

Notiziario Tecnico

Archivio

2/1995

 **TIM**

I protocolli di routing in ambiente TCP/IP

F. Antonelli, M. Carissimi, F. Iuso, F. Pugliese (*)

Le dimensioni e la complessità di Internet sono continuamente in crescita, determinando oltre all'aumento degli utilizzatori, la necessità di gestire automaticamente la registrazione dei nuovi punti terminali della rete e l'instradamento delle unità dati dalla sorgente alla destinazione attraverso i nodi intermedi della rete.

L'aggiornamento delle tabelle di routing, le tabelle cioè che governano il rilancio delle trame dati da un nodo al successivo, avviene attraverso l'utilizzo di particolari protocolli, detti protocolli di routing.

Nel seguito del documento vengono descritti i meccanismi di funzionamento dei principali protocolli di routing utilizzati in ambiente Transport Control Protocol/Internet Protocol, e che di fatto sono utilizzati per l'aggiornamento delle tabelle di routing della rete Internet.

1. L'instradamento in IP

La funzionalità di instradamento (*routing*) implementata nel protocollo IP (Internet Protocol) prevede l'individuazione della entità di rete successiva, sia essa l'host di destinazione o un router di transito, nel percorso che il pacchetto IP deve seguire per raggiungere la destinazione finale.

Sia gli host che i router partecipano dunque all'instradamento dei pacchetti. I router si interfacciano a due o più reti e provvedono a ricevere pacchetti da una interfaccia ed a rilanciarli attraverso un'altra interfaccia opportunamente individuata; gli host invece sono (generalmente) connessi ad una sola rete, per cui non trasferiscono pacchetti da una rete all'altra.

E' possibile fare una prima distinzione sulla base delle reti attraversate dal pacchetto IP per raggiungere la destinazione:

- si parla di *routing diretto* nello scambio di pacchetti tra host attestati ad una medesima rete;
- si parla di *routing indiretto* se nel trasferimento dei pacchetti a destinazione vengono interessati router IP.

In base alla distinzione ora fatta emergono due questioni. In primo luogo la modalità di riconoscimento di appartenenza alla medesima rete e in secondo luogo la determinazione del router successivo.

Ricordando il significato del **net_id** all'interno

dell'indirizzo IP [7], per determinare l'appartenenza ad una data rete sarà sufficiente confrontare il **net_id** con l'indirizzo di tale rete.

Nel caso di routing indiretto, il protocollo IP ricava l'indirizzo del router successivo a cui inviare il pacchetto consultando un'apposita tabella (*routing table*) contenente per un data rete di destinazione l'indirizzo del router successivo a cui deve essere inviato il pacchetto. In tale tabella oltre alle corrispondenze <rete-destinazione, router successivo> viene anche indicata la distanza della destinazione espressa in una metrica specificata. Tale metrica può essere ad esempio il numero di router da attraversare, il costo dei link, l'affidabilità dei link, il tempo di attraversamento, la massima lunghezza dei pacchetti.

E' possibile inoltre specificare, all'interno della *routing table*, instradamenti verso specifici host, e cioè corrispondenze del tipo <host destinazione, router successivo, distanza>. Questa opportunità può essere utilizzata dal Gestore a scopo di controllo e sicurezza.

Può essere previsto l'instradamento verso un *router di default* per tutti quei pacchetti IP destinati a reti non previste esplicitamente nella tabella; ciò consente di ridurne le dimensioni.

Per poter consentire l'effettiva trasmissione del pacchetto IP, questo viene incapsulato in un *frame* e indirizzato sulla base dell'*indirizzo fisico* della destinazione successiva, sia essa il router di transito o la destinazione finale. Si rende pertanto necessario far corrispondere all'indirizzo logico di rete (indirizzo IP)

(*) Ing. Ferruccio Antonelli, sig. Mauro Carissimi, ing. Francesco Iuso, ing. Francesco Pugliese -Telecom Italia DG- Roma

l'indirizzo fisico.

IP ricava questa corrispondenza invocando ARP (Address Resolution Protocol), che gestisce una apposita tabella denominata *ARP cache*. La descrizione del protocollo ARP è contenuta in [7]. L'ARP risolve le corrispondenze con un opportuno scambio di messaggi e le memorizza nell'ARP cache. E' prevista una cancellazione periodica delle informazioni contenute nell'ARP cache per garantire la coerenza delle informazioni con lo stato della rete.

Il pacchetto IP, nel processo di routing, rimane inalterato nell'attraversamento della rete, mentre nel trasferimento da un router all'altro cambia l'indirizzo fisico da inserire nell'header del frame.

Infine, al processo di routing prende parte anche l'ICMP (Internet Control Message Protocol, vd. [7]). Infatti oltre a notificare alla sorgente l'eventuale mancato recapito del pacchetto, il protocollo ICMP prevede un messaggio per la modifica delle informazioni contenute nella routing table. Qualora un pacchetto venga instradato erroneamente verso un router, quest'ultimo provvede ad inviare alla sorgente un opportuno messaggio ICMP di redirect, che modifica le informazioni della routing table.

Da quanto finora descritto emerge che il protocollo IP si avvale di due tabelle: la *routing table* e l'ARP cache.

Vi sono due modi per la loro gestione:

- la *gestione statica*: prevede che le tabelle siano costruite e gestite dal system manager mediante operazioni di management, effettuate manualmente alla console di gestione;
- la *gestione dinamica*: prevede che appositi protocolli, mediante opportuno scambio di informazioni, effettuino l'aggiornamento dinamico delle tabelle.

Di seguito vengono analizzati i protocolli per l'aggiornamento dinamico delle routing table.

2. Caratteristiche dei protocolli di instradamento

Come descritto in precedenza le informazioni di instradamento sono contenute all'interno delle tabelle di routing, la cui gestione può essere statica oppure dinamica. La gestione dinamica da un lato garantisce che le informazioni risultino in ogni momento coerenti con la realtà della rete, dall'altro essa si rende necessaria allorché le dimensioni della rete rendono la gestione statica delle tabelle eccessivamente onerosa.

In questo caso si instaura uno scambio di informazioni secondo un opportuno protocollo (protocollo di routing) in modo che, con opportuni algoritmi, le informazioni contenute nella tabella di instradamento sono aggiornate, registrando eventuali variazioni della topologia della rete.

Gli algoritmi utilizzati per effettuare queste operazioni sono gestiti in maniera distribuita con la cooperazione di tutti i router e posseggono in vario grado le seguenti caratteristiche:

- forniscono il percorso ottimo;
- introducono il minor overhead possibile;
- costituiscono un elemento di alta affidabilità;
- forniscono una soluzione nel minor tempo possibile;
- si adattano rapidamente agli imprevisti che si possono verificare durante il funzionamento della rete come i guasti o le modifiche della rete.

Gli algoritmi possono essere molto complessi dal momento che devono essere elaborate le informazioni riguardanti lo stato del nodo ed i messaggi provenienti dagli altri nodi della rete, con l'obiettivo di minimizzare i disallineamenti transitori che si creano tra le tabelle dei diversi nodi e tra queste e la situazione reale della rete, in seguito a malfunzionamenti od inserimenti e disinserimenti di nodi o linee.

Esistono numerosi metodi differenti di instradamento dinamico, ciascuno dei quali rappresenta un compromesso tra la potenza elaborativa, la quantità di memoria impiegata, la rapidità di convergenza.

Questi metodi prevedono tutti lo scambio delle seguenti tipologie di messaggio:

- *hello*: messaggio di presentazione di un nuovo nodo in rete mediante il quale il nodo comunica ai nodi vicini la propria presenza ed innesca con essi una fase di scambio di dati;
- *route updating*: messaggio mediante il quale un nodo comunica a quelli vicini l'acquisizione di nuove possibilità di raggiungere destinazioni o la perdita di possibilità di raggiungere destinazioni;
- *keep alive*: messaggio scambiato fra i nodi su base periodica per confermare gli uni agli altri la propria presenza (l'intervallo può essere dell'ordine dei secondi o delle decine di secondi in dipendenza dei casi).

Un difetto di cui soffrono in misura maggiore o minore tutti i metodi dinamici di instradamento adattativo sono i "*routing loop*".

I *routing loop* sarebbero evitati se tutte le tabelle di instradamento dei nodi di rete fossero aggiornate nello stesso istante. Questa prestazione però non è possibile per due ordini di motivi, sempre presenti negli algoritmi distribuiti: il primo consiste nell'inevitabile ritardo di trasferimento delle informazioni (tempi di trasmissione e di propagazione dei segnali), utilizzate nella fattispecie per l'aggiornamento delle tabelle di instradamento. Il secondo è legato alla velocità di convergenza dei metodi dinamici⁽¹⁾ di aggiornamento delle tabelle di instradamento verso una visione della rete consistente e rispondente alla reale connettività, ovvero al ritardo di propagazione delle informazioni di instradamento attraverso la rete. Nel caso in esame delle reti facenti uso dei protocolli TCP/IP, i valori tipici delle distanze internodali e dei ritmi binari delle linee trasmissive sono tali da rendere il primo contributo (dell'ordine dei ms)

(1) Cioè la velocità dell'algoritmo di aggiornamento delle tabelle.

senz'altro trascurabile rispetto al secondo (dell'ordine delle decine di s).

Allo scopo di contrastare l'insorgere di loop si adottano particolari regole, valide per tutti i metodi, e di queste si citano le più importanti:

- non si inviano messaggi, ad eccezione dei *keep alive*, dalla stessa parte da cui sono state ricevute le informazioni servite per aggiornare la tabella di instradamento;
- in seguito alla verifica di un guasto di rete, si dà notizia immediata ai nodi di rete adiacenti, in modo da far propagare l'informazione velocemente tra i nodi di rete in quanto in questo caso non sono richieste elaborazioni;
- l'informazione di irraggiungibilità di una rete è ignorata per un certo intervallo di tempo al fine di attendere che essa si propaghi attraverso la rete.

Le metriche di solito impiegate per valutare i costi degli instradamenti sono di seguito elencati:

- affidabilità dei link,
- tempo di transito dei nodi,
- larghezza di banda dei link,
- massima lunghezza dei pacchetti,
- costi dei link,
- numero dei nodi da attraversare.

Alcuni metodi di instradamento adattativo usano una sola metrica; altri più complessi usano metriche multiple e combinate. I vari metodi danno risultati pratici diversi tra loro, soprattutto per le metriche impiegate e/o per la possibilità di assegnarne arbitrariamente altre.

L'algoritmo eseguito sceglie quindi il cammino più "conveniente", oppure più "opportuno", oppure "ottimo", oppure a "distanza minima", oppure a "costo minimo" (tutti i termini sono equivalenti in questo contesto) in dipendenza della metrica o combinazione di metriche che in esso sono state inserite.

I metodi di instradamento adattativo non solo definiscono le metriche e la loro pesatura in una metrica combinata, ma definiscono anche la frequenza di scambio delle informazioni di instradamento, le modalità di comportamento dei manutentori nello scambio di informazioni ed il protocollo che precisa la dimensione, il contenuto ed il formato dei vari messaggi.

I protocolli per la gestione dinamica delle tabelle di instradamento possono essere classificati sulla base del tipo di algoritmo che implementano:

- distance-vector;
- link state/shortest path first (SPF).

Secondo i protocolli del tipo *distance-vector*, ciascun nodo della rete mantiene nella propria tabella, per ogni rete raggiungibile, l'indirizzo IP del nodo successivo e la distanza misurata in nodi da attraversare o in ritardo di attraversamento per raggiungere la destinazione finale.

Periodicamente le tabelle sono scambiate tra nodi adiacenti di modo che ciascuno di essi possa effettuare i dovuti aggiornamenti sulla base delle informazioni

contenute nelle tabelle dei nodi adiacenti: sono così aggiunti instradamenti verso nuove reti, sono aggiornate le informazioni di instradamento preesistenti, sono eliminate le informazioni non utilizzate da molto tempo.

I protocolli di routing di questo tipo sono caratterizzati da una estrema semplicità realizzativa, ma da una convergenza lenta. Essi si adattano bene ad ambienti statici e non eccessivamente estesi. All'aumentare delle dimensioni della rete lo scambio e la gestione delle tabelle richiede una quantità non indifferente di risorse. *La scalabilità dei protocolli basati su questo tipo di algoritmo è bassa.*

Per ovviare a quanto ora descritto, sono stati studiati i protocolli di routing di tipo *Shortest Path First (SPF)*. Questi prevedono che ogni nodo posseda una conoscenza completa della topologia della rete. In particolare la rete è rappresentata con un grafo pesato, i cui nodi rappresentano i router ed i rami rappresentano le reti che interconnettono i router. A questo scopo viene prevista:

- 1) la verifica dello stato delle connessioni verso i router adiacenti,
- 2) l'invio periodico delle informazioni sullo stato delle connessioni;
- 3) l'aggiornamento del grafo rappresentante la rete sulla base delle informazioni ricevute.

In particolare se lo stato di un link cambia, allora sono rielaborati gli instradamenti utilizzando l'algoritmo *Dijkstra (Shortest Path Algorithm)*.

I vantaggi possono essere così puntualizzati:

- non c'è dipendenza dall'elaborazione di macchine intermedie;
- l'elaborazione è locale (nel nodo) e ne è perciò garantita la convergenza;
- il messaggio sullo stato del link si propaga inalterato, quindi possono essere rilevati errori;
- la lunghezza del messaggio è indipendente dalla dimensione della rete.

La classificazione dei protocolli di instradamento non è effettuata solo in base al tipo di algoritmo, ma anche in base al loro impiego all'interno della struttura organizzativa gerarchica dei router. Per chiarire quest'ultima affermazione è necessario introdurre il concetto di *Autonomous System (AS)*.

Per ragioni tecniche, gestionali e di opportunità pratica, i gateway della rete sono organizzati in insiemi chiamati "Autonomous System" (AS), come mostrato in fig. 1.

I gateway all'interno di un AS agiscono in maniera autonoma dai gateway all'esterno, nel senso che sono liberi di organizzarsi come meglio credono per quello che riguarda lo scambio delle informazioni di routing. All'interno dell'AS è previsto un gateway delegato a raccogliere le informazioni di instradamento riguardanti il proprio AS ed a comunicarle ad analoghi gateway di altri AS mediante un protocollo standard. La suddivisione

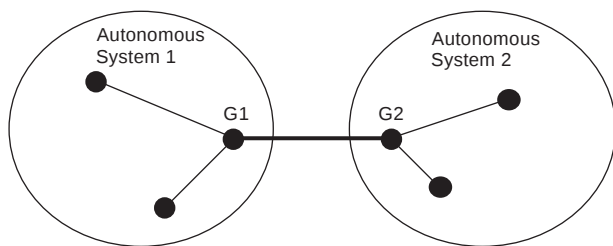


Figura 1 Autonomous Systems

in AS permette inoltre:

- la manutenzione ed il controllo di tutti i gateway costituenti il medesimo AS da parte di una singola organizzazione O&M;
- l'impiego di un protocollo di routing comune tra i gateway dell'AS, indipendentemente dalle scelte fatte per gli altri AS.

Sono pertanto individuati due tipi di protocollo: i primi impiegati all'interno degli AS, i secondi tra AS.

Per quanto riguarda il routing intra-AS sono state effettuate diverse scelte implementative, che hanno portato allo sviluppo di una serie di protocolli proprietari ai quali si fa generalmente riferimento indicandoli come "Interior Gateway Protocol".

Per lo scambio di informazioni topologiche inter-AS è previsto invece l'uso di un diverso tipo di protocollo di instradamento appartenenti alla famiglia dei protocolli denominata EGP "Exterior Gateway Protocol". Il primo protocollo standard appartenente a questa famiglia è l'omonimo protocollo EGP, ma è stato poi sviluppato un nuovo protocollo denominato BGP (Border Gateway Protocol).

Di seguito sono descritti i principali protocolli inter e intra AS.

2.1 Protocolli di routing intra-AS (IGP: Interior Gateway Protocol)

2.1.1 RIP

RIP (Routing Internet Protocol) è un protocollo di tipo vector-distance appartenente alla famiglia IGP per l'aggiornamento dinamico delle tabelle di instradamento che utilizza i servizi offerti dal livello UDP (User Datagram Protocol) per lo scambio di messaggi contenenti informazioni di instradamento.

Ciascun gateway può richiedere informazioni di instradamento inviando un messaggio di richiesta. Alla ricezione di una richiesta il protocollo prevede l'invio di un messaggio di risposta con il contenuto della tabella di instradamento.

Tuttavia in molte realizzazioni si verifica che i messaggi di risposta vengano trasmessi periodicamente

(ogni 30 s) in broadcast attraverso le interfacce disponibili, senza necessità di una esplicita richiesta.

I gateway confrontano le informazioni ricevute con quelle contenute nella propria tabella di instradamento, quindi inseriscono nella tabella gli instradamenti verso reti non ancora previste e aggiornano i costi degli instradamenti già esistenti.

Le informazioni non aggiornate per un tempo superiore a 3 minuti sono rimosse.

Il RIP non prevede meccanismi per evitare i loop. Per contrastare il fenomeno dei loop sono quindi necessarie funzionalità aggiuntive: ad esempio, in alcune implementazioni le modifiche sono effettuate dopo un intervallo di tempo (60 s), in modo da consentire la propagazione delle informazioni attraverso la rete.

Il RIP utilizza il numero di gateway da attraversare come misura della distanza di una data destinazione, mentre il valore 16 indica per convenzione che la destinazione è irraggiungibile.

Per chiarire il funzionamento del RIP e per evidenziare la scarsa velocità di convergenza viene di seguito presentato un esempio. Nella fig. 2 sono mostrati tre gateway ciascuno dei quali prevede, a costi differenti, un instradamento verso la rete R1.

In particolare nella tabella di instradamento di G1, direttamente collegato alla rete R1, è indicato che questa rete è raggiungibile a costo 1. G2 apprende dallo scambio delle tabelle con G1 che la rete R1 è raggiungibile attraverso G1 a costo 2 e inserisce questa informazione nella propria tabella di instradamento. Scambiando poi le informazioni di instradamento con G3, quest'ultimo apprende che la rete R1 è raggiungibile attraverso G2 a costo 3.

