale e leggerebbe quindi un UNO.

Se, invece, viene assicurata l'identità della polarità di trasmissione degli ZERO, il codice di linea garantisce una operazione di AND logico sull'informazione emessa dai TE. In queste condizioni i TE, che non hanno informazione da trasferire sul canale D possono emettere una sequenza di tutti UNO, senza disturbare l'emissione di un altro TE.

In ognuna delle due direzioni di trasmissione, i quattro bit del canale D debbono essere separati, in modo tale che l'eco di una cifra binaria trasportata sul canale D nella direzione da TE a NT sia ricevuta dai terminali prima che avvenga la trasmissione della cifra binaria successiva. Dato che strutture di trama con durata di 125  $\mu$ s non consentono questa separazione del canale D, si è resa necessaria l'adozione di trame con durata di 250  $\mu$ s.

I terminali trasmettono una trama sincronizzata in cifra binaria a quella ricevuta, ma con uno sfasamento di due intervalli di cifra in ritardo. Questo sfasamento consente di massimizzare l'intervallo temporale tra i bit del canale D e i corrispondenti bit d'eco. Tale intervallo temporale è significativo, dato che esso fissa il massimo tempo di propagazione di andata e ritorno tra TE e NT e, conseguentemente, la lunghezza massima del bus. Il protocollo di accesso al canale D è descritto nel par. VIII.4.

Per concludere, si precisa che l'NT ricava la sua temporizzazione di emissione dall'orologio di rete. Un TE estrae invece la temporizzazione di bit e di trama dal flusso di cifre binarie provenienti dall'NT e utilizza questa temporizzazione per sincronizzare il segnale trasmesso.

### VIII.3 Protocollo CSMA/CR

Il protocollo *CSMA-CR* (Carrier Sense Multiple Access - Collision Resolution) è impiegato nell'accesso al canale D dell'interfaccia utente-rete in una N-ISDN. Esso assicura che, nel caso in cui due o più TE (apparecchi terminali) tentino di accedere a tale canale simultaneamente, solo uno di loro ha successo nell'emissione della sua informazione. Sono previste due classi di priorità di accesso e, nell'ambito di ognuna di queste, sono attuati accorgimenti per assicurare equità nell'assegnazione della banda disponibile del canale D.

Questo protocollo MAC è basato su quattro punti:

- il canale D utilizza, nello strato 2 dell'architettura N-ISDN, un protocollo in accordo con l'HDLC;
- i terminali possono ascoltare il canale D nella direzione da TE a NT utilizzando il canale d'eco;
- il bus effettua una operazione di AND logico sull'informazione emessa da

differenti terminali;

- un terminale inattivo emette una sequenza di tutti UNO.

Descriviamo sinteticamente questo protocollo MAC. Un TE ascolta il canale d'eco e, quando ha informazione da emettere sul canale D, conta il numero di UNO binari consecutivi ricevuti dal canale d'eco. Quando questo numero raggiunge un valore prefissato  $X$  (*soglia di conteggio*), il TE assume che non esistano altri TE in fase di emissione sul canale D e inizia quindi la sua trasmissione. La rivelazione di uno ZERO binario ri-inizializza il processo di conteggio.

Quando è in fase di emissione, il TE continua a rimanere in ascolto sul canale d'eco e confronta l'ultima cifra binaria emessa con quella successiva ricevuta dal canale d'eco. Se la prima è la stessa della seconda, il TE ha la conferma della validità della sua assunzione e può quindi continuare la sua emissione. In caso contrario, che si verifica in presenza di collisione tra l'emissione in corso e una o più da parte degli altri apparecchi terminali, il TE cessa immediatamente la sua emissione e ritorna nello stato di ascolto del canale d'eco.

Il valore  $X$  della soglia di conteggio è fissato in sede di costruzione o di installazione del TE, in modo da assicurare un meccanismo di accesso con priorità. Più bassi valori di  $X$  corrispondono a una più elevata classe di priorità. Tra le due classi di priorità previste, quella più alta è riservata all'informazione di segnalazione. Per queste due classi i valori di  $X$  sono fissati a 8 e a 10.

Per assicurare una equa assegnazione della capacità del canale D tra TE che appartengono alla stessa classe di priorità, la soglia di conteggio di un TE è aumentata di una unità (cioè è portata a  $X + 1$ , diminuendo così la priorità di accesso), quando il TE ha completato con successo l'emissione di una unità informativa di strato 2. Questo valore  $X + 1$  viene conservato fintantoché il TE rivela, nel suo stato di ascolto, un numero  $X + 1$  di cifre UNO binarie consecutive sul canale d'eco. Quando ciò si verifica, il TE riassume la sua soglia di conteggio originaria.

#### **VIII.4 II LAP-D**

Descriviamo il *protocollo di strato 2 per l'accesso al canale D*, che nel seguito verrà indicato, per brevità, con l'acronimo *LAPD* (Link Access Procedure for the D-channel).

Scopo del LAPD è il trasferimento dell'informazione di segnalazione e/o di utente tra entità di strato 3 attraverso l'interfaccia utente-rete di una ISDN

utilizzando il canale D.

Il LAPD richiede un canale D pienamente duplice e trasparente al bit; è invece indipendente dalla capacità di trasferimento di questo canale. La procedura è concepita per consentire installazioni di più apparecchi terminali all'interfaccia utente-rete e per permettere l'interazione tra più entità di strato 3.

Le funzioni fondamentali del LAPD sono: a) la fornitura di una o più connessioni di strato 2 (*2-connessioni*) su un'unica connessione di strato 1 (*1-connessione*), e cioè nell'ambito di una *multiplazione di strato 2*; b) la delimitazione, l'allineamento e la trasparenza dell'unità informativa impiegata nello strato 2 e chiamata *trama*; c) il controllo di sequenza per conservare l'ordine sequenziale delle trame tra le estremità di ogni 2-connessione; d) la rivelazione di errori di trasmissione, di formato e di procedura su una 2-connessione; e) il recupero degli errori di cui in d); f) la notifica, alle entità di gestione nel piano M, di errori non recuperabili; g) il controllo di flusso.

Le funzioni del LAPD forniscono i mezzi per trasferire informazione tra combinazioni di più punti terminali di connessione. A tale scopo sono utilizzate, in alternativa, *connessioni punto-punto* o *connessioni a diffusione*. Nel caso di trasferimento punto-punto, una trama è diretta a un singolo punto terminale, mentre nel caso di trasferimento a diffusione, una trama è diretta a uno o più punti terminali.

Per il trasferimento di informazioni di strato 3 sullo strato 2, sono definiti due modi operativi, che possono coesistere su un singolo canale D: quello *con riscontro* e quello *senza riscontro* (Fig. VIII.14).

Per il funzionamento senza riscontro, che è applicabile in trasferimenti sia punto-punto che diffusivi, vengono usate *trame informative non numerate* (trame UI); per queste è possibile, quindi, solo la rivelazione di errori di trasmissione e di formato, mentre non sono definiti meccanismi di recupero di errore e di controllo di flusso.

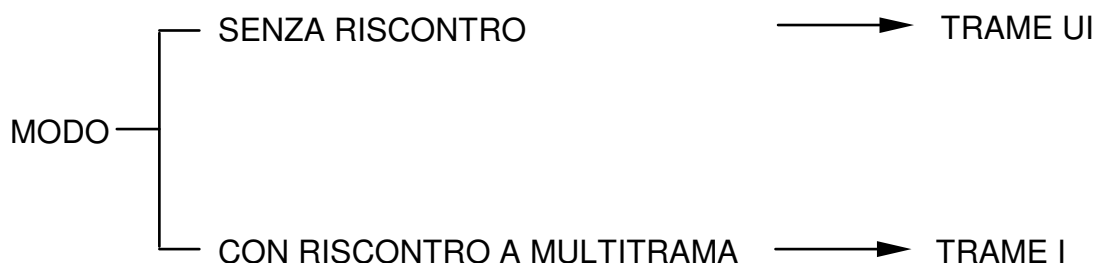


Fig. VIII.14 - Modi operativi del LAPD e relativi tipi di trama.