Nella seconda metà della figura si fa l'ipotesi che improvvisamente G1 venga scollegato (a causa di guasti o altro) da R1 e quindi G1 provveda a mettere ad infinito (distanza pari a 16) il costo verso R1. Col successivo broadcast della tabella di instradamento G1 dovrebbe informare di questo aumento di costo il gateway G2.

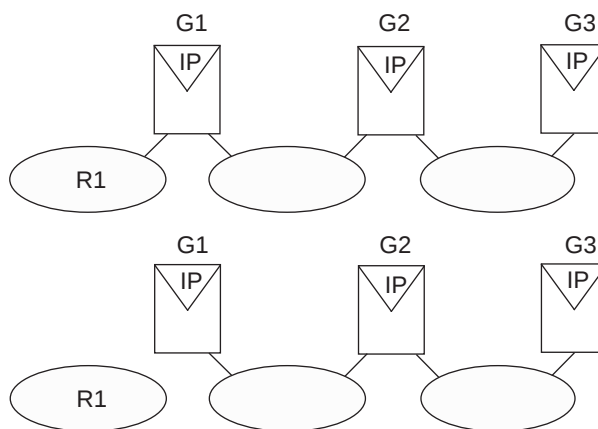


Figura 2 Esempio di convergenza del RIP

E' possibile tuttavia che G2 effettui il broadcast prima di G1 e che quindi quest'ultimo applichi l'algoritmo vector-distance sulla base delle informazioni ricevute: in particolare G1 rivelerebbe un instradamento verso R1 attraverso il gateway G2 a costo 2, inferiore del costo 16 dell'instradamento di sua conoscenza, e procederebbe ad aggiornare a 3 tale costo.

Con lo scambio successivo G2 verrebbe a conoscenza di un instradamento vero R1 attraverso G1 a costo 3 e procederebbe ad aggiornare il costo verso R1 a 4. Solo dopo varie iterazioni il costo giungerebbe a 16 e la rete R1 verrebbe dichiarata irraggiungibile dai gateway.

Questo esempio mostra come RIP non si adatti bene a contesti particolarmente dinamici e che comunque può essere utilizzato efficacemente in reti di piccole dimensioni (il numero massimo di nodi intermedi di un instradamento è limitato a 15).

I messaggi scambiati dal RIP, il formato dei quali è mostrato nella fig. 3, sono costituiti dai seguenti campi:

- Command (8 bit): specifica se il pacchetto è una richiesta o una risposta secondo lo schema:
 - 1 richiesta per informazione parziale o completa di routing;
 - 2 risposta contenente le coppie rete-distanza, tratte dalla routing table del mittente;
 - 3 attiva la modalità Trace;
 - 4 disattiva la modalità Trace;
 - 5 riservato per uso interno della SUN Microsystem.
- Version (8 bit): contiene il numero che indica la versione del protocollo;
- Family of net "i" (16 bit): identifica il protocollo di rete con cui interpretare l'indirizzo della rete "i";
- Net Address "i" (14 byte): contiene l'indirizzo di rete, la cui lunghezza può essere al massimo di 14 byte;

0	8	16	24	31
Command	Version	must be zero		
Family of net 1		must be zero		
IP Address of net 1				
must be zero				
must be zero				
Distance to net 1				
Family of net 2		must be zero		
IP Address of net 2				
must be zero				
must be zero				
Distance to net 2				
...				

Figura 3 Formato del messaggio RIP

- Distance to net "i" (32 bit): è un intero che indica quanti gateway è necessario attraversare per raggiungere la rete "i".

E' stata sviluppata una seconda versione del protocollo RIP, che prevede la gestione di sottoreti IP ed una serie di strumenti da impiegare come misure di sicurezza, nonché l'impiego del multicast al fine di alleggerire il carico elaborativo degli host che partecipano allo scambio di informazioni di routing.

2.1.2 HELLO

Anche il protocollo HELLO è di tipo distance-vector. Si differenzia dal RIP per il fatto che la distanza è misurata in ritardi di transito anziché in numero di nodi da attraversare. Il suo impiego è molto limitato: le funzioni svolte dal protocollo consistono nel sincronizzare i riferimenti temporali delle entità di routing e nel determinare il percorso più breve per raggiungere una data destinazione. Il protocollo si basa sulla stima dei riferimenti temporali delle entità adiacenti per valutare il ritardo di attraversamento di un collegamento ed individuare il percorso più breve per raggiungere una data destinazione. Anche HELLO non converge rapidamente e diviene instabile in caso di rapidi cambiamenti della configurazione di rete. Il messaggio HELLO, il cui formato è mostrato nella fig. 4, è costituito dai seguenti campi:

- Checksum (16 bit): contiene la checksum del messaggio;
 - Date (16 bit): contiene la data locale al mittente;
 - Time (32 bit): contiene l'ora secondo il riferimento del mittente;
 - Timestamp (16 bit): è utilizzato per calcolare il tempo di percorrenza;
 - Local Entry (8 bit): punta nella lista degli host per marcare il blocco di dati relativi alla rete locale;
 - #Hosts (8 bit): indica la lunghezza della lista.
- Ciascun elemento della lista contiene due campi:

0	8	16	24	31
Checksum		Date		
Time				
Timestamp		Local Entry	# Hosts	
Delay ₁		Offset ₁		
Delay ₂		Offset ₂		
...				
Delay _n		Offset _n		

Figura 4 Il formato del messaggio HELLO

- * Delay in cui è indicato il tempo per raggiungere un dato host;
- * Offset che contiene la stima fatta dal mittente della differenza tra i riferimenti temporali del mittente e dell'host.

2.1.3 OSPF

Il protocollo OSPF (*Open Shortest Path First*) è di tipo link-state, prevede lo scambio di messaggi impiegando i servizi di rete forniti da IP. E' caratterizzato da una maggiore velocità di convergenza rispetto ai precedenti protocolli IGP.

Per una assegnata destinazione il protocollo può prevedere una pluralità di instradamenti caratterizzati da differenti Quality of Service (QoS), così da poter utilizzare in modo appropriato il campo "Type of Service" contenuto nell'header del pacchetto IP. Se poi sono previsti più instradamenti caratterizzati da un medesimo QoS allora il protocollo prevede la distribuzione del traffico tra questi instradamenti. Possono essere previsti instradamenti verso specifici host, reti, sottoreti.

Il protocollo si adatta bene alla crescita della intera rete. Infatti l'insieme dei gateway può essere suddiviso in sottoinsiemi denominati *aree*. Ciascuna area non ha conoscenza della topologia interna delle altre e coopera nello scambio di informazioni di routing secondo OSPF; la modifica della topologia interna di una area non perturba le informazioni di routing all'interno delle altre aree.

OSPF prevede che lo scambio delle informazioni sia "autenticato", cioè che vengano previste esplicite misure di sicurezza, e consente massima flessibilità in quanto prevede di lavorare sulla rete virtuale IP, astraendosi dalla rete di trasporto. Le informazioni scambiate possono essere state acquisite anche da entità esterne. I messaggi scambiati dal protocollo sono di vario tipo. Essi sono costituiti da una parte comune e da una dipendente dal particolare tipo di messaggio. In fig. 5 è mostrata la struttura generale dei messaggi.

- Version (8 bit): specifica la versione del protocollo;
- Type (8 bit): definisce il tipo di messaggio scambiato;
- Message length (16 bit): indica la lunghezza del messaggio;
- Source Gateway IP Address (32 bit): contiene l'indirizzo del mittente;
- Area ID (32 bit): contiene l'identificativo dell'area;
- Checksum (16 bit): contiene il codice per la rivelazione di errori;
- Authentication Type (16 bit): specifica il meccanismo di sicurezza impiegato;
- Authentication (32 bit): contiene la parola di codice per il sistema di sicurezza.

I tipi di messaggio impiegati sono:

- *hello* (Type=1) è scambiato periodicamente per

0	8	16	24	31
Version	Type	Message Length		
Source Gateway IP Address				
Area ID				
Checksum		Authentication Type		
Authentication (byte 0-3)				
Authentication (byte 4-7)				
parte dipendente dal tipo di messaggio				

Figura 5 Struttura generale dei messaggi scambiati da OSPF

verificare la raggiungibilità delle entità adiacenti.

- *OSPF Database Description* (Type=2) è scambiato per inizializzare il database contenente la descrizione topologica della rete;
- *OSPF Link Status Request* (Type=3): è impiegato per richiedere alle entità adiacenti informazioni per aggiornare il database. In questo tipo di messaggio è possibile specificare una lista di link di interesse. Al messaggio di richiesta l'entità interlocutrice replica con un messaggio di risposta in base alle informazioni contenute nel proprio database.
- *OSPF Link status Update* (Type=4): è inviato in broadcast dai gateway per richiedere l'aggiornamento dello stato dei link.

2.2 Protocolli di routing inter-AS (EGP: Exterior Gateway Protocol)

2.2.1 EGP

EGP (*Exterior Gateway Protocol*) è il protocollo standard impiegato per lo scambio di informazioni di instradamento tra AS. In particolare è il protocollo con il quale i gateway G1 e G2 (cfr. fig. 1), delegati a raccogliere le informazioni di instradamento dei rispettivi AS di appartenenza, si scambiano informazioni di routing.

L'EGP svolge tre funzioni: 1) individuazione dei gateway adiacenti con cui scambiare le informazioni di routing; con riferimento alla fig. 1 i gateway G1 e G2 si accordano per scambiare informazioni secondo il protocollo EGP; per adiacenza si intende in questo caso quella logica: i gateway delegati allo scambio di informazioni possono infatti essere separati anche da altri gateway; 2) verifica continua della funzionalità dei gateway interlocutori; 3) scambio periodico delle informazioni di instradamento, contenute in appositi messaggi; queste riguardano solo la raggiungibilità delle reti, non la distanza.

Per svolgere queste funzioni il protocollo prevede lo scambio di vari tipi di messaggi (tab. 1). I messaggi hanno tutti un comune header così costituito:

- **Version (8 bit):** contiene un intero che identifica la versione del protocollo EGP. Il Gateway che riceve il messaggio verifica che venga impiegata la stessa versione del software che implementa il protocollo;
- **Type (8 bit):** identifica, insieme al campo successivo, il tipo di messaggio;
- **Code (8 bit):** distingue tra i diversi sottotipi nell'ambito dei vari tipi di messaggio;
- **Status (8 bit):** questo campo contiene informazioni sullo stato, in dipendenza dal messaggio;
- **Checksum (16 bit):** il protocollo impiega questo campo per verificare l'integrità del messaggio. E' impiegato lo stesso algoritmo dell'IP, consistente nel considerare il messaggio formato da una sequenza di interi di 16 bit e nel calcolarne la somma in complemento ad uno;
- **Autonomous System Number (16 bit):** questo campo contiene il numero assegnato all'AS di cui fa parte il gateway mittente;
- **Sequence Number (16 bit):** contiene un numero che gli interlocutori impiegano per sincronizzare le risposte con le rispettive richieste.

I messaggi possono essere suddivisi in 4 tipi distinti:

- **acquisizione dell'interlocutore vicino:** inizialmente i gateway, che scambiano informazioni secondo il protocollo EGP, stabiliscono una comunicazione EGP inviando un messaggio di richiesta di comunicazione (*Acquisition Request*) al quale

l'interlocutore può rispondere con un messaggio di conferma o di rifiuto (*Acquisition Confirm/Refuse*). Per interrompere lo scambio di informazioni viene impiegato il messaggio di richiesta di svincolo (*Cease Request*) a cui viene replicato con un messaggio di conferma (*Cease Confirm*). Il protocollo non specifica il modo in cui un gateway individua il corrispondente con cui intraprendere la procedura ora descritta: questo aspetto dovrà perciò essere curato dal system manager;

- **raggiungibilità dell'interlocutore:** per verificare che l'interlocutore vicino è attivo, il gateway invia periodicamente un messaggio (*Hello*) al quale il gateway interlocutore replica con (*I Heard You*) per confermare la sua funzionalità;
- **poll:** sono previsti i messaggi *poll request* e *poll response* rispettivamente per richiedere e per ottenere informazioni circa la raggiungibilità delle reti;
- **aggiornamento degli instradamenti:** per l'aggiornamento delle informazioni di routing è impiegato il messaggio *routing update*. In questo messaggio il gateway mittente riporta le informazioni sulle reti raggiungibili presenti nel proprio AS.

Nella fig. 6 è riportata la struttura generale dei messaggi scambiati.

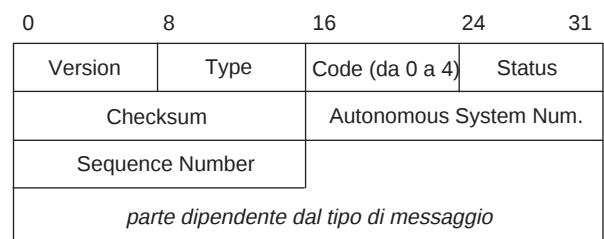


Figura 6 Struttura generale dei messaggi EGP

Messaggi EGP	Descrizione
<i>Acquisition Request</i>	richiesta di diventare un gateway pari per EGP
<i>Acquisition Confirm</i>	risposta positiva alla Acquisition Request
<i>Acquisition Refuse</i>	risposta negativa alla Acquisition Request
<i>Cease Request</i>	richiesta di termine del rapporto tra pari
<i>Cease Confirm</i>	risposta di conferma alla Cease Request
<i>Hello</i>	richiesta per verificare che il pari sia attivo
<i>I Heard You</i>	risposta al messaggio Hello
<i>Poll Request</i>	richiesta di aggiornamento degli instradamenti
<i>Routing Update</i>	informazioni sulla raggiungibilità delle reti
<i>Error</i>	risposta a un messaggio non corretto

Tabella 1 Tipi di messaggi scambiati dal protocollo EGP

2.2.2 BGP

Il protocollo BGP (*Border Gateway Protocol*) è di tipo distance-vector e può essere considerato una evoluzione dell'EGP, in quanto alle funzionalità di base aggiunge lo scambio di informazioni sulla raggiungibilità delle reti. Per lo scambio di messaggi tra le entità sono utilizzati i servizi di trasporto offerti dal Transport Control Protocol (TCP). Il protocollo effettua la verifica dello stato di un link o di un host inviando periodicamente messaggi di *keepalive* (il periodo raccomandato è di 30 s).

Le informazioni di raggiungibilità includono il percorso completo di AS che il traffico deve seguire per raggiungere determinate reti di destinazione. Queste informazioni sono sufficienti a costruire il grafo di connettività degli AS; eventuali *loop* possono essere poi rivelati ed eliminati con opportune politiche di gestione. L'informativa sulla raggiungibilità delle reti ha

naturalmente maggior impatto nel caso in cui sia presente, lungo il percorso tra l'AS sorgente e quello di destinazione, un numero di AS con funzioni di transito; in caso contrario l'impiego di BGP od EGP è pressoché equivalente.

Infatti il protocollo BGP consente di utilizzare opportune politiche di restrizione per il traffico in transito, stabilite dal gestore del sistema e codificate in file di configurazione, che sono poi trasferiti al protocollo. Ciò consente al protocollo di scegliere tra più instradamenti disponibili e di effettuare la ridistribuzione del traffico.

3. Conclusioni

I protocolli di routing descritti in questo documento sono stati sviluppati sin dall'avvio della rete Internet e della pila di protocolli utilizzata per il suo funzionamento, TCP/IP, ed hanno subito varie evoluzioni man mano che sono state superate frontiere di complessità e dimensioni crescenti.

E' evidente peraltro una perentoria evoluzione anche di questa tecnologia sulla spinta dello sviluppo e della realizzazione di nuove forme e tecniche di comunicazione (in primis l'introduzione della tecnica ATM).

In tale contesto sono stati avviati nuovi studi e realizzazioni prototipali di protocolli di routing in grado di utilizzare meglio (e quindi in modo non del tutto indipendente) i mezzi di comunicazione su cui si poggia la pila TCP/IP.

Tra questi ricordiamo: NHRP - Next Hop Routing Protocol -, la cui caratteristica peculiare è quella di introdurre il concetto di server di routing, in cui l'instradamento è gestito in modo concentrato; i protocolli di routing per la gestione dei traffici multicast, flussi di dati cioè indirizzati ad un gruppo di destinazioni distinte; infine i protocolli di routing per il supporto di traffici di tipo multimediale, attraverso la prenotazione e l'assegnazione delle risorse di rete e la definizione di classi di priorità per i diversi tipi di segmenti di informazione.

Questa evoluzione, ancora oggi in fase di definizione e di normalizzazione, formerà l'oggetto di un prossimo articolo.

Bibliografia

- [1] Comer, D.: *Internetworking with TCP/IP, Volume I - Principles, Protocols and Architecture*. Prentice Hall, 1991.
- [2] Stevens, W.R.: *TCP/IP Illustrated, Volume 1, The Protocols*. Addison Wesley, 1994.
- [3] Comer, D.; Stevens, D.: *Internetworking with TCP/IP,*

Volume II - Design, Implementtion and Internals. Prentice Hall, 1994.

- [4] Comer, D.; Stevens, D.: *Internetworking with TCP/IP, Volume III - Client-Server Programming and Applications for the AT&T TLI Version*. Prentice Hall, 1994.
- [5] Comer, D.; Stevens, D.: *Internetworking with TCP/IP, Volume III - Client-Server Programming and Applications for the BSD Socker Version*. Prentice Hall, 1993.
- [6] Perretti, E.; Antonelli, F.; Iuso, F.; Montechiarini, C.; Pugliese, F.; Carissimi, M. et al.: *"Progetto IP" Studio di fattibilità del servizio IP su rete pubblica e sperimentazione tecnologica*. Telecom Italia, Settembre 1994.
- [7] Antonelli, F.; Carissimi, M.; Iuso, F.; Pugliese, F.: *I protocolli TCP ed IP*. «Notiziario Tecnico Telecom Italia», Vol. 4, n. 1, Luglio 1995, pp. 50-62.