Per il funzionamento con riscontro, che è applicabile solo in trasferimenti

punto-punto, sono previste trame informative numerate; in questo caso sono possibili meccanismi di recupero di errore, di notifica degli errori non recuperabili e di controllo di flusso.

Per identificare una 2-connessione, una 2-entità utilizza un apposito *Identificatore*, che nel seguito indicheremo con *DLCI* (Data Link Connection Identifier). Questo è contenuto nel campo *Indirizzo* di ogni trama relativa a quella connessione ed è costituito da due elementi: il *SAPI* (Service Access Point Identifier) e il *TEI* (Terminal Endpoint Identifier). Il SAPI è utilizzato per identificare il punto di accesso al servizio del lato-rete o del lato-utente. Il TEI è invece impiegato per identificare uno specifico punto terminale di connessione all'interno di un SAP.

Il TEI può essere assegnato automaticamente per mezzo di un'apposita procedura di assegnazione oppure attribuito in sede di contratto di fornitura del servizio e introdotto nell'apparecchio terminale. In questo secondo caso, il valore del TEI deve essere verificato per assicurare che lo stesso valore non sia già usato in un altro apparato. La relazione tra SAPI, TEI e DLCI è illustrata in Fig. VIII.15.

L'identificazione di una 2-connessione da parte di una 3-entità è invece effettuata con un *Identificatore del punto terminale della 2-connessione* (CEI - Connection End point Identifier). Questo è costituito dal SAPI e da un *Suffisso del punto terminale di connessione* (CES - Connection End point Suffix). La corrispondenza tra CEI, SAPI e CES è mostrata in Fig. VIII.15.

Come mostrato in Fig. VIII.16, una 2-connessione punto-punto può essere in uno dei seguenti tre stati fondamentali: 1) *stato con TEI non assegnato*; in questo caso il TEI non è stato né assegnato, né verificato e conseguentemente non è possibile alcun trasferimento di informazione; 2) *stato con TEI assegnato*; in questo caso un TEI è stato assegnato/verificato e conseguentemente è possibile un trasferimento di informazione senza riscontro; 3) *stato a trama multipla*; in questo caso, che si verifica per effetto di un'apposita procedura, sono possibili trasferimenti senza riscontro o con riscontro a trama multipla. Una connessione diffusiva è sempre in uno stato in cui è consentito solo un trasferimento senza riscontro.

In accordo ai principi dell'architettura a strati, lo strato 2 fornisce servizi allo strato 3 e utilizza i servizi forniti dallo strato 1.

Per ciò che riguarda i servizi forniti allo strato 3, una loro descrizione, secondo il punto di vista dello strato 3, è data dalla specificazione delle primitive che sono scambiate tra gli strati 2 e 3 (Fig. VIII.17). Tra tali servizi si distinguono: a) il *servizio di trasferimento di informazione senza riscontro*; b) il

servizio di trasferimento di informazione con riscontro. Per questi è già stata data una presentazione sommaria; qui l'argomento viene ripreso per una descrizione più formale.

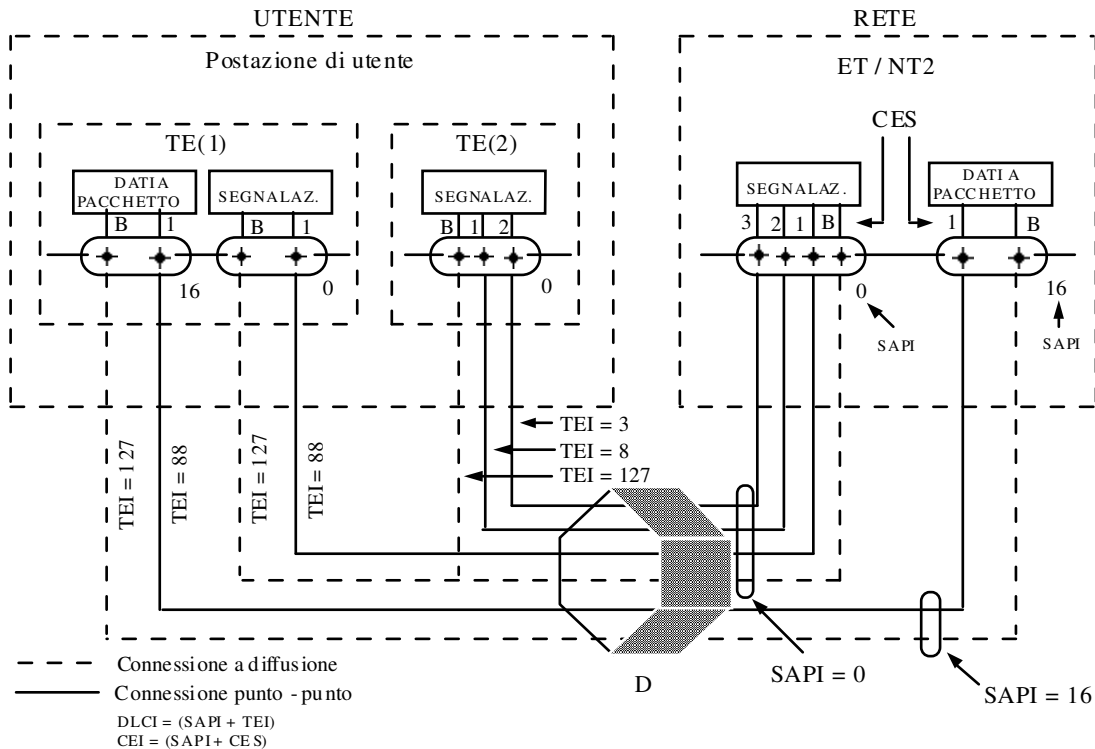


Fig. VIII.15 - Identificazione di 2-connessioni: relazione fra SAPI, TEI e DLCI.

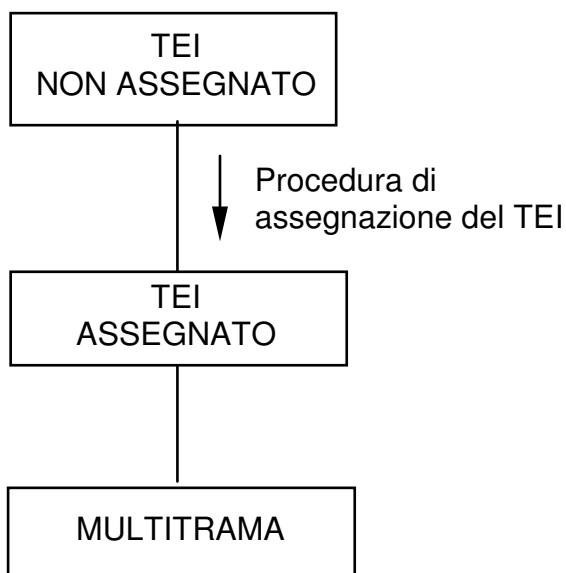


Fig. VIII.16 - Stati di una 2-connessione punto-punto.