Indirizzamento globale nelle reti di telecomunicazioni

L. Barbetta, G. Martini (*)

I moderni servizi di telecomunicazioni, in un contesto di crescente deregolamentazione, pongono in evidenza la necessità di far evolvere le tecniche di indirizzamento tradizionali verso l'utilizzo di una struttura di indirizzamento evoluta che, anche basandosi sui piani di numerazione pubblici, risulti indipendente dall'infrastruttura di rete specifica ed includa informazioni di indirizzamento utilizzabili anche in ambito privato. La tecnica che soddisfa tali esigenze è indicata con il termine "indirizzamento globale".

In questo articolo sono descritti, in una prima parte gli aspetti generali, teorici ed amministrativi, della normativa di pertinenza. Nella seconda sono analizzati gli aspetti di registrazione legati alla gestione ed all'assegnazione dello schema di indirizzamento globale individuato. Sono infine sottolineati gli aspetti strategici e di mercato legati a tale struttura di indirizzamento.

1. Introduzione

L'evoluzione delle reti e dei servizi, dovuto da un lato alla continua innovazione tecnologica e dall'altro al progressivo aumento delle esigenze di connettività da parte dell'utente, comporta la necessità di far evolvere il ruolo delle tecniche di indirizzamento tradizionali.

Tradizionalmente a causa della stretta associazione rete-servizio i piani di numerazione pubblici risultavano strettamente legati ad una particolare infrastruttura di rete: il piano di numerazione E.164 [1] alla rete telefonica/ISDN, quello X.121 [2] alla rete dati e l'F.69 [3] alla rete Telex. In questi casi la struttura del numero ed il significato delle sue cifre rispondono ai requisiti ed ai vincoli tecnologici dell'infrastruttura specifica.

Inoltre, tali forme di indirizzamento risultano limitate all'individuazione del punto di attacco utente-rete e non forniscono alcun meccanismo utilizzabile all'interno dei sistemi d'utente in scenari di interlavoro tra reti pubbliche e private.

I moderni servizi di TLC, che utilizzano infrastrutture di rete dati, richiedono sempre più una struttura di indirizzamento che, anche basandosi sui

piani di numerazione pubblici, risulti indipendente dall'infrastruttura di rete specifica ed includa informazioni di indirizzamento utilizzabili anche all'interno del dominio privato. La tecnica che soddisfa tale esigenze è indicata con il termine "indirizzamento globale".

La struttura ed il significato di una struttura di indirizzamento globale deve essere tale da realizzare l'instradamento in uno scenario di interlavoro tra differenti reti e deve essere elaborato da un opportuno protocollo indipendente dai livelli di trasporto dell'infrastruttura specifica.

Questi requisiti costituiscono la base del modello e dei protocolli specificati in ISO ed ITU relativamente all'OSI, Open System Interconnection [4].

Al suo interno è stato standardizzato un sistema di indirizzamento, noto come indirizzo N-SAP (Network Service Access Point), strutturato con l'obiettivo di assicurare la coesistenza dei differenti piani e schemi di numerazione pubblici e privati esistenti. Il dominio di appartenenza della struttura N-SAP è totalmente separato da quelli dei piani di numerazione utilizzati nelle reti esistenti (pubbliche e private, telefoniche e dati) ed è quindi capace di coesistere con questi. Ciò consente l'interoperabilità di sistemi attestati a reti diverse, sia pubbliche sia private, per la realizzazione di servizi di telecomunicazione comuni.

(*) Ing. Luca Barbetta -CSELT- Torino; ing. Giovanni. Martini - Telecom Italia DG- Roma

L'indirizzo N-SAP è oggi in fase di studio in vari ambiti di specifica, tra i quali Organismi finalizzati alla creazione di accordi per pervenire ad una comune implementazione (es. l'ATM Forum o la Comunità Internet); appare quindi opportuno procedere alla descrizione delle interpretazioni delle strutture emergenti in tali ambiti.

Per l'interpretazione delle caratteristiche e degli scopi dell'indirizzo N-SAP è opportuno conoscere in via preliminare la normativa, almeno limitatamente alla determinazione degli aspetti teorici ed amministrativi. La prima parte di questo articolo è quindi orientata principalmente ad una descrizione degli standard di pertinenza, in modo da fornire una chiave di lettura generale della tematica.

La seconda parte del documento analizza gli aspetti di registrazione legati alla gestione ed assegnazione dello schema di indirizzamento globale N-SAP.

Sono infine sottolineati gli aspetti strategici e di mercato legati a questa struttura di indirizzamento: infatti, le forme di indirizzamento evolute sono tendenzialmente aperte a gestire servizi che possono essere forniti sia da Gestori pubblici sia da Operatori privati. Questa opportunità, in un contesto di crescente liberalizzazione dei servizi di telecomunicazioni, incoraggiato dalla Comunità Europea, evidenzia la valenza strategica e commerciale della tematica.

2. Strutture e aspetti normativi

L'indirizzo N-SAP nasce come forma di indirizzamento OSI utilizzata per individuare, al di sopra delle funzionalità di trasporto, punti di accesso omogenei (Network Service Access Point, N-SAP, secondo la terminologia OSI) per i servizi o per gli applicativi d'utente [4]. In fig. 1 è illustrato l'approccio OSI alla realizzazione di un servizio di trasporto indipendente dalle specifiche risorse di trasmissione e di rete utilizzate.

La rete di telecomunicazione indicata in fig. 1 è in realtà costituita da un insieme di diverse infrastrutture di rete opportunamente interconnesse, ciascuna delle quali capace di prevedere meccanismi indipendenti per l'attivazione delle risorse di trasporto. Il protocollo del livello "rete", indicato in fig. 1, è omogeneo per tutti i sistemi d'utente e fa riferimento allo standard ISO 8348 [5].

Data la sua natura globale l'indirizzamento N-SAP è un meccanismo utilizzabile in rete in tutti quei casi in cui devono essere risolte problematiche di interlavoro tra piani/schemi di indirizzamento corrispondenti a

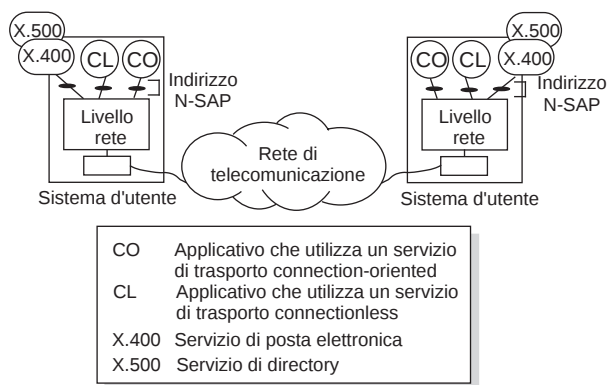


Figura 1 Ruolo dell'N-SAP e dell'indirizzo N-SAP

differenti infrastrutture di rete o, secondo la terminologia OSI, sottoreti⁽¹⁾. Infatti l'N-SAP può essere utilizzato per ricavare direttamente o indirettamente le informazioni necessarie per la realizzazione del collegamento o del servizio desiderato in uno scenario di interlavoro generico.

L'indirizzo N-SAP è in realtà formato da un insieme di schemi di indirizzamento indipendenti fra loro e coesistenti all'interno del dominio ISO/ITU degli indirizzi di rete. Questa caratteristica è alla base della sua globalità, in quanto risulta la naturale estensione delle capacità di indirizzamento dei piani tradizionali di numerazione pubblici ITU (E.164, X.121 e F.69) e di strutture di indirizzamento o identificazione ISO/IEEE utilizzate prevalentemente in ambito privato (ad esempio gli indirizzi LAN IEEE 802 a 48 bit [6]).

In fig. 2 è presentata la struttura dell'indirizzo N-SAP come definita congiuntamente da ISO e ITU [5]. Il campo IDP (Initial Domain Part) ha, mediante il sottocampo AFI (Authority and Format Identifier), la funzione di identificare l'autorità amministratrice tra ISO e ITU e, mediante il sottocampo IDI (Initial Domain Identifier), il formato specifico e l'autorità che gestisce la parte successiva dell'indirizzo (sottocampo DSP - Domain Specific Part).

I formati di N-SAP disponibili sono riassunti in tab.1 e devono essere interpretati secondo le seguenti modalità:

- formati relativi ai piani di numerazione pubblici ITU, cioè il campo IDI contiene un numero appartenente al piano E.164, X.121 o F.69. In questo caso l'autorità

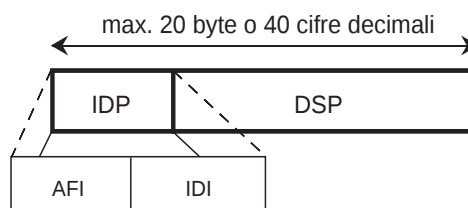


Figura 2 Formato dell'indirizzo N-SAP

(1) Nel seguito si utilizzerà il termine sottorete per riferirsi alle infrastrutture di rete esistenti (ad esempio la rete telefonica, una rete dati X.25, una LAN), che sono unicamente capaci di fornire una connettività limitata all'area geografica coperta dall'infrastruttura stessa.

di amministrazione di più alto livello gerarchico è l'ITU, poi agisce l'autorità nazionale che assegna i numeri dei piani di numerazione pubblici ed all'ultimo livello gerarchico l'ente o l'organizzazione individuata dal numero pubblico stesso. Di conseguenza il campo DSP è amministrato dall'utente della rete pubblica;

- formato ISO DCC (Data Country Code): l'ISO in passato ha definito nello Standard Internazionale 3166 [7] uno schema di identificazione, detto Destination Country Code (DCC), delle varie nazioni parallelamente al Country Code telefonico/ISDN (ad es. all'Italia è assegnato il valore 380). Questo valore di DCC individua il membro ISO principale all'interno di ciascuna nazione (ad es. l'UNINFO per l'Italia) che assume il ruolo di responsabile della gestione ed assegnazione degli indirizzi;
- formato ISO ICD (International Code Designer): utilizza un altro standard ISO, precisamente lo Standard Internazionale 6523 [8], che ha lo scopo di identificare organizzazioni su base mondiale senza associazioni geografiche. L'assegnazione di un valore ICD può essere richiesto al BSI (British Standard Institute), Organismo demandato dall'ISO come unica Autorità di registrazione, da parte di una qualsiasi organizzazione pubblica o privata previa sponsorizzazione da parte di un membro ISO. L'organizzazione, identificata dal valore di ICD, assume il ruolo di autorità responsabile per la strutturazione ed assegnazione della parte restante dell'indirizzo;
- formato locale a disposizione per accomodare forme di indirizzamento non OSI (proprietarie) nelle fasi iniziali di transizione verso l'introduzione globale dell'indirizzo N-SAP. Non è identificata alcuna Autorità amministratrice ed il suo significato è unicamente locale.

In tab. 1 sono riassunti i possibili formati dell'indirizzo N-SAP ed in ciascun caso l'autorità amministratrice del dominio, l'autorità che ha in carico

DOMINIO	AUTORITÀ SUL DOMINIO	AUTORITÀ SUL CAMPO DSP
X.121	ITU	Utente X.121
ISO DCC	ISO	Entità ISO Nazionale
F.69	ITU	Utente F.69
E.164	ITU	Utente E.164
ISO 6523-ICD	ISO	Organizzazione ICD
Locale	-----	Locale

Tabella 1 Formati ed autorità amministratrici

la strutturazione ed assegnazione dei valori del campo DSP e la lunghezza massima del campo DSP stesso.

In tab. 2 sono riportati i gruppi che attualmente negli organismi di standardizzazione ISO e ITU

Ente di Normativa	Gruppo di lavoro
ISO - JTC1	SC6 Working Group 2
ITU - TS Study Group 7	Working Party 5 Questione 23 Questione 19
ITU - TS Study Group 2	Working Party 1 Questione 5

Tabella 2 Gruppi di studio ISO e ITU rilevanti per la tematica N-SAP

hanno la responsabilità dell'attività di studio sull'indirizzamento N-SAP.

3. Specifiche per l'utilizzo dell'indirizzo N-SAP

In questo paragrafo sono illustrate una serie di normative che, basandosi sullo standard ISO/ITU presentato in precedenza, affrontano la problematica del supporto dell'indirizzo N-SAP nelle reti pubbliche esistenti. Inoltre sono presentate altre specifiche con requisiti particolari sulla struttura dell'N-SAP per il suo trattamento nei prodotti oggi in fase di definizione.

3.1 N-SAP e reti ISDN e PDN

L'utilizzo dell'indirizzo N-SAP in ISDN prevede unicamente il trasporto del segnale verso il sistema d'utente destinatario, in quanto tale indirizzo non contiene informazioni utili per l'instradamento all'interno della rete pubblica. Nella Raccomandazione ITU I.334 [9] sono illustrate tutte le possibili allocazioni degli schemi di indirizzamento N-SAP all'interno dei campi di indirizzamento disponibili nella segnalazione ISDN. L'utilizzo del meccanismo di sottoindirizzamento (Subaddressing) è quello che presenta la maggiore flessibilità e che permette appieno l'estensione delle capacità di indirizzamento fornite dall'indirizzo N-SAP. In fig. 3 è mostrata la struttura completa dell'indirizzo ISDN.

Nel caso delle sottoreti dati basate sul protocollo X.25 [10] il supporto dell'indirizzamento N-SAP è realizzato tradizionalmente mediante il suo inserimento in un campo di "Facility" all'interno del messaggio per

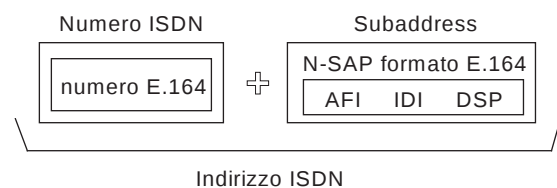


Figura 3 Supporto dell'indirizzo N-SAP formato E.164 per ISDN

la instaurazione della connessione (è un approccio molto simile al "Subaddress" ISDN).

In fig. 4 è presentata la struttura completa dell'indirizzo su Public Data Network (PDN) all'interno dei campi di indirizzamento del pacchetto di "Call Request" X.25.

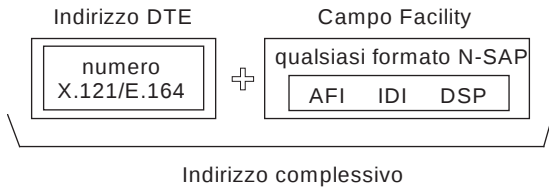


Figura 4 Supporto dell'indirizzo N-SAP per la PDN

3.2 Struttura e interpretazione dell'indirizzo N-SAP in ISO

Una volta definita la struttura dell'indirizzo N-SAP, l'ISO ha rilevato la necessità di fornire alcuni criteri per la strutturazione del campo DSP. In questo modo è possibile un effettivo utilizzo dell'indirizzo per derivare le informazioni necessarie per l'instradamento globale in scenari generali di interlavoro.

L'ISO ha ormai definito due standard relativi ai protocolli per l'instradamento in uno scenario di interlavoro tra reti [11, 12] con ulteriori requisiti, illustrati in fig. 5, sulla struttura dell'indirizzo N-SAP.

Tale struttura rispecchia una strategia di instradamento nella quale si individua un sistema d'utente e la sottorete a cui è attestato, localizzandoli in specifici domini di instradamento, identificati mediante la prima parte del campo DSP (Indirizzo di Area in fig. 5). La metodologia ISO di instradamento globale è descritta più in dettaglio nel § 4, ovvero nell'ambito degli aspetti evolutivi legati alla tematica N-SAP.



Figura 5 Struttura dell'indirizzo per l'instradamento ISO

3.3 Specifiche GOSIP (Government OSI Profile)

Taluni Governi ed Amministrazioni hanno realizzato delle specifiche nazionali, i cosiddetti GOSIP, relativi all'utilizzo dei protocolli OSI e quindi alla struttura degli indirizzi N-SAP.

3.3.1 Specifiche N-SAP in Gran Bretagna

In U.K. si è scelto di standardizzare il formato ISO DCC, cioè quello che presenta una connotazione geografica gestita dal membro ISO di riferimento nazionale. Di solito il membro ISO di riferimento nazionale coincide con l'entità di regolamentazione nazionale e quindi, nel caso britannico, con il BSI. Il BSI a sua volta ha investito del ruolo di autorità di gestione di tale schema l'ente EEA (Electronics and Business Equipment Association).

L'obiettivo principale della normativa britannica è la definizione di una struttura di N-SAP che assicuri principalmente la correttezza delle problematiche amministrative.

In fig. 6 è presentata la struttura dell'indirizzo N-SAP. All'interno della gerarchia definita dall'ISO, il BSI ha definito un'ulteriore gerarchia amministrativa nazionale al cui vertice ha posto l'ente EEA, che quindi ha la responsabilità di amministrare il campo UKDP. I valori di tale campo identificano un'Organizzazione o Ente che ha la responsabilità di strutturare ed assegnare il restante campo UKDSP. In tab. 3 sono illustrate le possibili strutture del campo UKDP.

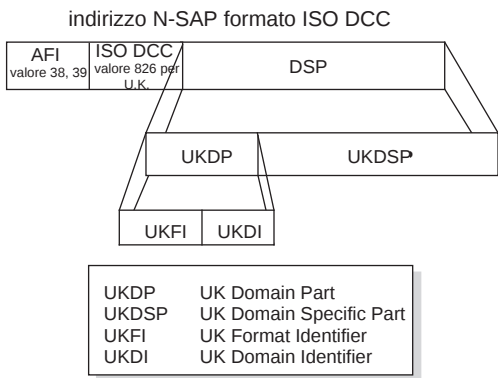


Figura 6 Struttura dell'indirizzo N-SAP formato ISO DCC per l'U.K.

Valore del sottocampo UKFI	Struttura dei sottocampi UKDI e UKDSP	Utilizzo
0, 4 ÷ 9	-----	riservati per usi futuri
UKFI =1	UKDI a 3 cifre UKDSP al massimo di 31 cifre decimali	da assegnare a organizzazioni di rilevanti dimensioni
UKFI =2	UKDI a 5 cifre UKDSP al massimo di 29 cifre decimali	da assegnare a organizzazioni di dimensioni medie
UKFI =3	UKDI a 8 cifre UKDSP al massimo di 26 cifre decimali	da assegnare a organizzazioni di ridotte dimensioni o al singolo utente

Tabella 3 Significato ed utilizzo dei campi UKDP e UKDSP

3.3.2 Specifiche N-SAP in U.S.A.

In questo caso per l'utilizzo all'interno del GOSIP americano, quindi principalmente per le agenzie e gli enti governativi americani, è stato scelto il formato ISO ICD (valore "0005"), cioè quello privo di un'esplicita connotazione nazionale (si veda la fig. 7).