Nome generico	TIPO			
	Richiesta	Indicazione	Risposta	Conferma
Strato 3 $\longleftrightarrow$ Strato 2				
DL-ESTABLISH	X	X	-	X
DL-RELEASE	X	X	-	X
DL-DATA	X	X	-	-
DL-UNIT-DATA	X	X	-	-
Piano M $\longleftrightarrow$ Strato 2				
MDL-ASSIGN	X	X	-	-
MDL-REMOVE	X	-	-	-
MDL-ERROR	-	X	X	-
MDL-UNIT-DATA	X	X	-	-
MDL-XID	X	X	X	X

Fig. VIII.17 - Primitive di servizio per l'interazione dello strato 2 con lo strato 3 e con il piano M.

Il servizio di trasferimento di informazione senza riscontro è utilizzato per connessioni punto-punto o a diffusione. Esso è caratterizzato da: A) la fornitura di una 2-connessione tra 3-entità per il trasferimento delle loro unità informative; B) l'identificazione dei punti terminali di ogni 2-connessione, per consentire a una 3-entità di identificare un'altra 3-entità; C) l'assenza di verifica, entro lo strato 2, dell'arrivo delle unità informative di cui in A). A questo elemento di servizio, che è di tipo non confermato, sono associate le primitive di richiesta e di indicazione DL-UNIT DATA.

Il servizio di trasferimento con riscontro è utilizzato solo per connessioni punto-punto ed è caratterizzato da:  $\alpha$ ) la fornitura di una 2-connessione tra 3-entità per il trasferimento con riscontro di unità informative di strato 3;  $\beta$ ) la identificazione dei punti terminali delle 2-connessioni, per consentire a una 3-entità di identificare un'altra 3-entità;  $\gamma$ ) l'integrità della sequenza delle unità informative di strato 2 in assenza di malfunzionamenti;  $\delta$ ) la notifica, alla entità paritetica, di malfunzionamenti, quali la perdita di sequenza;  $\epsilon$ ) la notifica, alla entità di gestione nel piano M, di errori non recuperabili rivelati dallo strato 2;  $\eta$ ) il controllo di flusso.

Nello svolgimento del servizio di trasferimento con riscontro, si

distinguono tre elementi e le corrispondenti primitive che sono usate per: 1) richiedere, indicare e confermare l'inizio del modo con riscontro (DL-ESTABLISH) a trama multipla; 2) richiedere, indicare e confermare la terminazione (DL-RELEASE) del modo con riscontro a trama multipla; 3) richiedere l'invio (indicare l'arrivo) di informazioni di strato 3 (DL-DATA), secondo il modo con riscontro a trama multipla.

Lo strato 2 fornisce servizi anche al piano M. In particolare viene consentita la comunicazione tra entità paritetiche di gestione; il relativo trasferimento di informazione è non riscontrato. La primitiva MDL - UNIT DATA è utilizzata per richiedere l'invio (per indicare l'arrivo) di una unità di messaggio.

Inoltre il piano M offre allo strato 2 i cosiddetti *servizi amministrativi*. Questi sono: a) l'assegnazione, la verifica e la rimozione di valori di TEI; b) il passaggio di parametri che caratterizzano la 2-connessione. Sono richiesti (indicati) con primitive di gestione, che sono denotate con il prefisso MDL (Management Data Link). Tali primitive, che sono elencate in Fig. VIII.17, sono utilizzate per l'assegnazione e la rimozione di un valore di TEI (MDL - ASSIGN e MDL - REMOVE, rispettivamente) e per la notifica di situazioni di errore tra lo strato 2 e il piano M (MDL - ERROR).

Per ciò che riguarda infine i servizi richiesti allo strato 1, questi sono: 1) la fornitura di una 1-connessione per il trasferimento trasparente di cifre binarie nello stesso ordine secondo cui queste sono presentate dallo strato 2; 2) l'indicazione dello stato fisico del canale D; 3) la trasmissione delle unità informative dello strato 2 in accordo alle loro priorità.

Si distinguono tre elementi di servizio e le associate primitive (Fig. VIII.18): a) per richiedere (per indicare) l'attivazione (PH-ACTIVATE) di una 1-connessione; b) per richiedere (per indicare) la disattivazione (PH-DEACTIVATE) di una 1-connessione; c) per richiedere l'invio (per indicare l'arrivo) di unità informative (PH-DATA) di strato 2.

Per l'interazione tra lo stato fisico e il piano M, si hanno anche (cfr. Fig. VIII.18) tre primitive di gestione, denotate con il prefisso MPH (Management Physical).

In Fig. VIII.19 sono mostrati i due possibili formati di trama per il LAPD: il *formato A* per trame che non contengono informazione d'utente e il *formato B* per trame che la contengono. Nelle seconde è quindi presente un campo informativo, che è invece assente nelle prime.

Nome generico	TIPO			
	Richiesta	Indicazione	Risposta	Conferma
Strato 2 $\longleftrightarrow$ Strato 1				
PH-DATA	X	X	-	-
PH-ACTIVATE	X	X	-	-
PH-DEACTIVATE	X	X	-	-
Piano M $\longleftrightarrow$ Strato 1				
MPH-ACTIVATE	-	X		
MPH-DEACTIVATE	X	X		
MPH-INFORMATION	-	X		

Fig. VIII.18 - Primitive di servizio per l'interazione dello strato 1 con lo strato 2 e con il piano M.

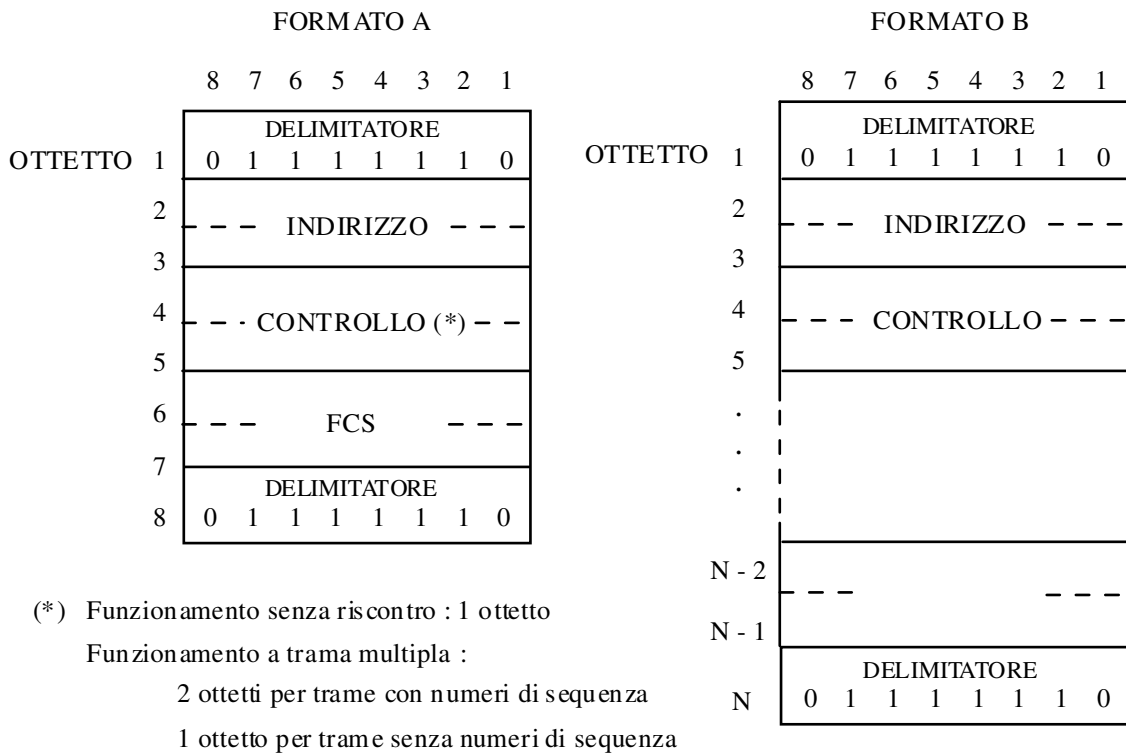


Fig. VIII.19 - Formati di trama per il LAPD.