L'autorità responsabile di tale dominio di indirizzi è il NIST (National Institute of Standards and Technology), Ente di normativa americano, che a sua volta ha delegato il GSA (General Services Administration) come responsabile per l'assegnazione operativa degli indirizzi.

Si noti come il campo DSP abbia una struttura gerarchica attenta unicamente a riprodurre il metodo OSI di instradamento globale. Infatti, secondo tale approccio, l'individuazione di ciascun sistema d'utente deriva dalla sua localizzazione in un dominio di instradamento (individuati nel DSP dal sottocampo RD) ed all'interno di esso in un'area (individuata nel DSP dal sottocampo Area). All'interno di ogni area i terminali d'utente sono distinti in modo non ambiguo da valori del sottocampo ID e, all'interno di un sistema d'utente, il sottocampo Sel individua un particolare N-SAP.

L'autorità GSA assegna alle agenzie Governative un valore del sottocampo Administrative Authority (AA).

Può essere osservato che ciascuna autorità, alla quale il GSA ha assegnato un valore AA, procederà alla definizione di un'opportuna ripartizione della sua rete in domini di instradamento ed aree. In questo modo si avrà la coesistenza di più ripartizioni in domini di instradamento all'interno della stessa area geografica, ciascuna facente capo alla corrispondente autorità Amministratrice.

3.4 Struttura dell'indirizzo N-SAP definita dall'ANSI

Parallelamente al GOSIP americano, l'ANSI (American National Standards Institute), ambito in cui confluiscono interessi dei più diversi settori pubblici e

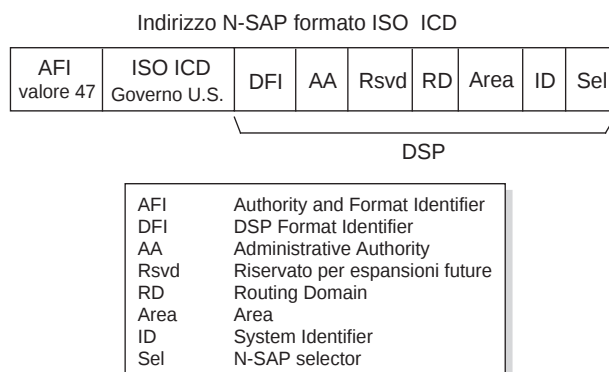


Figura 7 Struttura dell'indirizzo N-SAP per il GOSIP americano

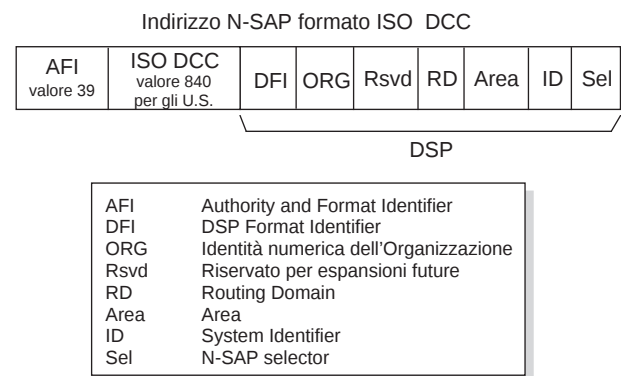


Figura 8 Struttura ANSI dell'indirizzo N-SAP

privati, ha definito una struttura di N-SAP basata sul formato ISO DCC (formato con connotazione geografica: gli Stati Uniti sono individuati con il valore 840) e mirata all'utilizzo in ambito privato o in enti non appartenenti al settore governativo americano. In fig. 8 è presentata la struttura come definita dall'ANSI.

In questo caso l'obiettivo principale dell'indirizzo N-SAP è quello di individuare organizzazioni private con nomi a sintassi numerica (sottocampo ORG). La restante parte del campo DSP coincide con quella GOSIP.

3.5 Comunità Internet ed indirizzo N-SAP

La Comunità Internet (ISOC, Internet Society Community) è un'organizzazione completamente autonoma rispetto agli ambiti tradizionali di standardizzazione e prevede un organismo specifico (Internet Engineering Task Force, IETF) per lo sviluppo nel breve/medio periodo delle normative necessarie. Ha una specifica procedura per la definizione di uno standard molto flessibile e veloce, anche se spesso formalmente non impeccabile.

Il documento specifico su cui si basa tale procedura è la proposta di standard, detto Request For Comments (RFC). Questi RFC sono documenti completamente pubblici che, una volta concordati, attivano un processo di trasformazione in standard. Al momento tali standard hanno una scarsa validità formale per le realtà di normativa ufficiale, ma rappresentano un importante punto di riferimento per la notevole penetrazione che hanno sul mercato⁽²⁾.

L'IETF ha da tempo promosso un'attività di studio per identificare la struttura più adatta dell'indirizzo N-SAP in relazione all'insieme di standard prodotti dall'ISO [13]. I formati proposti sono analoghi a quelli descritti in precedenza relativamente all'ANSI ed all'U.K.

(2) Sono oggi in corso contatti tra l'ISOC e l'ISO affinché la Comunità Internet e i relativi standard possano in un prossimo futuro essere riconosciuti formalmente come standardizzazione internazionale.

4. Aspetti evolutivi

Come si è detto, obiettivo principale dell'N-SAP è quello di fornire all'utente una crescente possibilità di interconnettersi con altri utenti, superando gli ostacoli posti dall'eterogeneità delle reti di trasporto che così utilizzano (ad es. reti telefoniche, reti dati, MAN, LAN ed ATM).

Un forte impulso a tale approccio è stato dato dalla Comunità Europea che già a partire dal 1987 con la Decisione 87/95/CEE [14] e la Risoluzione 88/257/CEE [15], poi riconfermata e sostenuta nella Direttiva 90/387/CEE [16] relativamente alla realizzazione dell'Open Network Provision (ONP), ha promosso l'utilizzo degli Standard Internazionali ISO, di cui l'OSI ne rappresenta la parte dominante, per individuare i prodotti da acquisire nell'ambito delle reti governative e delle Amministrazioni pubbliche.

È importante ricordare che gli standard ISO, ormai quasi sempre armonizzati con l'ITU, rappresentano l'unico punto di riferimento a livello di normativa internazionale per la realizzazione di applicativi capaci di assicurare la compatibilità tra sistemi di costruttori diversi su scala mondiale. E' inoltre importante considerare che, relativamente alle specifiche del mondo Internet, come è stato già ricordato, sono in corso contatti tra l'ISOC e l'ISO per il riconoscimento reciproco delle normative. Ciò potrà permettere in futuro l'armonizzazione tra due mondi finalizzati al raggiungimento di obiettivi analoghi, ma che oggi viaggiano su binari diversi.

La problematica della connettività globale, cioè della possibilità per un utente di interconnettersi con altri utenti attestati a differenti infrastrutture di rete, dal punto di vista della numerazione e dell'indirizzamento, trova negli Enti di normativa internazionale uno specifico approccio, che individua due aspetti predominanti nelle procedure di indirizzamento. Il primo riguarda l'indirizzamento nell'ambito delle tradizionali sottoreti, nelle quali si prevede nel caso delle reti pubbliche l'utilizzo di numeri dei piani di numerazione pubblici (X.121 e E.164) e nel caso delle reti private schemi di indirizzamento ISO/IEEE (indirizzi LAN a 48 bit [6]). Questi piani di numerazione/indirizzamento si possono definire come dei meccanismi di indirizzamento locali a ciascuna infrastruttura di rete per realizzare il trasporto delle informazioni limitatamente all'area di servizio corrispondente. Il secondo aspetto che affronta l'indirizzamento al confine tra i servizi di trasporto e gli applicativi d'utente, ha significato globale ed è indipendente dall'infrastruttura specifica. Lo standard ISO/OSI che "copre" tale aspetto è proprio l'indirizzo N-SAP.

La Comunità Internet utilizza già da tempo un approccio simile basato sui noti protocolli proprietari TCP/IP e, per quanto riguarda l'indirizzamento,

sull'indirizzo proprietario IP. La rete Internet internazionale offre già alcuni servizi di interconnessione e trasferimento di informazioni, limitatamente a tutti i siti in cui sono installati i protocolli TCP/IP (in Europa sostanzialmente nei Centri di ricerca e nelle Università).

A partire dalle normative illustrate nei precedenti paragrafi si ritiene utile proporre un'esemplificazione delle conseguenze concrete che tali normative comportano nell'utilizzo dell'indirizzo N-SAP, con particolare riferimento alla strategia di instradamento globale da utilizzare.

In fig. 9 è presentato lo scenario indicato dall'ISO, sulla base dei protocolli di instradamento illustrati nei paragrafi precedenti. Nella figura è presentata una rete di TLC, facente capo ad un certo fornitore di servizi di telecomunicazione, formata da un insieme di infrastrutture di reti pubbliche e private interconnesse fra loro mediante unità Gateway di interlavoro (elementi G).

L'indirizzo N-SAP, come già indicato, non contiene informazioni utili per instradare all'interno delle varie infrastrutture di rete, ma è un identificativo "globale" degli oggetti in rete al confine del dominio degli applicativi d'utente. I meccanismi di indirizzamento da utilizzare localmente sono derivati dall'indirizzo N-SAP e tradotti in forme di indirizzamento di sottorete.

Secondo la normativa ISO la prima parte dell'indirizzo N-SAP (campo IDP), che individua l'Autorità amministratrice (cioè nel caso di fig. 9 il fornitore di servizi di TLC), non ha informazioni direttamente utilizzabili per l'individuazione del sistema destinatario all'interno dell'insieme di sottoreti che costituiscono la rete di TLC. E' quindi necessario "leggere" parte del campo DSP, anche se la struttura di tale campo non è standardizzata dalla ISO ed è affidata alla responsabilità della corrispondente Autorità amministratrice.

In concreto ciascuna autorità identificata dal campo

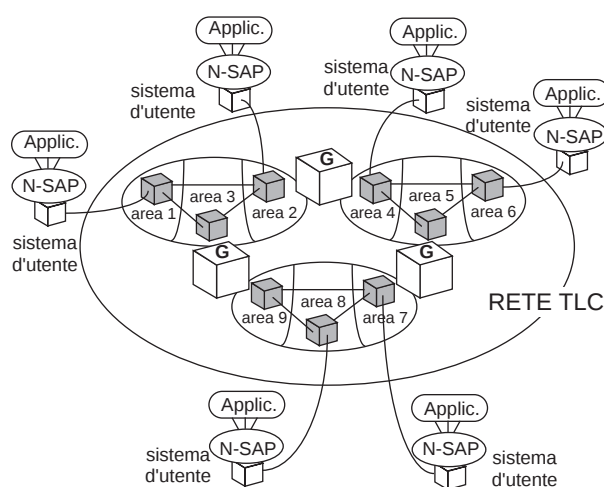


Figura 9 Interpretazione OSI dell'indirizzo N-SAP

IDP deve definire una ripartizione della rispettiva rete di TLC in aree di indirizzamento ed assegnare in modo non ambiguo a ciascuna area una combinazione delle prime cifre del campo DSP (come presentato nelle figg. 7 e 8).

Infine la restante parte del DSP deve contenere informazioni non ambigue all'interno di ciascuna area di indirizzamento per l'individuazione specifica del terminale d'utente ed eventualmente dell'applicativo al suo interno.

Questa metodologia di instradamento in uno scenario di interlavoro di rete rispecchia una suddivisione del processo in due livelli gerarchici: il primo è l'instradamento all'interno di un'area ed il secondo riguarda l'instradamento tra aree di indirizzamento diverse. Come già indicato in precedenza, l'ISO ha sviluppato due protocolli differenti [11,12] per realizzare tali livelli di instradamento che, per quanto riguarda le informazioni di indirizzamento, richiedono una strutturazione del campo DSP nelle due parti gerarchicamente distinte illustrate nelle figg. 7 e 8.

La costruzione precisa del campo DSP è quindi compito dell'Autorità amministratrice in relazione alle effettive necessità delle infrastrutture di rete e dei servizi che utilizzerà. Ciascuna singola Autorità deve perciò procedere alla definizione di una sua specifica normativa.

4.1 Comunità Internet

La Comunità Internet ha da tempo promosso un'attività di studio nell'ambito della evoluzione dei meccanismi di indirizzamento nella rete Internet internazionale o "Big Internet". Infatti già verso la metà degli anni '80 erano stati evidenziati alcuni aspetti di criticità nel medio termine dell'attuale forma di indirizzamento utilizzata nei protocolli TCP/IP, cioè il cosiddetto indirizzo IP (Internetworking Protocol). È stata oggi confermata, una scarsa disponibilità di indirizzi per una rete con copertura mondiale e, aspetto più grave, l'assenza di una struttura gerarchica sta rendendo ingestibili le tabelle di instradamento nei gateway.

L'IETF ha avviato le attività di valutazione di possibili soluzioni, anche indagando la possibilità di utilizzare una forma di indirizzamento strutturalmente aperta verso il mondo ISO/OSI.

La scelta di un nuovo schema di indirizzamento deve considerare non solo aspetti di opportunità tecnica, ma anche problematiche politico-commerciali ed aspetti di compatibilità con lo schema di indirizzamento IP esistente (minimizzazione degli impatti sull'utenza, ecc.). Inoltre il consolidarsi del protocollo OSI/ITU connectionless CLNP (ConnectionLess Network Protocol) [17], molto simile dal punto di vista del servizio offerto al protocollo IP e basato sull'utilizzo dell'indirizzamento N-SAP, ha comportato la sua naturale e progressiva introduzione in molti gateway Internet come nuova "pila" TCP/CLNP in parallelo alla tradizionale "pila" TCP/IP, come è illustrato in fig. 10.

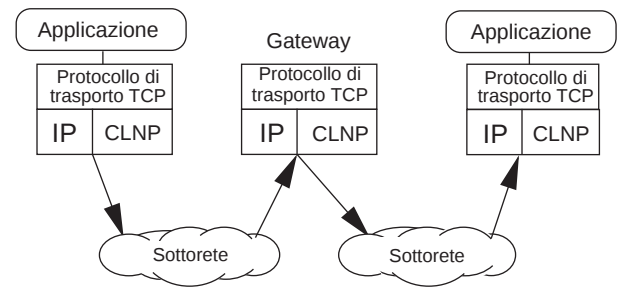


Figura 10 Tendenza evolutiva della "Big Internet"

Allo stato attuale la Comunità Internet sta definendo una nuova struttura di indirizzo IP che prevederà la coesistenza di più formati tra i quali l'indirizzo N-SAP relativamente al mondo ISO/OSI. La tendenza è quella di estendere l'indirizzo IP attuale senza comportare l'automatica sostituzione degli indirizzi a tutti gli utenti.

È importante tenere presente che tali aspetti prettamente tecnici nascondono conflitti politico-commerciali tra la Comunità Internet e gli Organismi tradizionali di standardizzazione ISO e ITU che, pur con notevoli ritardi, hanno ormai disponibili protocolli standard con il "marchio" OSI e Standard Internazionale.

4.2 ATM Forum

La tematica N-SAP è da considerare particolarmente rilevante nelle attività di standardizzazione e specifica di nuove tecniche di rete, quali la tecnica ATM (Asynchronous Transfer Mode). In ambito ATM Forum il mondo del privato, contrapposto a quello di fornitura di servizi pubblici, si sta muovendo molto rapidamente nelle attività di specifica per le enormi capacità prestazionali che le nuove tecniche di rete offrono e di conseguenza per l'ampia gamma di servizi che esse offrono: dai servizi dati tradizionali ed evoluti, ai prospettati servizi in fonìa.

In tale sede l'obiettivo perseguito riguarda la definizione di specifiche implementative con un prevalente influsso da parte del mondo nordamericano. Infatti l'approccio ATM Forum risulta polarizzato su una topologia di rete basata su un insieme di sottoreti di media e di piccola copertura interconnesse tra loro, la cui proprietà può essere indifferentemente pubblica/governativa o privata con topologia analoga ad Internet.

L'approccio ITU risulta invece basato su una topologia di rete costituita tradizionalmente da grosse sottoreti pubbliche, solitamente per una copertura nazionale, che interconnettono "isole" private, sostanzialmente basate su reti LAN. In tali scenari si inserisce la tecnica ATM come una nuova tecnologia che inizialmente si affiancherà a quelle esistenti (si veda la fig. 11).

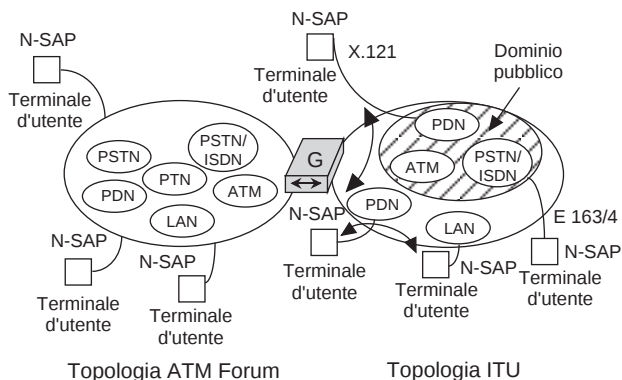


Figura 11 Topologia ITU e ATM Forum a confronto

Nella topologia dell'ATM Forum la struttura di indirizzo N-SAP ha assunto una notevole importanza soprattutto per il suo utilizzo in ambito privato. Il dominio di indirizzi N-SAP è stato ritenuto sufficientemente ampio da risolvere qualsiasi problema di quantità di indirizzi e, nel contempo, adatto ad una topologia di rete formata dall'unione di una notevole quantità di sottoreti. Si noti che ciascuna sottorete ha normalmente una ridotta copertura territoriale, ma esse considerate nel complesso possono raggiungere un ambito di copertura sovranazionale paragonabile a quello delle tradizionali reti internazionali.