Descriviamo brevemente i campi di queste trame, con l'avvertenza che questi sono gli stessi del protocollo HDLC, ma con contenuti che si diversificano sotto vari aspetti.

Il campo di delimitazione (Flag) ha lo stesso formato di quello omonimo del protocollo HDLC, con funzioni del tutto analoghe (cfr. § III.3). Esso apre e chiude ogni trama.

Il campo di indirizzo (Address), come mostrato in Fig. VIII.20, può essere esteso a più di un otetto con la seguente convenzione: la prima cifra binaria di ogni otetto, detta *bit EA* (Extended Address), indica se quello è o non è l'ultimo otetto dell'indirizzo

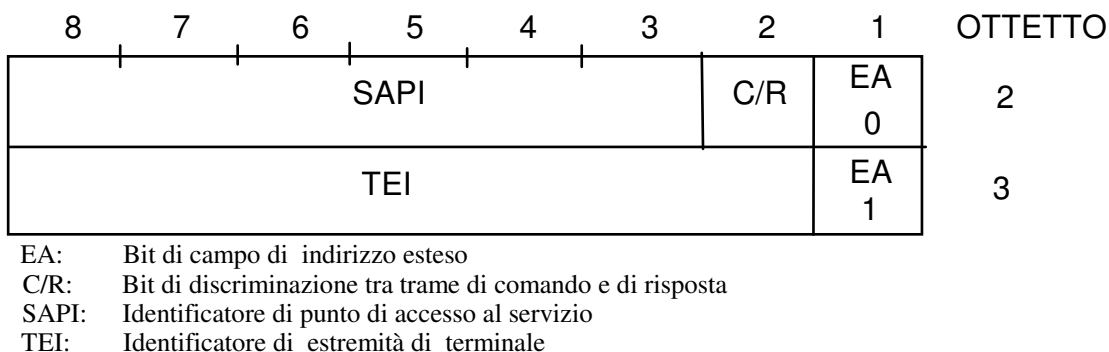


Fig. VIII.20 - Formato del campo di indirizzo in una trama del LAPD.

Il bit *C/R* distingue tra *trame di comando* e *trame di risposta*, secondo la regola illustrata in Fig. VIII.21. In conformità con l'HDLC, i comandi contengono l'indirizzo della stazione ricevente, mentre le risposte contengono quello della stazione emittente. L'indirizzo vero e proprio consta, come già detto, di un valore del SAPI e di uno del TEI. Per connessioni diffusive il valore del TEI è sempre rappresentato da una configurazione di sette UNO (corrispondente a 127).

	Lato Rete	Lato Utente
<b>Comandi da</b>	1	0
<b>Risposte a</b>	1	0
<b>Comandi a</b>	0	1
<b>Risposte da</b>	0	1

Fig. VIII.21 - Valore del bit C/R del campo di indirizzo in una trama del LAPD.

Il campo di controllo (Control) identifica il tipo di trama ovvero uno dei possibili elementi di procedura. La Fig. VIII.22 illustra tutti i suoi possibili formati. Le trame, come nel protocollo HDLC, si distinguono in: *trame*

informative (I), trame supervisive (S) e trame non numerate (U).

Cifre binarie del campo di controllo	8	7	6	5	4	3	2	1		
TRAMA I	N(S)							0	OTTETTO	4
	N(R)							P		
TRAMA S	X	X	X	X	S	S	0	1	OTTETTO	4
	N(R)							P/F		
TRAMA U	M	M	M	P/F	M	M	1	1	OTTETTO	4

- N(S) Numero:di sequenza in trasmissione
- N(R): Numero di sequenza in ricezione
- S: Identificatore del tipo di trama S
- M: Identificatore del tipo di trama U
- P/F: Bit "poll" quando emesso come un comando; Bit "final" quando emesso come una risposta
- X Bit riservato e posto a "zero"

Fig. VIII.22 - Formati del campo di controllo in una trama del LAPD.

In Fig. VIII.23 sono riportati tutti i possibili tipi di trame impiegate nel LAPD. In particolare i numeri  $N(R)$  e  $N(S)$  sono i noti numeri di sequenza in trasmissione e in ricezione, che sono modulo 128.

Applicazione	Formato	Comandi	Risposte	Codifica								Otteto	
				8	7	6	5	4	3	2	1		
Trasferimento di informazione riscontrata a trama multipla o non riscontrata	Trasferimento di informazione	I		N(S)							0	4	
				N(R)							P	5	
	Supervisivo	RR	RR	0	0	0	0	0	0	0	1	4	
				N(R)							P/F	5	
		RNR	RNR	0	0	0	0	0	1	0	1	4	
				N(R)							P/F	5	
		REJ	REJ	0	0	0	0	1	0	0	1	4	
				N(R)							P/F	5	
	Non numerato	SABME		0	1	1	P	1	1	1	1	4	
					DM	0	0	0	F	1	1	1	4
		UI		0	0	0	P	0	0	1	1	4	
					DISC	0	1	0	P	0	0	1	4
			UA		0	1	1	F	0	0	1	1	4
						FRMR	1	0	0	F	0	1	1
Gestione connessione	XID	XID	1	0	1	P/F	1	1	1	1	4		

- |       |   |      |                              |
|-------|---|------|------------------------------|
| I     | - Information                             | UI   | - Unnumbered information     |
| R     | - Receive ready                           | DISC | - Disconnect                 |
| RNR   | - Receive non ready                       | UA   | - Unnumbered acknowledgement |
| REJ   | - Reject                                  | FRMR | - Frame reject               |
| SABME | - Set Asynchronous balanced mode extended | XID  | - Exchange identification    |
| DM    | - Disconnected mode                       |      |                              |

Fig. VIII.23 - Comandi e risposte nel LAPD con numerazione modulo 8.

Il *campo informativo* (Information), quando presente (come nelle trame con formato B), segue il campo di controllo e precede il campo di controllo di errore. Esso contiene un numero intero di ottetti, che, in mancanza di diversa scelta, è uguale a 128 ottetti nel caso in cui l'informazione trasportata sia segnalazione e a 260 ottetti nel caso di dati.

Il *campo FCS* (Frame Check Sequence) è costituito da due ottetti e la sequenza in esso contenuta è calcolata in emissione o utilizzata in ricezione con le stesse modalità del protocollo HDLC.

L'informazione di strato 3 da non riscontrare viene convogliata su particolari trame U, chiamate *UI* (Unnumbered Information). Per ciò che riguarda poi il modo con riscontro a trama multipla, vengono utilizzate trame I, che possono essere riscontrate con altre trame I o con trame RR (Receive Ready) (cfr. Fig. VIII.14).

La Fig. VIII.24 illustra uno schema funzionale dello strato 2 supportato da un singolo canale D, quale può esistere sul lato-rete dell'interfaccia utente-rete. Lo schema relativo al lato-utente può essere un sotto-insieme di questo.

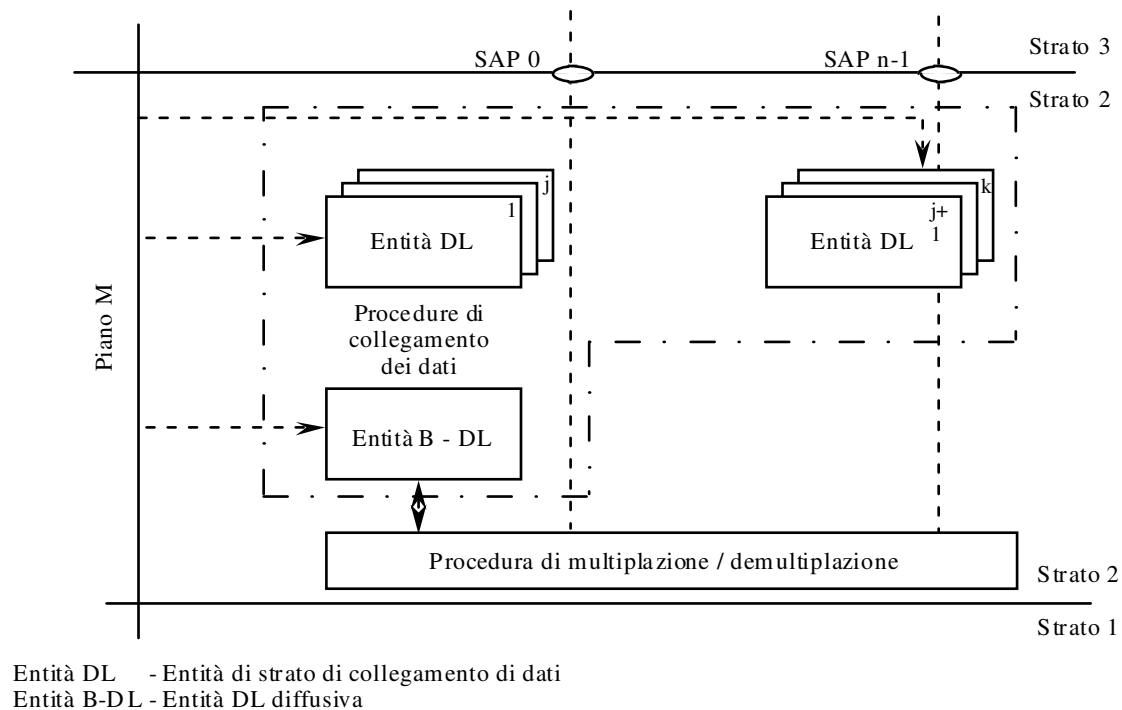


Fig. VIII.24 - Schema funzionale dello strato 2.

Si distinguono due tipi di procedure: quella di *moltiplicazione/demoltiplicazione* (multiplex procedure) e quella di *collegamento* (data link procedure).

La procedura di moltiplicazione/demoltiplicazione analizza, per ogni trama ricevuta dallo strato 1, il delimitatore, il campo FCS e il campo di indirizzo. Se la trama è corretta, la procedura provvede a smistarla ad un appropriato blocco di procedura di collegamento, in base al DLCI contenuto nel campo indirizzo. La procedura risolve anche le contese della 1-connessione da parte dei vari blocchi di procedura di collegamento, che abbiano trame da inviare; la risoluzione di queste contese è basata sul SAPI, dando priorità all'informazione di segnalazione. Una procedura di collegamento analizza il campo di controllo della trama ricevuta, invia risposta appropriata nell'ambito dello strato e indicazioni da strato a strato. Analizza inoltre le primitive di servizio dello strato 2 (quelle con il prefisso DL) ed emette appropriati comandi o risposte nell'ambito dello strato. In Fig. VIII.25 è riportato il modello funzionale della procedura di collegamento: si distinguono uno o più blocchi funzionali per connessioni punto-punto e a diffusione.

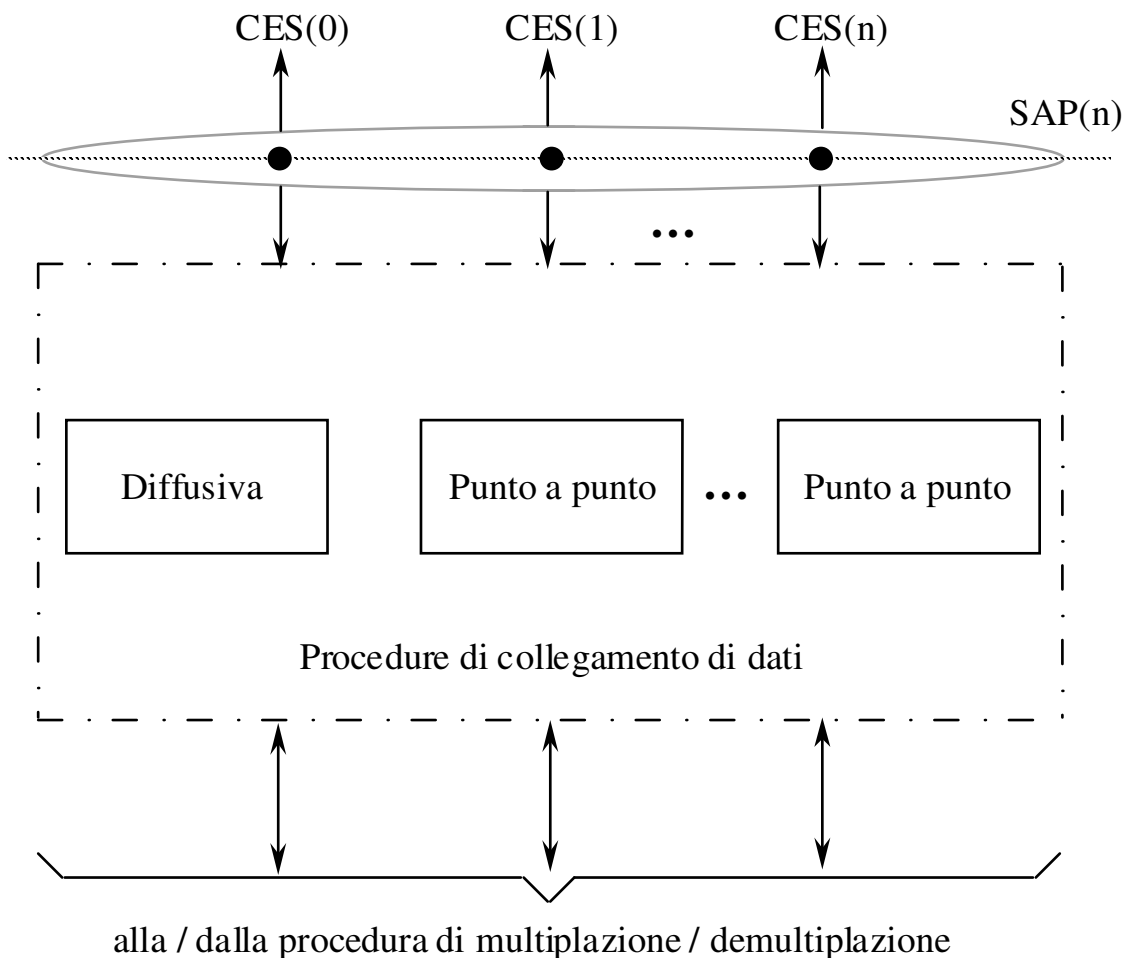


Fig. VIII.25 - Struttura della procedura di collegamento.

Ognuno di questi blocchi consiste di tre unità funzionali, e cioè di un controllo di emissione, di un controllo di ricezione e di un controllo dello stato della 2-conessione. Si nota che il SAPI identifica un blocco di procedura di collegamento, mentre il TEI identifica, nell'ambito del predetto blocco, una singola 2-conessione.