Nel corso dell'ultimo anno le posizioni dell'ATM Forum ed ITU si sono notevolmente avvicinate, in quanto è stata riconosciuta l'importanza di produrre specifiche compatibili. L'ATM Forum tende sempre di più ad essere un ambito di "implementation agreement" a partire dalle specifiche prodotte in ITU, più che Organismo che produce autonomamente specifiche ATM.

5. Aspetti di registrazione per gli schemi di indirizzamento N-SAP

La corretta comprensione degli schemi di indirizzamento N-SAP comporta l'approfondimento delle problematiche di amministrazione e registrazione connesse con una forma di indirizzamento con le suddette caratteristiche di globalità.

Nel seguito sono analizzati in maggiore dettaglio gli aspetti di assegnazione e gestione connessi con la struttura di indirizzi N-SAP.

In fig. 12 è presentata mediante una struttura ad albero, detto albero di naming, la gerarchia dell'indirizzo N-SAP come normalizzata congiuntamente da ISO ed ITU [5]. I vari formati dell'indirizzo N-SAP, individuati da valori differenti del campo AFI, sono amministrati internazionalmente al massimo livello gerarchico dagli enti indicati in fig. 12, che hanno la responsabilità di assegnare i valori del formato corrispondente (campo IDI).

Nel caso dei formati di N-SAP legati ai piani di numerazione pubblici ITU l'autorità responsabile è la stessa ITU e per ciascun formato essa è delegata allo Study Group (SG) responsabile della Raccomandazione corrispondente: lo SG2 assegna i Country Code E.164 (reti telefoniche/ISDN), lo SG7 i Data Country Code X.121 (reti dati pubbliche PDN) e lo SG1 i Telex Destination Code F.69 (rete pubblica telex).

In ciascuno di tali casi l'Amministrazione/PTT di ogni nazione, in quanto responsabile dell'assegnazione dei numeri appartenenti ai piani pubblici, individua le autorità preposte alla strutturazione ed assegnazione della parte restante dell'indirizzo. In altre parole tutti gli utenti connessi ad una rete pubblica, e che quindi sono

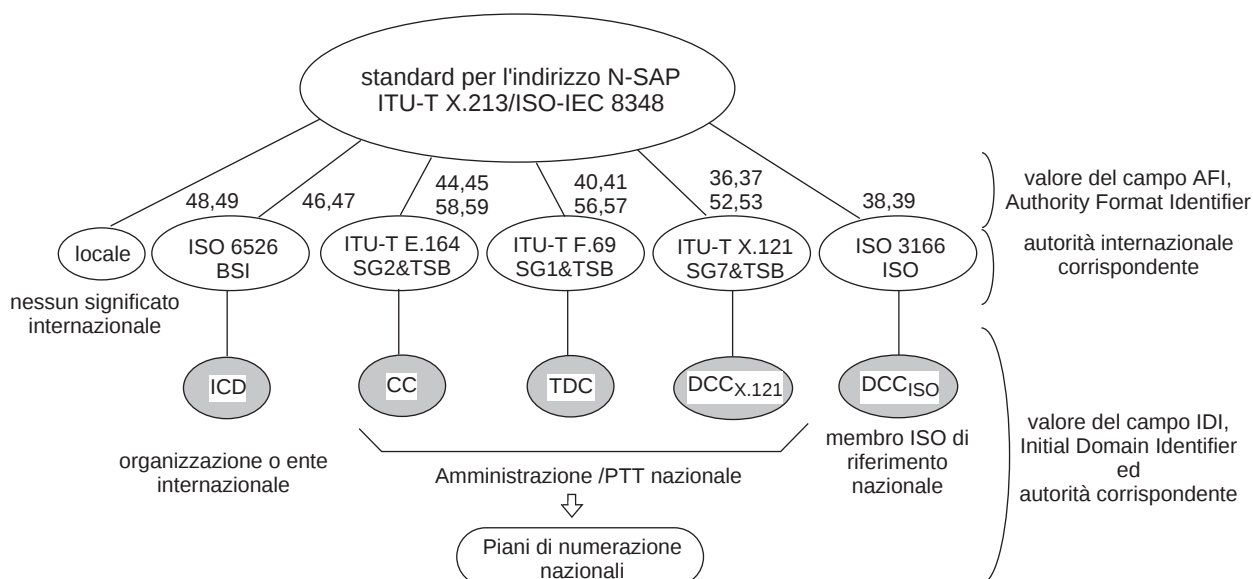


Figura 12 L'albero di naming per l'indirizzo N-SAP

identificati con un numero E.164, X.121 o F.69, sono automaticamente potenziali autorità per l'assegnazione della parte restante dell'indirizzo N-SAP (campo DSP).

Il formato di indirizzo N-SAP ISO DCC è analogo al precedente, dal momento che per ciascuna nazione è allocato da ISO un numero identificativo a 3 cifre decimali sulla base delle regole fissate dallo Standard Internazionale ISO 3166 [7]. L'ente associato nazionalmente al valore DCCISO è questa volta il Membro ISO principale o di riferimento nazionale, che in generale può differire dall'ambito prettamente Amministrativo/PTT (nel caso dell'Italia ad esempio tale ruolo è assunto dall'organizzazione UNINFO). Anche in questo caso tale Ente assume il ruolo di autorità responsabile della strutturazione ed assegnazione della parte restante dell'indirizzo (campo DSP).

Nel caso del formato non geografico ISO ICD l'ISO ha delegato il BSI, l'ente di normativa nazionale britannico (paragonabile come ruolo all'ente italiano UNI), come autorità responsabile dell'assegnazione dei valori ICD sulla base dello standard ISO 6253 [8]. Le richieste di identificatori internazionali ICD devono essere patrocinate da un membro ISO (non necessariamente il Membro ISO principale di ciascuna nazione) e un'organizzazione, nel momento in cui riceve un identificatore ICD, diviene l'autorità amministratrice della parte restante dell'indirizzo (campo DSP).

Infine il formato locale non ha alcuna struttura che risponda a normative internazionali e, di conseguenza, chiunque può utilizzarlo limitatamente ad ambiti ristretti all'interno dei quali è concordato il suo utilizzo sia a livello sintattico che semantico.

Riassumendo le autorità, individuate dalle "foglie" dell'albero di naming di fig. 12 (indicate in figura mediante un cerchio scuro), rappresentano le autorità preposte internazionalmente alla gestione dei diversi formati di indirizzo N-SAP. Dal momento che alla base dello standard sull'indirizzo N-SAP vi è la realizzazione di una gerarchia dell'indirizzo e della sua amministrazione, le suddette autorità possono ulteriormente strutturare l'indirizzo N-SAP, in particolare il campo DSP, demandando a sotto-autorità il compito dell'effettiva gestione dell'indirizzo. In fig. 13 è illustrato tale approccio gerarchico.

Ciò significa che le "foglie" dell'albero di fig. 12 sono in generale le radici di uno o più sottoalberi che definiscono la struttura completa dell'indirizzo N-SAP ed in particolare del campo DSP. Tali sottoalberi

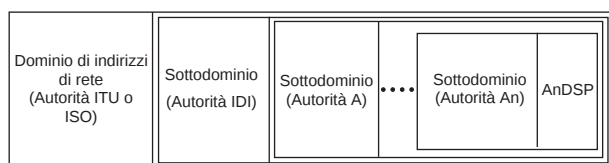


Figura 13 Struttura gerarchica dell'indirizzo N-SAP

possono avere aree di validità diverse e sono decisi dalle autorità corrispondenti alle suddette foglie dell'albero di fig. 12 (le autorità individuate dal contenuto del campo IDI).

In fig. 14 sono illustrate le due strutture di sottoalbero del campo DSP, che internazionalmente stanno emergendo in ambito Governativo U.S. (GOSIP), ANSI ed Internet (fig. 14.a), e in ambito BSI (fig. 14.b).

Sono approcci compatibili, ma che presentano differenti margini di flessibilità. In entrambi i casi è individuata un'autorità/organizzazione (prime cifre del campo DSP) che ha in carico la strutturazione e l'assegnazione della parte restante del campo DSP e solo nel caso a) è definita una precisa ripartizione sul territorio (nazionale e/o internazionale) del dominio di indirizzi in domini di instradamento ed aree di indirizzamento. Nel caso b) invece la ripartizione del dominio di indirizzi sul territorio non ha particolari requisiti. Ciò è anche dovuto alle necessità di adattamento a differenti topologie di rete a cui si fa riferimento nei due casi esaminati: il caso a) nasce ed è principalmente utilizzato in Internet mentre il caso b) considera la presenza tradizionale in Europa di una topologia di rete costituita da grosse infrastrutture di reti a copertura prevalentemente nazionale interconnesse tra loro da "carrier" internazionali.

I sottoalberi di fig. 14 possono essere applicati a qualsiasi formato di indirizzo N-SAP, quindi l'albero di fig. 12 può essere considerato il prolungamento di una qualsiasi delle "foglie": infatti tali sottoalberi si riferiscono unicamente al formato del campo DSP.

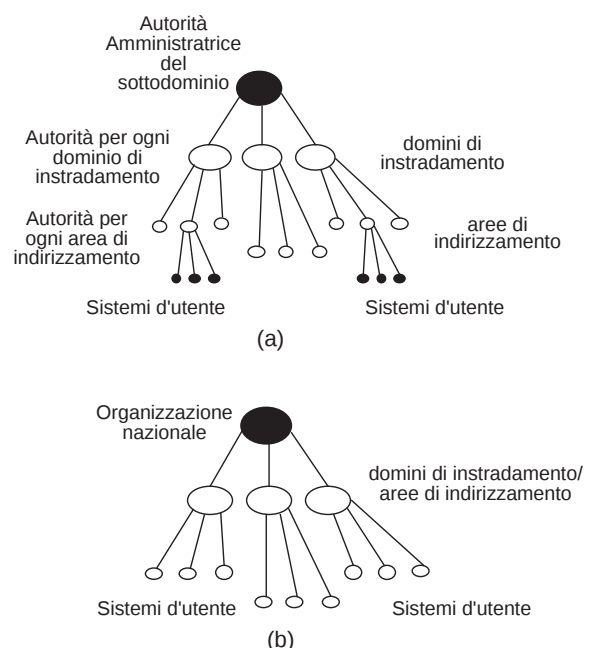


Figura 14 Gerarchia di assegnazione ed amministrazione del campo DSP

5.1 Autorità nazionali per la gestione ed assegnazione degli indirizzi N-SAP

Da quanto illustrato precedentemente, in ambito nazionale devono essere identificati opportuni Enti che assumano il ruolo di Autorità di gestione dell'indirizzo N-SAP. Queste Organizzazioni possono assegnare direttamente l'indirizzo N-SAP, definendo opportune procedure, o delegano altre Organizzazioni pubbliche o private per la gestione operativa dell'indirizzo.

I requisiti minimi per la gestione ed assegnazione degli indirizzi N-SAP a livello nazionale possono essere riassunti nei seguenti punti:

- **formato ISO ICD:** le Organizzazioni interessate ad utilizzare tale formato devono farne richiesta diretta all'Ente di standardizzazione britannico BSI. La domanda dovrà essere patrocinata da un qualsiasi membro ISO, nazionale o estero. In questo caso non sono necessarie naturalmente autorità nazionali specifiche;
- **formati E.164/X.121/F.69:** l'Autorità nazionale attualmente responsabile della gestione, strutturazione ed assegnazione del piano di numerazione telefonico/ISDN (E.164), del piano di numerazione Telex (F.69) e del piano di numerazione per le reti dati (X.121) (rete dati) è il Ministero delle Poste e Telecomunicazioni. Nel momento in cui un'Organizzazione o un utente residenziale è identificato da un numero appartenente a tali piani, diventa automaticamente l'autorità che ha il carico della strutturazione e della assegnazione del campo DSP. Anche in questo caso non risulta necessaria alcuna Autorità nazionale;
- **formato ISO DCC:** in questo caso l'Autorità nazionale che ha la responsabilità dell'Amministrazione, strutturazione ed assegnazione del campo DSP è quella individuata dall'identificatore geografico ISO DCC (380 per l'Italia) [7]. Si deve identificare nell'ambito di ciascun Paese un membro Principale ISO (detto membro "P") che per l'Italia è l'UNINFO.

5.2 Amministrazione e gestione degli indirizzi Internet

Come esempio di gerarchia di assegnazione e gestione di indirizzi globali, attualmente attivo, può essere considerato quello della Comunità Internet.

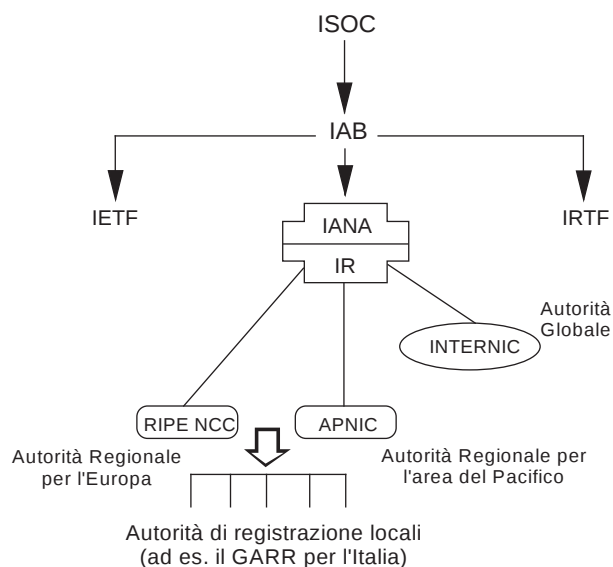
Infatti, in conseguenza della criticità di indirizzi IP disponibili, individuata sul finire degli anni '80, la Comunità Internet ha promosso un radicale cambiamento nelle modalità di gestione ed assegnazione degli indirizzi, in precedenza scarsamente formalizzate e controllate. Nell'arco temporale 1992-1993 l'IETF, anima tecnica della

Comunità Internet, ha definito un insieme di precise procedure formali ed una nuova gerarchia di assegnazione basata sulla delegazione regionale dei ruoli di autorità di registrazione. In fig. 15 è illustrata la suddivisione delle responsabilità di "Registration Authority" recentemente instaurata all'interno della Comunità Internet per l'assegnazione degli indirizzi IP.

Nella figura è presentato un possibile organigramma di Internet relativamente agli Organismi preposti al coordinamento e alla gestione dell'infrastruttura di rete Internet (IAB, IETF ed IRTF) ed è illustrata in dettaglio la parte relativa all'autorità di registrazione.

La IANA è l'autorità responsabile per la gestione formale dei vari identificatori necessari al funzionamento della rete e dei servizi e l'IR è la parte operativa che sovrintende alla gestione quotidiana delle procedure di assegnazione.

L'ente che operativamente rappresenta l'Authority è l'INTERNIC e che, oltre ad essere l'autorità globale di Internet, ricopre il ruolo di autorità di registrazione per il Nord America e per le aree che non hanno istituito un'autorità regionale (RIPE NCC per l'Europa e l'APNIC per l'area del Pacifico).



APNIC	Area Pacific Network Information Centre
GARR	Gruppo di Armonizzazione delle Reti di Ricerca
IAB	Internet Architecture Board
IANA	Internet Assigned Network Authority
IETF	Internet Engineering Task Force
IR	Internet Registry
IRTF	Internet Research Task Force
ISOC	Internet Society
RIPE NCC	Reseaux IP Europeens Network Coordination Centre

Figura 15 Gerarchia delle autorità di registrazione in Internet

6. Conclusioni

Nella realizzazione di un servizio di telecomunicazione fra due utilizzatori finali, dopo aver instaurato la connessione in rete tra i due sistemi finali è necessario individuarne, al loro interno, l'opportuna applicazione. Il meccanismo di indirizzamento N-SAP per il suo significato globale può essere utilizzato per realizzare e mantenere un legame tra le singole applicazioni del sistema d'utente ed i vari indirizzi di sottorete che permettono di realizzare l'instradamento nella catena di reti coinvolte.

In una rete pubblica si può scegliere di realizzare il solo trasporto degli indirizzi N-SAP con funzioni di sottoindirizzamento e di delegare ai privati il suo effettivo utilizzo, oppure si può utilizzare tale risorsa per la fornitura all'utenza di servizi a valore aggiunto e di teleservizi. Per esempio si potrebbero offrire direttamente servizi evoluti di interconnessione di LAN, di posta elettronica o di accesso a directory che richiedono l'indirizzamento diretto di applicativi residenti in sistemi pubblici o privati.

Un meccanismo di estensione delle funzionalità di indirizzamento ha anche altri vantaggi: fornisce un potente strumento per semplificare le problematiche di interlavoro relativamente agli aspetti di numerazione e indirizzamento. Si ha infatti l'opportunità di offrire una procedura unica per la realizzazione di qualsiasi interlavoro all'interno o no del mondo OSI, presentando meccanismi di compatibilità con forme di indirizzamento proprietarie. Ciò presenta l'ulteriore vantaggio, sia dal punto di vista gestionale sia tecnico, di permettere di mantenere separate le informazioni di indirizzamento esterne al dominio pubblico rispetto a quelle proprie di tale dominio. In tal modo si riduce al minimo l'utilizzo della risorsa di numerazione pubblica per scopi privati (ad esempio selezione passante, multinumero).

Di particolare importanza risultano gli aspetti di registrazione e, in particolare, la definizione di opportune "authority" internazionali e nazionali per la gestione ed assegnazione corretta dei diversi schemi di indirizzamento.

In conclusione lo sviluppo delle attività di standardizzazione sullo schema di indirizzamento globale N-SAP vede emergere possibili utilizzi concreti all'interno delle reti dati attuali, nelle reti dati ad alta velocità ed in prospettiva nella B-ISDN.

Acronimi

AFI	Authority and Format Identifier
ANSI	American National Standards Institute
APNIC	Area Pacific Network Information Centre
ATM	Asynchronous Transfer Mode
BSI	British Standard Institute
CC	Country Code

CEE	Comunità Economica Europea
CLNP	ConnectionLess Network Protocol
DCC	Data Country Code
DSP	Domain Specific Part
EEA	Electronics and Business Equipment Association
GARR	Gruppo di Armonizzazione delle Reti di Ricerca
GOSIP	Government OSI Profile
GSA	General Services Administration
IAB	Internet Architecture Board
IANA	Internet Assigned Network Authority
ICD	International Code Designator
IDI	Initial Domain Identifier
IDP	Initial Domain Part
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
IR	Internet Registry
IRTF	Internet Research Task Force
ISDN	Integrated Services Digital Network
ISO	International Standards Organization
ISOC	Internet Society
ITU	International Telecommunication Union
ITU T	International Telecommunications Union - Telecommunications Standard Sector
LAN	Local Area Network
MAN	Metropolitan Area Network
N-SAP	Network Service Access Point
NDC	National Destination Code
NIST	National Institute of Standards and Technology
NUA	Numbering and Addressing
ONP	Open Network Provision
OSI	Open System Interconnection
PDN	Public Data Network
PSTN	Public Switched Telephone Network
PTN	Private Telecommunication Network
PTT	Postal, Telegraph and Telecommunications administration
RFC	Request For Comments
RIPE NCC	Reseaux IP Europeens Network Coordination Centre
SG	ITU-T Study Group
TDC	Telex Destination Code
TSB	Telecommunication Standardization Bureau
U.S.	United States
U.K.	United Kingdom

Bibliografia

- [1] *Numbering plan for the ISDN era*. Racc. ITU-T E.164, 1991.
- [2] *International Numbering Plan for the Public Data Network*. Racc. ITU-T X.121, 1992.
- [3] *The international telex service - Service and operational provisions of telex destination codes and telex network identification codes*. Racc. ITU-T F.69, 1994.
- [4] *Information technology - Open Systems Interconnection - Basic reference*. Racc. ITU-T X.200, 1994.
- [5] *Information Processing Systems - Data Communications - Network Service Definition*. ISO/IEC 8348, 1992. *Information Technology - Network Service Definition for OSI*. Racc. CCITT X.213, 1992.
- [6] *IEEE Standards for Local and Metropolitan Area Networks*. IEEE std. 802, 1990.
- [7] *Codes for the Representation of Names of Countries*. ISO 3166, 1981.
- [8] *Data Interchange - Structure for the Identification of Organizations*. ISO 6523, 1984.
- [9] *Principles relating ISDN numbers/subaddresses to the OSI reference model network layer addresses*. Racc. CCITT I.334, Blue Book, 1988.
- [10] *Interface between data terminal equipment (DTE) and data circuit-terminating equipment (DCE) for terminals operating in the packet mode and connected to public data networks by dedicated circuit*. Racc. ITU-T X.25, 1993.
- [11] *Information Technology - Telecommunications and Information Exchange between Systems - Intermediate System to Intermediate System Intra-Domain Routing Information Exchange Protocol for Use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)*. Draft ISO/IEC 10589, 1992.
- [12] *Information Technology - Telecommunications and Information Exchange between Systems - Protocol for Exchange of Inter-Domain Routing Information among Intermediate Systems to Support Forwarding of ISO 8473 PDUs*. Draft ISO/IEC 10747, 1993.
- [13] *Guidelines for OSI N-SAP Allocation in the Internet*. RFC 1237, 1993.
- [14] Decisione 87/95/CEE, relativa alla normalizzazione nel settore delle tecnologie dell'informazione e delle telecomunicazioni.
- [15] Risoluzione 88/257/CEE, sullo sviluppo del mercato comune dei servizi e delle apparecchiature di telecomunicazione entro il 1992.
- [16] Direttiva 90/387/CEE, sull'istituzione del mercato interno per i servizi delle telecomunicazioni mediante la realizzazione della fornitura di una rete aperta di telecomunicazioni (Open Network Provision - ONP).
- [17] *Protocol for Providing the Connectionless-mode Network service*. ISO/IEC 8473, 1988.

Una nuova generazione di Sistemi di Trasmissione⁽¹⁾

G. De Marchis, F. Matera, M. Settembre (*)

Sin dai primi anni in cui le fibre furono utilizzate per sostituire i cavi in rame ci si trovò dinanzi ad una situazione che si ripresenta con puntualità: ogni volta che si cerca di sfruttarne al massimo la capacità trasmissiva, vale a dire una banda pressappoco infinita, ci si imbatte in un fenomeno fino a quel momento sottovalutato e che diviene il fattore limitante. Prima la dispersione modale, poi la dispersione cromatica, poi gli effetti non lineari, poi il rumore introdotto dagli amplificatori ottici, ora la dispersione di polarizzazione e prossimamente la dispersione di polarizzazione del secondo ordine.

Poiché la maggior parte di questi fenomeni sono oltre che ineliminabili anche di difficile comprensione per i non addetti ai lavori (si pensi ad esempio all'equazione di Schrödinger non lineare che descrive con la migliore approssimazione che conosciamo la propagazione in fibra nei casi divenuti oggi di interesse) abbiamo tentato nel seguito di spiegare i meccanismi che ne sono alla base e, con l'ausilio di un programma di simulazione numerica, di mostrare come i vari effetti interagiscano, in alcuni casi compensandosi almeno parzialmente.

L'intento è quello di fornire indicazioni utili per il progetto di sistemi per collegamenti che rappresentano il limite tecnologico oggi raggiungibile: migliaia di chilometri nei collegamenti transoceanici a velocità di cifra medio-alta (2.5 Gbit/s) o alcune centinaia in reti di trasporto trasparente ad altissima velocità (10 e 40 Gbit/s).

Poiché poi tecniche di trasmissione, giudicate appena ieri "esotiche" e di scarsa applicabilità pratica, hanno invece dimostrato piena vitalità e sono in sperimentazione già ora o in un futuro molto prossimo, si è affiancata all'analisi di sistemi convenzionali operanti con segnali NRZ quella dei sistemi a solitoni, mentre per le tecniche di multiplexazione si è presa in considerazione sia la tecnica nel dominio del tempo con sistemi singolo canale sia quella nel dominio della frequenza.

1. Introduzione

Un sistema per comunicazioni ottiche differisce dagli altri sistemi di comunicazione soprattutto per la frequenza della portante e, poiché la larghezza di banda trasmissibile è in prima approssimazione proporzionale alla frequenza della portante stessa, l'utilizzo di portanti ottiche permette in linea di principio la trasmissione di segnali con bande dell'ordine del centinaio di GHz. Proprio l'enorme capacità potenziale offerta dai sistemi di comunicazione ottica è stata la molla che ha innescato negli ultimi anni uno sviluppo sempre crescente della ricerca di nuove tecnologie per la generazione, la trasmissione e l'elaborazione di segnali ottici ad alta capacità.

Le aspettative sono state confermate dalla sperimentazione, specialmente grazie alla invenzione di alcuni dispositivi ottici fondamentali come le fibre

ottiche, i laser e gli amplificatori ottici: in particolare è stata dimostrata la possibilità di trasmettere segnali a 10 Gbit/s su 9000 km [1-2].

La comunicazione ottica si attua tramite un trasmettitore, un mezzo trasmissivo e un ricevitore: oggetto di questo lavoro è il mezzo trasmissivo, la fibra ottica, di cui verranno analizzate le proprietà fisiche rilevanti per la propagazione dei segnali ad alta velocità e lunga distanza. In linea con le tendenze attuali la linea di trasmissione si supporrà costituita dalla fibra e da amplificatori ottici che sostituiscono i rigeneratori elettrici usati in precedenza: questa si propone come la soluzione tecnologicamente più opportuna nei sistemi di trasmissione di nuova generazione in quanto la linea viene ad essere trasparente (a meno delle non idealità dell'amplificatore ottico) al segnale che vi transita, lasciando così la possibilità di miglioramenti futuri senza dover sostituire una gran mole di apparati.

(*) Ing. Giancarlo De Marchis, ing. Francesco Matera, dott.ssa Marina Settembre, -Fondazione Ugo Bordoni- Roma

(1) Lavoro eseguito nell'ambito della convenzione tra la Fondazione Ugo Bordoni e Telecom Italia.

Gli amplificatori ottici, in particolare quelli realizzati con fibre drogate con ioni erbio, presentano interessantissime proprietà: elevato guadagno, insensibilità alla polarizzazione, elevata potenza di saturazione, larga banda di guadagno, bassa figura di rumore. Queste caratteristiche, unitamente alla trasparenza rispetto al formato di modulazione del segnale e alla velocità di cifra, rendono questi dispositivi utili sia come preamplificatori che come amplificatori in linea o di potenza.

L'inserimento di tali amplificatori nei collegamenti in fibra ottica permette di annullare le perdite, ma al contrario dei rigeneratori opto-elettronici non compensa la distorsione accumulata dal segnale durante la propagazione in fibra e, inoltre, sovrappone al segnale il rumore dovuto all'emissione spontanea, cosa che peggiora il rapporto segnale rumore.

In una linea amplificata le principali limitazioni alle prestazioni di un sistema ottico di trasmissione sono dovute alla sovrapposizione di quattro effetti: la dispersione cromatica e la dispersione di polarizzazione della fibra, che tendono ad allargare temporalmente un impulso [3-4], la non linearità ottica della fibra, che tende ad allargare spettralmente l'impulso ed il rumore ottico di amplificazione [5].

Quando si ha la presenza simultanea di questi quattro fenomeni il segnale trasmesso è alterato sia spettralmente sia temporalmente ed in maniera tra l'altro diversa a seconda della regione di dispersione cromatica della fibra e del tipo di segnale stesso.

Nel seguito verranno descritti questi effetti con particolare riferimento alle loro implicazioni sulle prestazioni dei sistemi ottici ed inoltre verranno confrontati diversi sistemi di trasmissione allo scopo di individuare come sfruttare al meglio la grande larghezza di banda della fibra [6-8].

Nella sezione 2 si può trovare una rassegna sui principali effetti che limitano la propagazione dei segnali in fibra ottica insieme alla descrizione del metodo numerico utilizzato per studiare le prestazioni dei sistemi. La sezione 3 riporta alcuni risultati sulle prestazioni dei sistemi ottici con amplificazione periodica, risultati che ovviamente vogliono solo offrire uno spaccato delle varie situazioni che si possono presentare. La sezione 4 tratta la degradazione indotta dalla dispersione di polarizzazione e mostra come tale degradazione possa essere compensata con l'effetto Kerr.

2. Propagazione di un segnale nella fibra ottica

2.1 Effetti dovuti alla dispersione cromatica

Una fibra ottica trasmette le diverse componenti spettrali del segnale con ritardi diversi. Dal punto di

vista analitico, responsabile di tale fenomeno è la derivata seconda della costante di propagazione β , che sviluppata in serie di Taylor intorno alla frequenza centrale ω_0 fornisce

$$\beta(\omega) = \frac{\omega n(\omega)}{c} = \beta_0 + \beta_1 \Delta\omega + \frac{1}{2} \beta_2 \Delta\omega^2 + \frac{1}{6} \beta_3 \Delta\omega^3 + \dots (1)$$

dove $\Delta\omega = \omega - \omega_0$ e β_i rappresenta la derivata i -esima di β rispetto ad ω calcolata in ω_0 [29].

Il termine β_1 determina solo un ritardo nella propagazione del segnale e nel caso di fibre singolo modo non altera la forma del segnale mentre era responsabile della dispersione modale nelle fibre multimodo. Il termine β_2 , cui siamo interessati e diamo il nome di dispersione cromatica, è generalmente il contributo di distorsione più importante nei sistemi attuali e tiene in conto sia la dispersione del materiale sia la struttura guidante della fibra (il profilo dell'indice di rifrazione).

La dispersione cromatica o Dispersione della Velocità di Gruppo (DVG) gioca un ruolo critico nella propagazione di impulsi in fibra in quanto introduce un "chirp" nella frequenza istantanea della portante. La DVG cambia segno attraversando il valore zero per una certa lunghezza d'onda λ_0 , che per le fibre progettate per lavorare a 1300 nm vale circa 1310 nm (fibre SMF o standard) mentre si sposta nell'intorno di 1550 nm per le fibre cosiddette a dispersione spostata (DS o dispersion shifted) di cui l'Italia ha fatto uso abbondante nella rete a lunga distanza della ex-ASST. Per lunghezze d'onda inferiori a λ_0 si parla di dispersione normale ($\beta_2 > 0$) e la parte iniziale dell'impulso tende ad avere frequenze istantanee più basse rispetto alla coda dell'impulso, mentre nell'altro caso si parla di dispersione anomala ($\beta_2 < 0$) e la situazione si inverte. In entrambi i casi il processo tende ad allargare temporalmente gli impulsi nel corso della loro propagazione.

Ad esempio, un impulso gaussiano mantiene la sua forma durante la propagazione in z , ma la sua larghezza a metà altezza $T(z)$ evolve secondo la legge:

$$T(z) = \frac{T_0}{\sqrt{1 + \left(\frac{z}{L_D}\right)^2}} \quad (2)$$

dove la lunghezza di dispersione, $L_D = T_0^2 / |\beta_2|$ determina l'entità dell'allargamento. A parità di lunghezza di fibra, impulsi più corti si allargano di più, perché la lunghezza di dispersione è più corta.

In un sistema di trasmissione in fibra ottica a modulazione d'intensità, con codifica PCM standard, l'allargamento temporale indotto dalla dispersione cromatica provoca interferenza intersimbolica e la relazione che lega la massima velocità di cifra R_{max} alla distanza di propagazione L può essere descritta dalla seguente espressione [4]:

$$R_{max} = \sqrt{\frac{0.1}{|\beta_2|L}} \quad (3)$$

I sistemi in terza finestra (1550 nm) sono stati introdotti per limitare l'attenuazione indotta dalla fibra, che, come noto, è minima in questa regione. D'altro canto, la dispersione cromatica in terza finestra è sufficientemente elevata (17 ps/km•nm) da divenire il fattore limitante: per questo furono introdotte in rete le fibre a dispersione spostata (DS, $\lambda_0 \approx 1550$ nm) e nella fig. 1 si vede come tali fibre abbiano sensibilmente ridotto il degrado delle prestazioni per sistemi ad alta capacità su distanze più corte di 1000 km, mentre permane una severa limitazione per sistemi che trasmettono su più lunghe distanze.

Oltre a ciò, ogni fibra singolo modo presenta due modi degeneri, con diverse costanti di propagazione β , la cui differenza dà origine alla dispersione di polarizzazione (DP) di cui si parlerà nella sezione 4. Tale effetto risulta molto importante quando la lunghezza d'onda del segnale coincide con lo zero di dispersione della fibra o per durate di impulso molto corte (tipicamente <10 ps).

E' utile ricordare che, sebbene nella maggior parte dei casi di interesse β_3 sia trascurabile, talvolta è necessario includere anche gli effetti dovuti agli ordini superiori dell'espansione in serie della dispersione cromatica.

La degradazione dovuta alla dispersione cromatica può essere limitata utilizzando particolari schemi di compensazione: si può agire sul trasmettitore (tecniche di pre-chirping), si possono impiegare dispositivi basati su fibre con lo zero della dispersione spostato in una regione opportuna, si possono realizzare collegamenti con dispersione cromatica variabile lungo il cammino, con un valore medio complessivamente uguale a zero, ed infine si può utilizzare la stessa non linearità della fibra. Da citare inoltre una tecnica molto promettente, che consiste nell'invertire lo spettro in frequenza del segnale a metà del collegamento.

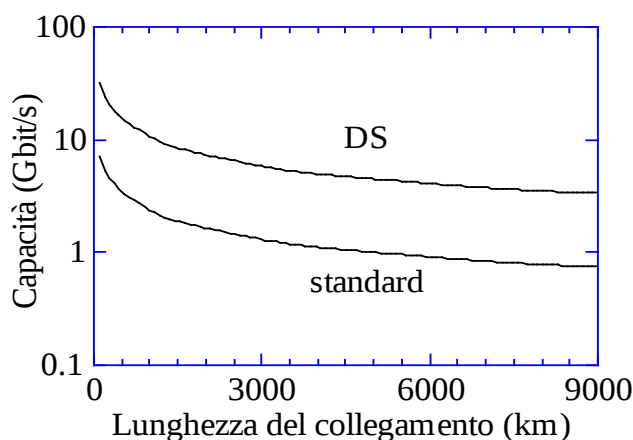


Figura 1 Massima capacità ottenibile in collegamenti in presenza di fibre DS e fibre SMF in assenza di ASE e effetto Kerr

2.2 Effetti dovuti alla non linearità della fibra

Anche se le fibre non sono fatte di materiali che esibiscono comportamenti altamente non lineari, la geometria della fibra che confina la luce in una piccola sezione su lunghe distanze rende i fenomeni non lineari molto importanti a causa della grande lunghezza di interazione. Come verrà illustrato nel seguito gli effetti non lineari non sono sempre dannosi, anzi in particolari condizioni possono anche risultare estremamente vantaggiosi ai fini di realizzare sistemi di telecomunicazione ad altissime capacità.

I principali effetti non lineari presenti nelle fibre sono l'effetto Brillouin, l'effetto Raman e l'effetto Kerr. I primi due sono dovuti a processi di scattering anelastico con la partecipazione di fononi⁽²⁾ ottici nel caso dell'effetto Raman e di fononi acustici nell'effetto Brillouin. Entrambi i processi inducono una perdita di energia nel segnale ogni volta che la potenza incidente supera un valore di soglia e costituiscono quindi un meccanismo addizionale di attenuazione. L'effetto Raman e l'effetto Kerr possono essere spiegati in termini del vettore di induzione elettrica $\mathbf{D}$ definito come

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (4)$$

dove ϵ_0 è la costante dielettrica del vuoto, $\mathbf{E}$ il campo elettrico e $\mathbf{P}$ il vettore di polarizzazione. In particolare il vettore $\mathbf{P}$ ha una dipendenza lineare e una dipendenza non lineare dal campo elettrico. Nel caso della fibra, che risulta un materiale centro simmetrico⁽³⁾, la parte non lineare della rappresentazione nel dominio della frequenza (trasformata di Fourier) del vettore $\mathbf{P}$ è data dal prodotto dalla terza potenza della trasformata di Fourier del campo elettrico per una funzione della frequenza, denominata suscettibilità del terzo ordine, $\chi^{(3)}$. La parte reale della $\chi^{(3)}$ è responsabile dell'effetto Kerr, mentre la parte immaginaria dell'effetto Raman.