## **IX Bibliografia**

- [1] A. Roveri: "Reti di Telecomunicazioni", Principi Generali, Scuola Superiore G. Reiss Romoli, L'Aquila, 1995.
- [2] R. Gallager, D. Bertsekas: "Data Networks", Prentice Hall, 1992.
- [3] D. Comer: "Internetworking with TCP/IP", Volume I - Principles, Protocols and Architecture. Prentice Hall, 1991.
- [4] D. Comer, D. Stevens: "Internetworking with TCP/IP", Volume II - Design, Implementation and Internals. Prentice Hall, 1994.
- [5] D. Comer, D. Stevens: "Internetworking with TCP/IP", Volume III - Client-Server Programming and Applications for the AT&T TLI Version. Prentice Hall, 1994.
- [6] J. Cypser: "Communications for Cooperating Systems, OSI, SNA and TCP/IP, Addison-Wesley, 1992.
- [7] Ramteke: "Networks", Prentice Hall, 1994
- [8] W. Stallings: "ISDN and B-ISDN", Prentice Hall
- [9] A. Tanenbaum, "Computer Networks", Prentice Hall '89

## IX.1 Lista delle figure, tabelle ed equazioni

### IX.1.1 Figure

Fig. I.1 - Rete per dati a pacchetto.....	11
Fig. I.2 - Architettura funzionale della B-ISDN LFC. Local Function Capabilities, TE: Terminal Equipment.....	17
Fig. I.3 - Architettura del modello OSI.....	28
Fig. I.4 - Esempio di funzionamento di NFS.....	40
Fig. II.1 - Relazioni tra lo strato fisico e gli strati adiacenti dell'architettura OSI.....	46
Fig. II.2 - Schematizzazione di una connessione di strato fisico comprendente una entità di rilegamento con controllo operato dallo strato di rete.....	47
Fig. II.3 - Esempio di connessione fisica: .....	48
Fig. II.4 - Connessioni fisiche punto-punto e punto-multipunto.....	49
Fig. II.5 - Modello astratto del servizio di strato fisico.....	53
Fig. II.6 - Diagramma di transizione di stato per una connessione di strato fisico.....	55
Fig. II.7 - Schema generale di riferimento per una connessione fisica.....	57
Fig. II.8 - Schemi funzionali di: .....	58
Fig. II.9 - Illustrazione del principio di generazione di un segnale di banda base a quattro livelli, generato utilizzando l'impulso elementare rappresentato in (a), le ampiezze di impulso descritte in (b) e gli istanti significativi riportati in (c). .....	62
Fig. II.10 - Esempio di codifica a quattro livelli.....	63
Fig. II.11 - Esempi di segnale di banda base a due livelli: .....	65
Fig. II.12 - Spettri di densità di potenza di segnali di banda base a due livelli (a) o a tre livelli (b), ottenuti con varie codifiche di linea.....	66
Fig. II.13 - Schema funzionale di un ricevitore in banda base.....	69
Fig. II.14 - Esempio di interferenza intersimbolica:.....	72
Fig. II.15 - Esempi di risposta in ampiezza a coseno rialzato con coefficiente di arrotondamento uguale a: .....	74
Fig. II.16 - Diagramma ad occhio (b) di un canale trasmissivo ad interferenza intersimbolica nulla con segnale di linea a tre livelli e con impulso elementare rappresentato in (a). ....	77
Fig. II.17 - Spettro di densità di potenza di un segnale modulato FSK a due livelli, per vari valori dell'indice di modulazione $\mu$ . .....	81
Fig. II.18 - Diagrammi delle costellazioni per un segnale modulato PSK a due livelli (a), a quattro livelli (b) e a otto livelli (c). .....	83
Fig. II.19 - Diagramma delle costellazioni per un segnale modulato misto in ampiezza e in fase a 16 livelli secondo la Racc. V.29 (a) e diagrammi ridotti con 8 livelli (b) e con 4 livelli (c). .....	86
Fig. II.20 - Caratteristiche meccaniche dei connettori:.....	91
Fig. II.21 - Schemi di riferimento per la definizione delle interfacce elettriche: .....	92
Fig. II.22 - Definizione dei fili dell'interfaccia V.24 e loro corrispondenza con i piedini del connettore ISO 2110.....	97
Fig. II.23 - Configurazione dell'interfaccia X.24.....	99
Fig. II.24 - Esempio di procedura relativa alla Racc. V.24.....	102
Fig. II.25 - Connessione terminale-elaboratore tramite "null modem". .....	106
Fig. II.26 - Convenzioni per la rappresentazione dei diagrammi di transizione di stato relativi alla Racc. X.21. ....	107

Fig. II.27 - Diagramma di transizione di stato della Racc. X.21 relativo alla <i>fase di riposo</i> .	108
Fig. II.28 - Diagramma di transizione di stato della Racc. X.21 relativo alla <i>fase di controllo di chiamata</i> .....	110
Fig. II.29 - Diagramma di transizione di stato della Racc. X.21 relativo alla <i>fase di disconnessione</i> .....	111
Fig. II.30 - Diagramma temporale relativo alla fase di instaurazione di una chiamata mediante il protocollo X.21. ....	113
Fig. III.1 - Relazione tra lo stato di collegamento e gli strati adiacenti dell'architettura OSI. ....	116
Fig. III.2 - Modello astratto di una DL-connessione. ....	119
Fig. III.3 - Diagramma di transizione di stato relativo ad una estremità di DL-connessione per la fornitura di un DL-servizio orientato alla connessione. ....	124
Fig. III.4 - Sequenza di primitive nel caso di instaurazione di una DL-connessione con riferimento a: (a) instaurazione iniziata da un singolo DLS-utente; (b) instaurazione iniziata simultaneamente da due DLS-utenti.....	125
Fig. III.5 - Sequenza di primitive nel caso di mancata instaurazione di una DL-connessione con riferimento a: (a) instaurazione rifiutata di un DLS-utente; (b) instaurazione non consentita da un DLS-fornitore. ....	125
Fig. III.6 - Sequenza di primitive nel caso della fase di trasferimento di dati.....	126
Fig. III.7 - Sequenza di primitive nel caso di attivazione di una procedura di re-inizializzazione, con riferimento a re-inizializzazione: (a) iniziata da un singolo DLS-utente; (b) iniziata simultaneamente da due DLS-utenti; (c) iniziata dal DLS-fornitore; (d) iniziata simultaneamente da un DLS-utente e dal DLS-fornitore.....	126
Fig. III.8 - Sequenza di primitive nel caso di abbattimento di una DL-connessione, con riferimento ad abbattimento: (a) iniziato da un DLS-utente; (b) iniziato simultaneamente da due DLS-utenti; (c) iniziato dal DL-fornitore; (d) iniziato simultaneamente da un DLS-utente e dal DLS-fornitore.....	127
Fig. III.9 - Modello astratto di un DL-servizio nella modalità senza connessione.....	128
Fig. III.10 - Sequenza di primitive nel trasferimento di dati con riferimento al caso di un DL-servizio con modalità senza connessione. ....	129
Fig. III.11 - Diagramma di transizione di stato relativo a un DL-SAP per la fornitura di un DL-servizio nella modalità senza connessione. ....	129
Fig. III.12 - Esempio di colloquio con un protocollo orientato al carattere. ....	133
Fig. III.13 - Struttura generale di una trama in un protocollo orientato al bit. ....	134
Fig. III.14 - Delimitazione della trama con il metodo del conteggio dei caratteri: (a) procedimento di ricostruzione delle trame; (b) ricostruzione delle trame in caso di errore su un carattere indicatore di lunghezza. ....	135
Fig. III.15 - Operazione di riempimento di carattere: (a) sequenza originaria di caratteri; (b) sequenza dopo l'operazione di riempimento; (c) sequenza dopo l'operazione di svuotamento. ....	137
Fig. III.16 - Operazione di riempimento di bit: (a) sequenza originaria di cifre binarie; (b) sequenza dopo l'operazione di riempimento; (c) sequenza dopo l'operazione di svuotamento. ....	137
Fig. III.17 - Controllo di parità a blocchi: (a) blocco di carattere originario e configurazione dei bit di parità; (b) errore trasmissivo non rivelato dal controllo di parità. ....	141
Fig. III.18 - Esempio di utilizzazione di riscontri positivi e negativi.....	143
Fig. III.19 - Esempi di uso dei temporizzatori e della numerazione delle trame: (a) recupero di errore su una trama; (b) duplicazione di una trama in seguito a perdita di un riscontro; (c) uso del numero di sequenza in emissione. ....	144

Fig. III.20 - Esempio di utilizzazione della modalità a riscontro positivo con ri-emissione..	146
Fig. III.21 - Esempio di aggiornamento della finestra di trasmissione.....	147
Fig. III.22 - Esempio di aggiornamento della finestra di ricezione.....	149
Fig. III.23 - Esempio di utilizzazione della modalità con ri-emissione per il recupero di errore: (a) ri-emissione non selettiva; (b) ri-emissione selettiva.....	151
Fig. III.24 - Controllo del trasferimento di informazione: (a) con invito a ricevere; (b) con invito a emettere.....	154
Fig. III.25 - Configurazione asimmetrica con procedura sbilanciata.....	155
Fig. III.26 - Configurazione simmetrica con procedura sbilanciata.....	156
Fig. III.27 - Procedura bilanciata.....	157
Fig. III.28 - Formato della trama del protocollo X.25 di livello 2.....	159
Fig. III.29 - Diagramma semplificato di transizione di stato delle procedure di instaurazione e di abbattimento di una DL-conessione mediante il protocollo X.25 di livello 2.....	162
Fig. III.30 - Esempi di applicazione della procedura di recupero degli errori nel protocollo X.25 di livello 2, con riferimento a: (a) trasferimento corretto delle trame; (b) errore su una trama informativa; (c) errore su un riscontro.....	163
Fig. IV.1 - Interazione tra instradamento e controllo di flusso.....	172
Fig. IV.2 - Interazione tra instradamento e controllo di flusso.....	173
Fig. IV.3 - Applicazione dell'algoritmo di instradamento del percorso più breve; A e D sono i nodi di origine e di destinazione.....	178
Fig. IV.4 - Limiti dell'algoritmo di instradamento del percorso più breve.....	180
Fig. IV.5 - Esempio di algoritmo diffusivo.....	183
Fig. IV.6 - Esempio di spanning tree.....	184
Fig. IV.7 - Esempio di inter-rette.....	185
Fig. IV.8 - Diagramma di transizione di stato relativo a una estremità di N-conessione per la fornitura di un N-servizio orientato alla connessione.....	191
Fig. IV.9 - Possibili sequenze di primitive dell'N-servizio con riferimento a:.....	192
Fig. IV.10 - Multiplazione di più circuiti virtuali (N-conessioni) su un'unica DL-conessione. .....	194
Fig. IV.11 - Classificazione dei pacchetti definiti nel protocollo X.25 di livello 3.....	196
Fig. IV.12 - Formato dei pacchetti di "richiesta di chiamata" (CAR) e di "chiamata entrante" (INC) nel protocollo X.25 di livello 3.....	197
Fig. IV.13 - Formato dei pacchetti di dati (D) nel protocollo X.25 di livello 3, nei casi di: (a) numerazione modulo 8; (b) numerazione modulo 128.....	198
Fig. IV.14 - Procedure di instaurazione (a) e di abbattimento (b) di un circuito virtuale nel protocollo X.25 di livello 3.....	204
Fig. IV.15 - Diagramma semplificato di transizione di stato, relativo all'interfaccia DTE-DCE durante l'instaurazione di una chiamata virtuale nel protocollo X.25 di livello 3.....	205
Fig. IV.16 - Diagramma semplificato di transizione di stato relativo all'interfaccia DTE-DCE durante l'abbattimento di una chiamata virtuale nel protocollo X.25 di livello 3.....	205
Fig. IV.17 - Evoluzione di una chiamata virtuale nel protocollo X.25 di livello 3.....	208
Fig. IV.18 - Procedure di trasferimento dei dati (a) e di interrupt (b) nel protocollo X.25 di livello 3.....	210
Fig. IV.19 - Esempio di finestre di trasmissione (a) e di ricezione (b) secondo il protocollo X.25 di livello 3.....	213
Fig. V.1- Procedura di connessione eseguita con successo.....	221
Fig. V.2 - Procedura di connessione fallita per rifiuto del chiamato.....	222
Fig. V.3 - Procedura di connessione fallita per rifiuto del fornitore.....	222
Fig. V.4 - Procedura di disconnessione per iniziativa di un utente.....	223

Fig. V.5 - Procedura di disconnessione per iniziativa del fornitore .....	224
Fig. V.6 - Diagramma di stato del servizio di trasporto .....	228
Fig. V.7 - Un esempio di automa generatore di stringhe.....	228
Fig. V.8 - Relazione tra tipi di reti e classi di protocollo di trasporto.....	232
Fig. V.9 - Funzioni di trasporto e classi di protocollo.....	233
Fig. V.10 - Formato generale di una T-PDU .....	234
Fig. V.11 - Formato delle parti fisse delle T-PDU CR e CC .....	236
Fig. V.12 - Formato delle T-PDU DT e ED .....	239
Fig. VI.1 - Topologia a bus bidirezionale.....	246
Fig. VI.2 - Struttura dell'interfaccia tra mezzo trasmissivo e MAU nella rete Ethernet. ....	246
Fig. VI.3 - Topologia ad albero costituita da bus bidirezionali interconnessi.....	247
Fig. VI.4 - Topologie a bus unidirezionale: a) bus unidirezionale ripiegato; b) bus unidirezionale doppiamente ripiegato. ....	248
Fig. VI.5 - Riconfigurazione, in caso di guasto, di una rete a bus unidirezionale: a) funzionamento normale; b) rete riconfigurata.....	249
Fig. VI.6 - Topologia a doppio bus unidirezionale. ....	249
Fig. VI.7 - Riconfigurazione, in caso di guasto, di una rete a doppio bus unidirezionale: a) funzionamento normale; b) rete riconfigurata.....	250
Fig. VI.8 - Topologia ad anello. ....	251
Fig. VI.9 - Riconfigurazione, in caso di guasto, di una rete ad anello: a) funzionamento normale; b) rete riconfigurata. ....	252
Fig. VI.10 - Topologia a stella.....	253
Fig. VI.11 - Architettura logica di un terminale LAN.....	254
Fig. VI.12 - Norme IEEE 802 e loro relazione con il modello OSI. ....	255
Fig. VI.13 - Portata media del protocollo ALOHA .....	261
Fig. VI.14 - Ritardo medio di accesso D del protocollo ALOHA.....	262
Fig. VI.15 - Portata media del protocollo Slotted ALOHA.....	263
Fig. VI.16 - Ritardo medio di accesso del protocollo Slotted ALOHA. ....	264
Fig. VI.17 - Effetto delle procedure di persistenza: .....	266
Fig. VI.18 - Illustrazione del tempo di rivelazione di una collisione. ....	268
Fig. VI.19 - Struttura della MAC PDU del protocollo CSMA/CD.....	270
Fig. VI.20 - Creazione di un anello logico su un bus bidirezionale. ....	272
Fig. VI.21 - Struttura della MAC PDU del protocollo a testimone su bus.....	273
Fig. VI.22 - Formato della PDU token nel protocollo a testimone su bus. ....	275
Fig. VI.23 - Fasi del protocollo a testimone su anello: .....	276
Fig. VI.24 - Struttura della MAC PDU, del token e della sequenza di arresto del protocollo a testimone su anello. ....	277
Fig. VI.25 - Formato dei singoli campi della PDU del protocollo a testimone su anello. ....	279
Fig. VI.26 - Primitive di servizio definite per lo strato LLC.....	282
Fig. VI.27 - Struttura della LLC PDU. ....	287
Fig. VI.28 - Struttura del campo di indirizzo di una LLC PDU.....	287
Fig. VI.29 - Codifica del campo di controllo delle PDU definite per il funzionamento di tipo 1.....	289
Fig. VI.30- Codifica del campo di controllo delle PDU definite per il funzionamento di tipo 2. ....	291
Fig. VI.31 - Architettura logica dei protocolli IEEE 802.6. ....	293
Fig. VI.32 - Struttura di una stazione DQDB.....	294
Fig. VI.33 - Struttura di uno slot nel protocollo DQDB.....	294
Fig. VI.34 - Struttura del campo ACF.....	295