2.2.1 Effetto Brillouin

L'effetto Brillouin consiste nel trasferimento di energia ad un'onda che si propaga in direzione contraria a quella del segnale ed è la non linearità della fibra con la soglia più bassa: fortunatamente può essere soppresso nella maggior parte dei sistemi di trasmissione [9], dipendendo la soglia dalla larghezza di banda della sorgente.

Allargando opportunamente la larghezza di riga della

(2) I fononi possono essere definiti come "quantità di energia vibrazionale".

(3) Ovvero, che ha simmetria di rotazione rispetto ad un asse. Nel caso della fibra ottica ideale (lineare, senza impurità né imperfezioni), le caratteristiche elettriche (es. l'andamento dell'indice di rifrazione) sono le stesse in ogni direzione ortogonale all'asse centrale del nucleo della fibra.

sorgente laser, la soglia aumenta di un fattore $B/\Delta\nu_B$, dove $\Delta\nu_B$ è la larghezza di banda dell'effetto Brillouin (tipicamente $\Delta\nu_B=20$ MHz) e B la larghezza di riga della sorgente. Poiché la larghezza di banda spettrale aumenta con la velocità di trasmissione, aumentando la velocità di cifra si induce un allargamento della curva di guadagno dell'effetto Brillouin riducendone il picco. Lo scattering Brillouin è indipendente dal numero dei canali e nel caso di sistemi multi canale la potenza in ciascun canale deve essere tenuta sotto la soglia dell'effetto Brillouin.

2.2.2 Effetto Raman

Al contrario dell'effetto Brillouin, nell'effetto Raman la radiazione spuria si propaga nella stessa direzione del segnale. La risposta Raman della fibra è caratterizzata da una curva con un guadagno basso e molto allargata (>5 THz). Questa caratteristica può limitare severamente le prestazioni di un sistema con più canali a diverse lunghezze d'onda (WDM), trasferendo l'energia da un canale a quelli vicini. E' possibile dimostrare che l'effetto Raman è trascurabile quando è soddisfatta la seguente relazione [9]

$$P_T^{(W)} B_T^{(Hz)} L_e^{(m)} < 9 \cdot 10^{15} \quad (5)$$

dove $L_e \sim z/(L_a \alpha)$ è la lunghezza efficace del collegamento che tiene conto delle perdite della fibra α , della lunghezza della fibra z , della distanza tra amplificatori L_a , mentre P_T e B_T sono rispettivamente la potenza ottica e la banda ottica totale del segnale. Si noti che le grandezze che compaiono nella (5) vanno misurate nelle rispettive unità del sistema MKS.

2.2.3 Effetto Kerr

L'effetto Kerr è invece dovuto alla dipendenza dall'intensità e dell'indice di rifrazione secondo la nota legge

$$n(E) = n_0 + n_2 |E|^2 \quad (6)$$

dove $n(E)$ rappresenta l'indice di rifrazione della fibra, n_0 l'indice lineare e $n_2 = 3\chi^{(3)}/(8n_0)$ è l'indice non lineare.

Dalla (6) segue che, se un segnale monocromatico a lunghezza d'onda λ_0 si propaga in presenza di effetto Kerr e in assenza di perdite, la fase ϕ del segnale si modifica lungo la distanza z secondo le legge

$$\phi(z) = \frac{2\pi n_0 z}{\lambda_0} + \frac{2\pi n_2 |E|^2 z}{\lambda_0} \quad (7)$$

Questo effetto risulta molto importante quando il segnale presenta una modulazione di intensità in quando la variazione della potenza del segnale stesso induce

una modulazione di fase o chirp: tale effetto è denominato Self Phase Modulation (SPM). A differenza della dispersione cromatica, la non linearità produce un chirp sempre dello stesso tipo.

Una prima conseguenza dell'effetto della SPM è che, in assenza di dispersione cromatica, un impulso durante la propagazione presenta allargamento sia spettrale sia temporale. Tale comportamento si modifica di molto in presenza della dispersione cromatica, poiché il chirp indotto dalla SPM può essere incrementato o diminuito dalla presenza della dispersione cromatica. In particolare nella regione di dispersione normale gli effetti della SPM e della GVD si accumulano e l'impulso tende ad allargarsi nel dominio del tempo. Viceversa nella regione anomala i due effetti si contrappongono ed in particolari condizioni si possono anche annullare permettendo all'impulso di mantenere inalterata la sua forma nel dominio sia del tempo sia della frequenza. Ciò avviene in particolare quando l'impulso presenta la forma di una secante iperbolica e la potenza di picco P_k è data dalla seguente relazione

$$P_k = \frac{\beta_2}{\gamma \tau_s^2} \quad (8)$$

dove γ è il coefficiente non lineare, $\gamma = n_2 \omega_0 / (c A_e)$, ω_0 la pulsazione corrispondente alla portante ottica, A_e l'area efficace del nucleo della fibra e τ_s rappresenta la larghezza totale a metà altezza dell'impulso, che in questo caso è denominato solitone.

Quando si ha la propagazione di segnali a frequenze diverse, a causa dell'effetto Kerr, la variazione di potenza di un segnale induce una modulazione di fase negli altri. In questo caso l'effetto è denominato Cross Phase Modulation (CPM). Tale effetto può essere semplicemente spiegato dal fatto che l'indice di rifrazione visto dal segnale "a" durante la propagazione dipende non solo dalla sua potenza, ma anche dalla potenza di un altro segnale, "b", secondo l'espressione

$$n_a(E) = n_0 + n_2 \left(|E_a|^2 + 2|E_b|^2 \right). \quad (9)$$

Quando un segnale è composto da due o più righe spettrali possono generarsi delle nuove componenti spettrali e tale fenomeno è denominato miscelazione a quattro onde (Four Wave Mixing, FWM). In particolare una nuova frequenza ω_4 è generata a partire da tre frequenze ω_1 , ω_2 e ω_3 incidenti nella fibra se sono verificate alcune condizioni (conservazione dell'energia e del momento). Nella propagazione in fibra la frequenza ω_4 data da:

$$\omega_1 - \omega_2 = \omega_3 - \omega_4. \quad (10)$$

Tale frequenza è generata con una efficienza che dipende dal fattore

$$\Delta\beta = [\beta(\omega_1) - \beta(\omega_2)] - [\beta(\omega_3) - \beta(\omega_4)]. \quad (11)$$

Nel caso di frequenze equispaziate di $\Delta\omega$ l'efficienza η del FWM risulta pari a [10]

$$\eta = \left(\frac{\sin(2\pi^2\beta_2\Delta\omega^2L_e)}{2\pi^2\beta_2\Delta\omega^2L_e} \right)^2. \quad (12)$$

Dalla (12) si deduce che la degradazione dovuta al fenomeno del FWM diminuisce all'aumentare della spaziatura tra i canali e della dispersione cromatica.

2.3 Rumore ottico

L'amplificatore ottico oltre a ripristinare la potenza ottica e quindi compensare le perdite dovute alla propagazione, aggiunge al segnale anche rumore dovuto all'emissione spontanea degli ioni erbio (ASE) [5].

Le caratteristiche più importanti associate a questo tipo di rumore sono la sua natura stocastica, l'accumulo della potenza di rumore durante la propagazione in fibra ed infine la non linearità del battimento tra rumore e segnale.

Queste caratteristiche dell'ASE sono dannose nei collegamenti in fibra per due motivi principali. Il primo è legato alle fluttuazioni di energia che introducono un limite inferiore all'energia del segnale per rendere distinguibili gli "1" dagli "0". Il secondo è dovuto all'interazione tra ASE e segnale tramite effetto Kerr in termini di FWM. Questa interazione, come mostrato dalla (12), è molto efficiente alla dispersione zero e può portare ad un grande allargamento spettrale [11]. Nel caso di propagazione solitonica il valore della GVD della fibra non è zero e quindi la banda di efficienza del FWM è molto stretta e come conseguenza lo spettro del segnale non si allarga. In questo caso l'ASE è però responsabile di una fluttuazione del tempo di arrivo degli impulsi, nota come effetto Gordon-Haus [12-13]. Gli effetti dannosi dell'ASE possono essere limitati inserendo dei filtri ottici nel collegamento in fibra [14-17].

2.4 Evoluzione di un impulso in fibra

L'evoluzione di un segnale in fibra ottica, in presenza di dispersione cromatica, dispersione del terzo ordine, perdite ed effetto Kerr, può essere studiata risolvendo (in genere solo numericamente) l'equazione di Schrödinger non lineare [3]

$$i \left(\frac{\partial U}{\partial \xi} + \frac{\alpha}{2} L_D U \right) = \frac{1}{2} \text{sign}(\beta_2) \frac{\partial^2 U}{\partial \tau^2} + \frac{i}{6} \frac{L_D}{L_D} \frac{\partial^3 U}{\partial \tau^3} - \frac{L_D}{L_{NL}} |U|^2 U \quad (13)$$

dove:

- $L_D = T_0^2/|\beta_2|$, $L'_D = T_0^3/|\beta_3|$ e $L_{NL} = (\gamma P)^{-1}$ sono tre grandezze che definiscono la scala delle lunghezze su cui gli effetti della dispersione cromatica al secondo e al terzo ordine (β_2 e β_3) e la non linearità della fibra

(γ) diventano importanti;

- T_0 è una costante arbitraria;
- α rappresenta l'attenuazione della fibra;
- $\xi = z/L_D$ e $\tau = t/T_0$ sono rispettivamente le variabili spaziale (ascissa lungo l'asse della fibra) e temporale normalizzate;
- $U = E/P^{1/2}$ essendo E l'intensità del campo elettrico e P la potenza d'ingresso.

Infine, $\text{sign}(\beta_2)$ è uguale ad 1 in condizioni di dispersione normale ($\beta_2 > 0$) e a -1 per dispersione anomala ($\beta_2 < 0$).

L'equazione (13) è stata risolta analiticamente solo in alcuni casi particolari mentre in genere occorre procedere numericamente. Le ipotesi fatte in questo lavoro considerano amplificatori ottici spaziali di una lunghezza L_a con il guadagno che compensa esattamente le perdite introdotte dal tratto di fibra precedente: $G = \exp(\alpha L_a)$; il rumore ottico di amplificazione è tenuto in conto aggiungendo a ciascuna componente spettrale termini di rumore statisticamente indipendenti la cui parte reale e immaginaria sono variabili gaussiane [5] con varianza $\sigma^2 = N_{sp} h \nu (G-1) \Delta \nu / 2$, in cui N_{sp} rappresenta l'inversione di popolazione ($N_{sp} = 1$ per inversione completa, nel seguito si assumerà $N_{sp} = 2$), h è la costante di Planck, ν è la frequenza della portante ottica e $\Delta \nu$ è il passo di campionamento in frequenza del segnale.

Il campionamento dello spettro si rende necessario per mere esigenze di calcolo e simulazione del collegamento: infatti occorre rappresentare gli andamenti continui dello spettro con un numero finito di valori (1024 nel caso in esame).

L'analisi numerica mostra che nel caso in cui la lunghezza del collegamento L sia molto minore della lunghezza di dispersione L_D , l'evoluzione del segnale è dominata dall'interazione non lineare tra il segnale ed il rumore di amplificazione, mentre nel caso inverso l'evoluzione è dominata dall'interazione tra effetto Kerr e dispersione cromatica.

In fig. 2 è riportato lo spettro di uscita di un segnale binario modulato in ampiezza a 5 Gbit/s con potenza di 1 mW dopo la propagazione in una fibra lunga 9000 km con amplificatori spaziali di 40 km per due diversi valori della dispersione cromatica della fibra: $\beta_2 = -1 \text{ ps}^2/\text{km}$ e $\beta_2 = -0.1 \text{ ps}^2/\text{km}$ (i valori nelle unità di misura abituali, $\text{ps}^2/\text{km} \cdot \text{nm}$, si ottengono moltiplicando per il fattore 0.8 e prendendo il valore assoluto).

Nel caso a) il rumore si comporta come nel caso di regime lineare e non induce un grande allargamento spettrale ed analizzando anche il comportamento nel dominio temporale è possibile osservare una parziale compensazione della distorsione dovuta alla dispersione cromatica mediante l'effetto Kerr [11]. Nel caso b) l'interazione non lineare tra segnale e rumore è in grado di generare nuove componenti spettrali su una banda di frequenza molto più vasta che nel caso precedente a causa dell'efficienza del FWM molto più alta a

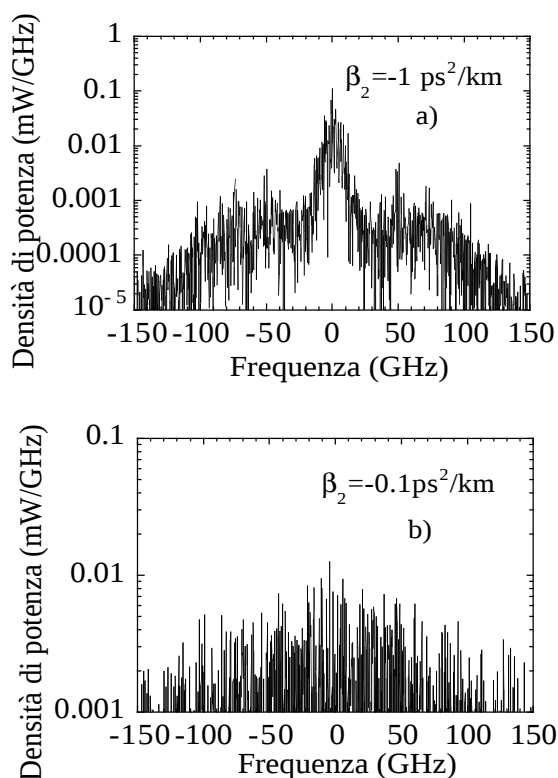


Figura 2 Spettro di uscita di un segnale modulato in ampiezza ASK a 5 Gbit/s con potenza di 1 mW dopo la propagazione in una fibra lunga 9000 km con amplificatori spaziali ogni 40 km per due diversi valori della dispersione cromatica della fibra: $\beta_2 = -1 \text{ ps}^2/\text{km}$ e $\beta_2 = -0.01 \text{ ps}^2/\text{km}$

dispersione bassa. In queste condizioni la compensazione non lineare della dispersione cromatica non è particolarmente evidente.

E' utile ricordare che l'allargamento spettrale è un importante parametro da controllare quando si vuole sfruttare la non linearità della fibra per compensare la degradazione indotta dalla dispersione cromatica, ma non è un criterio generale per valutare le prestazioni di un sistema. Infatti, quando la dispersione cromatica è molto bassa si possono ottenere buone prestazioni del sistema utilizzando filtri al ricevitore sufficientemente larghi da ricoprire la banda del segnale trasmesso allargata dall'interazione non lineare tra segnale e rumore [18].

3. Confronto tra le prestazioni di sistemi ottici

La soluzione numerica della (13) consente di analizzare e confrontare le prestazioni di diversi sistemi di trasmissione, con segnali sia solitonici sia convenzionali tenendo conto della mutua interazione

tra dispersione cromatica, non linearità della fibra e rumore ottico.

Nel seguito verranno presi in esame solo i risultati riguardanti i segnali modulati in intensità, in quanto i segnali modulati in fase o in frequenza sono maggiormente degradati dall'interazione non lineare tra segnale e rumore [11] e non sono, inoltre, molto diffusi al momento.

Saranno mostrati confronti tra la propagazione in collegamenti con fibre SMF ($\beta_2 = -20 \text{ ps}^2/\text{km} \Rightarrow 17 \text{ ps/km}\cdot\text{nm}$), con fibre DS ($\beta_2 = -1 \text{ ps}^2/\text{km} \Rightarrow 0.8 \text{ ps/km}\cdot\text{nm}$) e collegamenti in cui la dispersione cromatica delle fibre impiegate vari lungo il percorso con valore medio uguale a zero.

A seconda dei sistemi presi in considerazione è stata utilizzata rivelazione diretta (tipo Intensity Modulation - Direct Detection, IM-DD) o rivelazione coerente (ASK). Nel caso di sistemi IM-DD, il ricevitore è costituito da un filtro ottico di Butterworth del secondo ordine con banda pari a $4R$ (dove R è la velocità di cifra), seguito da un quadratore e da un filtro elettrico di Butterworth del secondo ordine, con una banda uguale a $0.8R$. Nel caso di un sistema coerente ASK asincrono, il ricevitore è costituito da un sommatore, in cui il segnale di ingresso viene aggiunto al campo dell'oscillatore locale, seguito da un filtro di Butterworth del secondo ordine con una larghezza di banda pari a R , da un quadratore e da un filtro di Butterworth del secondo ordine con una banda pari a $0.8R$.

Nelle simulazioni il segnale di ingresso è costituito da una sequenza pseudocasuale di 32 bit. Nel caso del formato NRZ, il segnale è modulato in ampiezza con la forma d'onda di un coseno rialzato con tempo di salita⁽⁴⁾ pari a $T_r = 1/4R$ [19]. Nel caso dei solitoni, la forma del segnale è una secante iperbolica e la potenza di picco necessaria per mantenere le sue caratteristiche "naturalmente non dispersive" anche in presenza di perdite si ottiene modificando la (8) come segue [13]

$$P_k = \frac{\alpha L_a G \beta_2}{(G-1) \gamma \tau_s^2} \quad (14)$$

In presenza di effetto Kerr la probabilità d'errore può essere valutata solo per quelle condizioni di propagazione per le quali la statistica del rumore è nota analiticamente: assenza di ASE [19] oppure di dispersione [18]. In generale, le prestazioni sono state valutate in termini del fattore Q , definito come rapporto segnale rumore all'ingresso del circuito di decisione, tenendo in conto gli effetti dovuti alla successione degli uni e degli zeri [20]. Tale fattore è stato calcolato su 50 realizzazioni del processo, cambiando ogni volta il seme per la

(4) Intervallo di tempo necessario per passare dal 10% al 90% del valore massimo.

generazione dell'ASE. E' noto che un valore di Q uguale a 6 corrisponde, nell'ipotesi di rumore gaussiano, ad una probabilità di errore pari a 10^{-9} .