Fig. VI.35 - Algoritmo di coda distribuita: a) aggiornamento del contatore RQ; b) aggiornamento del contatore CD.....	296
Fig. VI.36 - Inserimento dei segmenti nella coda distribuita. ....	298
Fig. VI.37 - Trasmissione dei segmenti in coda distribuita.....	299
Fig. VIII.1 - Interfacce di accesso a una ISDN.....	304
Fig. VIII.2 - Configurazioni di riferimento per l'interfaccia utente-rete in una ISDN.....	306
Fig. VIII.3 - Esempi di configurazioni di accesso. ....	308
Fig. VIII.4 - Modello architetturale di base di una ISDN.....	314
Fig. VIII.5 - Gruppi funzionali e punti di riferimento per l'accesso a una ISDN. ....	315
Fig. VIII.6 - Accesso, tramite il canale B, a una rete per dati a pacchetto. ....	317
Fig. VIII.7 - Accesso, tramite il canale B, per la fornitura del servizio di circuito virtuale in una ISDN.....	318
Fig. VIII.8 - Accesso, tramite il canale D, per la fornitura del servizio di circuito virtuale in una ISDN.....	318
Fig. VIII.9 - Struttura semplificata di un PB per una N-ISDN.....	322
Fig. VIII.10 - Interazioni di un PB con l'esterno. ....	323
Fig. VIII.11 - Contesti di comunicazione in una connessione commutata a circuito attraverso il canale B. ....	324
Fig. VIII.12 - Configurazioni di base per l'accesso a una ISDN: configurazioni punto-punto (a) e punto-multipunto (b). ....	326
Fig. VIII.13 - Struttura della trama di strato 1 nei punti di riferimento S e T dell'interfaccia utente-rete per l'accesso a una ISDN. ....	327
Fig. VIII.14 - Modi operativi del LAPD e relativi tipi di trama.....	331
Fig. VIII.15 - Identificazione di 2-connessioni: relazione fra SAPI, TEI e DLCI.....	333
Fig. VIII.16 - Stati di una 2-connessione punto-punto. ....	333
Fig. VIII.17 - Primitive di servizio per l'interazione dello strato 2 con lo strato 3 e con il piano M. ....	334
Fig. VIII.18 - Primitive di servizio per l'interazione dello strato 1 con lo strato 2 e con il piano M. ....	336
Fig. VIII.19 - Formati di trama per il LAPD. ....	336
Fig. VIII.20 - Formato del campo di indirizzo in una trama del LAPD.....	337
Fig. VIII.21 - Valore del bit C/R del campo di indirizzo in una trama del LAPD. ....	337
Fig. VIII.22 - Formati del campo di controllo in una trama del LAPD.....	338
Fig. VIII.23 - Comandi e risposte nel LAPD con numerazione modulo 8.....	339
Fig. VIII.24 - Schema funzionale dello strato 2. ....	340
Fig. VIII.25 - Struttura della procedura di collegamento. ....	341

### IX.1.2 Tabelle

Tab. II.1 - Classi di utenza nelle reti pubbliche per dati e nell'ISDN in accordo alla Racc. X.1 .....	51
Tab. II.2 - Primitive di servizio dello strato fisico. ....	54
Tab. II.3 - Standard ISO relativi alla definizione delle caratteristiche meccaniche dei connettori.....	90
Tab. II.4 - Specifiche elettriche del generatore definite .....	95
Tab. II.5 - Raccomandazioni ITU-T relative alla definizione delle caratteristiche procedurali dello strato fisico. ....	101
Tab. III.1 - Primitive di un DL-servizio con modalità orientata alla connessione. ....	123
Tab. III.2 - Elenco dei caratteri di controllo nei protocolli orientati al carattere.....	132

Tab. III.3 - Corrispondenza tra modalità di trasmissione dei dati e modalità di utilizzazione del collegamento. ....	158
Tab. III.4 - Tipi di trame definite nel protocollo X.25 di livello 2. ....	160
Tab. IV.1 - Primitive del servizio di rete e loro parametri. ....	189
Tab. IV.2 - Tipi di pacchetti definiti nel protocollo X.25 di livello 3. ....	202
Tab. V.1 - Elenco degli elementi di servizio di trasporto. ....	220
Tab. V.2 - Elenco delle unità dati di protocollo di trasporto. ....	234
Tab. V.3 - Elenco dei parametri delle T-PDU CR e CC. ....	238
Tab. VI.1 - Primitive della classe di servizio senza connessione. ....	283
Tab. VI.2 - Primitive della classe di servizio con connessione. ....	285
Tab. VI.3 - Primitiva di scambio dati con lo strato MAC. ....	286
Tab. VI.4 - LLC PDU gestite dalle due classi di LLC-entità. ....	288
Tab. VIII.1 - Capacità di trasferimento dei canali B e H (in kbit/s) ....	310
Tab. VIII.2 - Strutture di interfaccia con il canale B. ....	311
Tab. VIII.3 - Strutture di interfaccia con i canali H. ....	311

*IX.1.3 Equazioni*

Eq. II.1 $C = B \lg_2 \left(1 + \frac{S}{N}\right)$ .....	59
Eq. II.2 $F_L = F_S / \log_2 q$ .....	64
Eq. II.3 $\mu = \frac{F_M - F_I}{F_S}$ .....	80
(Eq. VI.1): $G = T \Lambda$ .....	260
(Eq. VI.2): $S = G P_0$ .....	260
(Eq. VI.3): $P_0 = \Pr\{\text{nessuna trasmissione in } 2T\} = e^{-2\Lambda T} = e^{-2G}$ .....	260
(Eq. VI.4): $S = G e^{-2G}$ .....	260
(Eq. VI.5): $D = T + K (R_1 + R_2)$ .....	261
(Eq. VI.6): $1 + K = G/S$ .....	261
(Eq. VI.7): $K = e^{2G} - 1$ .....	261
(Eq. VI.8): $D = T + (e^{2G} - 1) (R_1 + R_2)$ .....	261
(Eq. VI.9): $S = G e^{-G}$ .....	262
(Eq. VI.10): $D = T + T/2 + K (R_1 + T/2 + R_2)$ .....	263
(Eq. VI.11): $K = e^G - 1$ .....	263
(Eq. VI.12): $D = T + T/2 + (e^G - 1) (R_1 + T/2 + R_2)$ .....	263

Future versioni:

Architetture di interconnessione nel capitolo 1  
Relativa modifica delle LAN con arch Protoll  
LAN ad alta velocità

Appendice con organizzazione di Internet ed esempi di comandi

Distinzione tra host e router: usa sistema  
Distingui meglio tra Internet e inter-rete  
memorie allocate in TCP nei SAP: definire meglio  
Segnalazione in banda di IP  
definire meglio controllo di flusso e di congestione  
  
sostituire time-out con fuori tempo massimo