I parametri comuni a tutte le simulazioni sono: $\alpha=0.057 \text{ km}^{-1}$ (0.25 dB/km), $\beta_3=0.1 \text{ ps}^3/\text{km}$, $L_a=40 \text{ km}$ e $\gamma=2.7 \text{ (W}\cdot\text{km)}^{-1}$; gli altri parametri sono stati riportati nel testo.

Per migliorare le prestazioni si è considerato l'uso lungo la linea di filtri ottici, i quali, comunque, introducono delle perdite che debbono essere compensate dagli amplificatori ottici [14-17]. Tali filtri riducono l'interazione non lineare tra segnale ed ASE e nel caso dei sistemi a solitoni limitano il jitter temporale dovuto all'effetto Gordon-Haus [13-15] e le instabilità indotte dall'amplificazione periodica e dall'interazione non lineare tra gli impulsi [17].

I filtri ottici possono avere il picco di trasmissione ad una frequenza fissa o gradualmente variabile lungo il collegamento (filtri a scivolamento di frequenza). Quest'ultimo tipo di filtri, sebbene aumenti il grado di complessità del sistema, è particolarmente utile nei sistemi a solitoni per evitare la crescita esponenziale delle componenti di rumore; infatti lo spettro del solitone segue lo spostamento in frequenza del filtro, mentre il rumore ASE viene tagliato via.

3.1 Sistemi singolo canale

In fig. 3 sono riportate le prestazioni in termini del fattore Q per quattro sistemi ottici, con un formato di modulazione NRZ, in un collegamento lungo 9000 km considerando diversi profili di dispersione cromatica del collegamento con $L_a=40 \text{ km}$. Le linee continue si riferiscono ad un collegamento con fibra DS, mentre le linee tratteggiate ad un collegamento con un valore medio della dispersione uguale a zero. Il caso a) si riferisce ad un sistema NRZ ASK a 5 Gbit/s in assenza di filtri ottici, che sfrutta la compensazione non lineare della dispersione cromatica; il caso b) si riferisce allo stesso sistema del caso precedente, ma in presenza di filtri ottici lungo il collegamento con banda pari a 150 GHz; il caso c) ad un sistema NRZ IM-DD a 6 Gbit/s in presenza di dispersione cromatica variabile con distribuzione gaussiana a valore medio nullo e deviazione standard di $0.3 \text{ ps}^2/\text{km}$ (il valore della dispersione cromatica è variato ogni 20 km), mentre il caso d) rappresenta un sistema IM-DD a 10 Gbit/s in presenza di una dispersione cromatica con una distribuzione a dente di sega che è ottenuta alternando fibre DS con fibre SMF (il tratto delle fibre DS è lungo 300 km).

L'aumento del fattore Q con la potenza è dovuto ad un miglioramento del rapporto segnale rumore e, nei casi a) e b), alla compensazione non lineare della dispersione cromatica; quando si supera un certo livello di potenza gli effetti non lineari predominano e degradano il sistema.

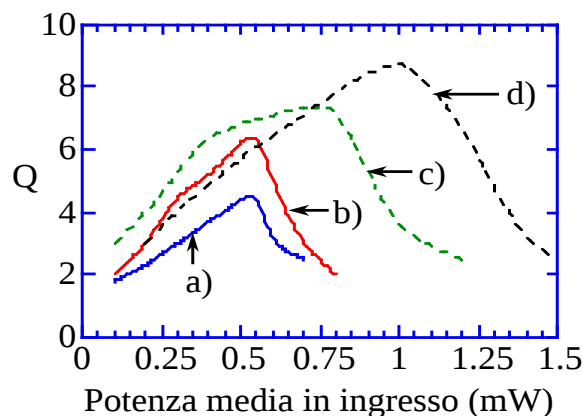


Figura 3 Fattore Q per quattro sistemi ottici, con un formato di modulazione NRZ, in un collegamento lungo 9000 km. I dettagli sono nel testo

La presenza dei filtri migliora le prestazioni del sistema perché l'interazione non lineare tra segnale e ASE è ridotta. Le prestazioni dei sistemi c) e d) sono migliori perché la dispersione cromatica è in media nulla, anche se è localmente diversa da zero e questo limita l'allargamento dello spettro consentendo di utilizzare un ricevitore a banda più stretta [21].

Un ulteriore miglioramento delle capacità su lunga distanza può essere ottenuto per mezzo dei solitoni, come si evince dalla fig. 4 in cui il fattore Q è riportato in funzione della potenza di ingresso normalizzata, P/P_k , in presenza di filtri ottici a scivolamento di frequenza (linee continue) e filtri ottici a frequenza fissa (linee tratteggiate) in un collegamento lungo 9000 km su fibra DS con $L_a=40 \text{ km}$. La distanza temporale tra i solitoni è pari a $\Delta T=6\tau_s$ per i sistemi a filtri fissi, mentre $5.5\tau_s$ per i sistemi con filtri a scivolamento di frequenza. La figura evidenzia l'importanza dei filtri a scivolamento di frequenza nelle trasmissioni a lunga distanza ad alta capacità.

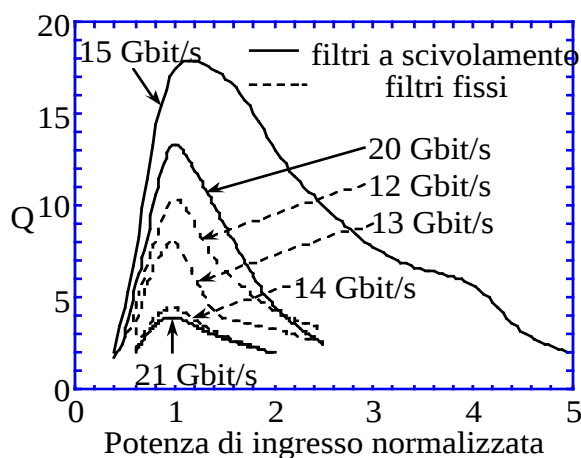


Figura 4 Fattore Q per sistemi a solitoni operanti in un collegamento lungo 9000 km con fibra DS

Un confronto ancora più diretto tra le prestazioni dei vari sistemi è stato riportato in fig. 5, che mostra la massima capacità di trasmissione in funzione della lunghezza del collegamento in una fibra DS; nella stessa figura sono stati anche inclusi due casi relativi a collegamenti con dispersione variabile a valore medio nullo.

Le capacità riportate sul grafico sono state ottenute ottimizzando i parametri del sistema (potenza di ingresso, durata degli impulsi) ed aumentando la velocità di trasmissione fino a raggiungere un valore di Q non inferiore a 6.

La curva a) si riferisce al caso puramente dispersivo, b) ad un sistema a solitoni in assenza di filtri, c) a un sistema a solitoni in presenza di filtri a frequenza fissa, d) a un sistema a solitoni in presenza di filtri a scivolamento di frequenza, e) a un sistema NRZ IM-DD dove la dispersione cromatica è parzialmente compensata dalla non linearità della fibra, f) a un sistema NRZ IM-DD con un collegamento a dispersione variabile con distribuzione gaussiana del tipo di quello già assunto in fig. 3, g) a un sistema NRZ IM-DD in un collegamento a dispersione variabile del tipo a dente di sega descritto in fig. 3. Nel caso e), per distanze superiori a 6000 km, sono stati utilizzati filtri ottici a frequenza fissa con una larghezza di banda 30R. Nel caso e) e g) sono stati utilizzati filtri ottici davanti al fotodiodo con una banda pari a 4R, mentre nel caso f) 10R.

Tutte le curve mostrano una diminuzione della

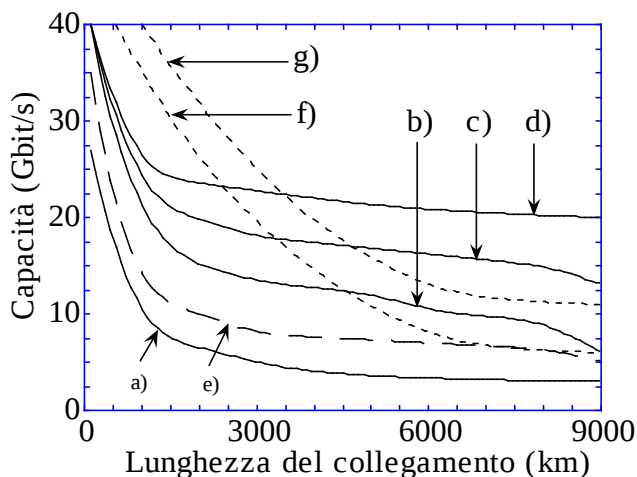


Figura 5 Massima capacità nel caso di sistemi singolo canale con fibra DS: a) caso puramente dispersivo, b) sistema a solitoni senza di filtri, c) sistema a solitoni con filtri a frequenza fissa, d) sistema a solitoni con filtri a scivolamento di frequenza, e) sistema NRZ IM-DD con dispersione cromatica parzialmente compensata dalla non linearità della fibra, f) sistema NRZ IM-DD con collegamento a dispersione variabile con distribuzione gaussiana, g) sistema NRZ IM-DD in collegamento a dispersione variabile del tipo a dente di sega

capacità all'aumentare della lunghezza del collegamento. Dal confronto delle curve b), c), d), in presenza di amplificazione periodica, si evidenzia il ruolo dei filtri che tra 2000 e 7000 km stabilizzano il solitone, mentre per lunghezze maggiori limitano l'effetto Gordon-Haus. La curva e) mostra un andamento simile alla curva c) anche se le capacità raggiunte sono più basse per il fatto che la compensazione non lineare della dispersione cromatica non è completa come nel caso solitonico.

I sistemi a dispersione variabile mostrano una capacità molto alta per distanze inferiori a 5000 km: tale tecnica, infatti, presenta il duplice vantaggio di avere un collegamento complessivamente a dispersione zero e di limitare fortemente l'aumento del rumore dovuto all'interazione non lineare tra segnale e rumore, essendo la dispersione localmente diversa da zero. Per lunghezze superiori a 5000 km gli effetti cumulativi dovuti all'interazione non lineare tra segnale e rumore e la distorsione indotta dall'effetto combinato della dispersione cromatica e dell'effetto Kerr riducono la capacità del sistema.

In fig. 6 è riportato lo stesso tipo di grafico relativo però a fibre SMF: in questo caso le curve sono riportate solo fino ad una distanza di 4000 km, a causa dell'alto valore della dispersione cromatica. La curva a) si riferisce al regime lineare puramente dispersivo, la curva b), c) e d) si riferiscono rispettivamente ad un sistema a solitoni in assenza di filtri, in presenza di filtri a frequenza fissa e di filtri a scivolamento di frequenza. Le curve e) e f)

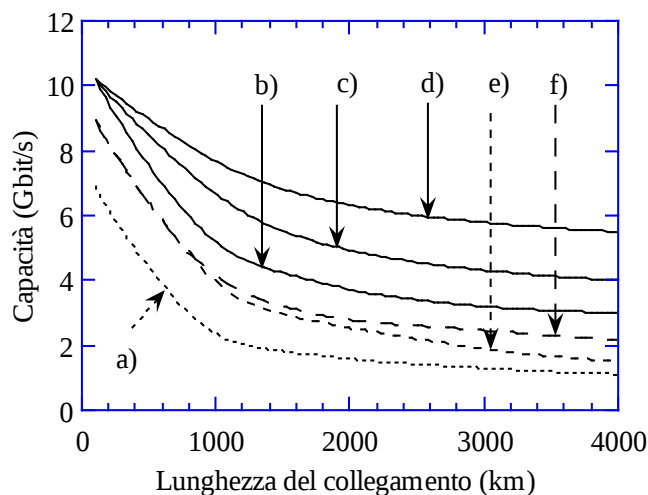


Figura 6 Massima capacità nel caso di sistemi singolo canale con fibra SMF. La curva a) si riferisce al regime lineare puramente dispersivo, la curva b), c) e d) si riferiscono rispettivamente ad un sistema a solitoni in assenza di filtri, in presenza di filtri a frequenza fissa e di filtri a scivolamento di frequenza. Le curve e) ed f) si riferiscono ad un sistema NRZ IM-DD dove la dispersione cromatica è parzialmente compensata dalla non linearità della fibra in assenza e in presenza di filtri a frequenza fissa (con banda 30 GHz)

ed f) si riferiscono ad un sistema NRZ IM-DD con dispersione cromatica parzialmente compensata dalla non linearità della fibra in assenza e in presenza di filtri a frequenza fissa con banda di 30 GHz. Confrontando le curve valide per collegamenti in fibre DS ed in fibre SMF si nota che il rapporto tra le massime capacità ottenibili nei due casi è inversamente proporzionale, in prima approssimazione, alla radice quadrata del rapporto delle rispettive dispersioni cromatiche, in accordo con la forma della (13) e, in particolare, della grandezza L_D . Infatti, per L_D costante, il modulo del coefficiente β_2 , relativo alla dispersione cromatica del secondo ordine, è proporzionale al quadrato della durata dell'impulso T_0 .

3.2 Sistemi WDM

Un altro modo di sfruttare la capacità trasmissiva della fibra ottica consiste nell'impiego di sistemi con più canali a differenti lunghezze d'onda (Wavelength Division Multiplexing, WDM). Le prestazioni dei sistemi WDM sono principalmente limitate dal FWM [10]. Come ricordato nella sezione precedente, l'efficienza di tale effetto diminuisce incrementando la spaziatura in frequenza tra i canali (Δf) o aumentando il valore della dispersione cromatica della fibra. L'effetto può essere ulteriormente ridotto adottando una spaziatura non uniforme dei canali, anche se in questo caso la banda ottica occupata può essere molto superiore rispetto al caso di canali uniformemente spaziati [22].

In linea di principio, la degradazione dovuta al FWM può essere eliminata mediante l'impiego di sistemi WDM a solitoni [23]. Infatti, in assenza di perdite, solitoni a differenti frequenze sono praticamente trasparenti gli uni agli altri, a parte uno spostamento temporale, dovuto alle collisioni, che può essere molto piccolo se i canali non sono troppo vicini in frequenza. Nel caso reale, la presenza di perdite e la necessaria amplificazione periodica introducono una asimmetria di guadagno durante la collisione, che determina uno spostamento in frequenza dei solitoni, che si traduce, a sua volta, in uno spostamento temporale. Questo effetto è molto critico se la collisione avviene per una lunghezza che sia confrontabile con la distanza a cui si trovano gli amplificatori. La lunghezza di collisione L_C è definita dall'espressione $L_C = 0.561 \tau_c / (\beta_2 \Delta f)$. Tale effetto può essere sensibilmente ridotto se vengono inseriti filtri ottici lungo il collegamento.

In fig. 7 il fattore Q è riportato per diversi sistemi WDM-ASK con segnali NRZ operanti su una distanza di 4000 km in fibra DS. La frequenza di cifra di ciascun canale è pari a 5 Gbit/s, le linee continue si riferiscono a sistemi equispaziati in frequenza (125 GHz) e N indica il numero di canali. La linea a tratteggio rappresenta un sistema con spaziatura in frequenza non uniforme: 60, 120, 72, 96, 108, 132, 84, 144, 156 GHz.

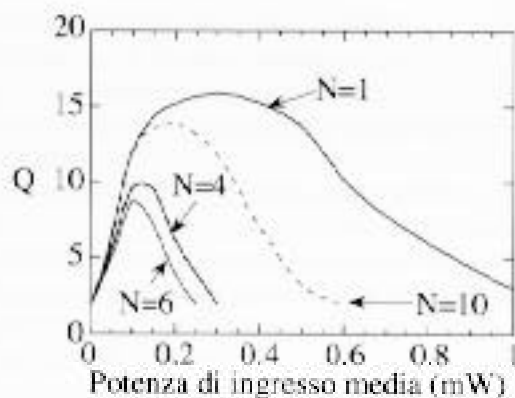


Figura 7

Fattore Q per differenti sistemi WDM-ASK con segnali NRZ in un collegamento lungo 4000 km in fibra DS. La frequenza di cifra di ciascun canale è pari a 5 Gbit/s. Le linee continue si riferiscono a sistemi equispaziati in frequenza, mentre le tratteggiate rappresentano una spaziatura non uniforme

Anche in questi sistemi il fattore Q cresce in funzione della potenza fino a che l'effetto Kerr, che in questo caso si manifesta principalmente come FWM, non induce un sensibile deterioramento del sistema.

In fig. 8 il fattore Q è riportato in funzione della spaziatura in frequenza tra i canali per un sistema a solitoni 2x5 Gbit/s. La curva a) è valutata in assenza di filtri ottici, la curva b) in presenza di filtri a frequenza fissa con banda di 120 GHz, mentre la curva c) in presenza di filtri a scivolamento di frequenza con banda pari a 100 GHz e uno spostamento in frequenza pari a 34 MHz ad ogni passo di amplificazione.

In assenza di filtri se la separazione in frequenza tra i canali assume valori tali che la lunghezza di collisione è confrontabile con la distanza tra gli amplificatori, le

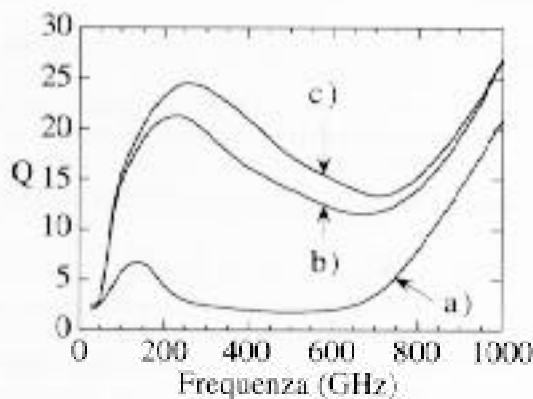


Figura 8

Fattore Q in funzione della spaziatura in frequenza per un sistema IM-DD a due canali solitonici con $R=5$ Gbit/s: a) senza filtri ottici; b) con filtri fissi e c) con filtri a scivolamento di frequenza

prestazioni del sistema sono notevolmente ridotte: dalla figura si può apprezzare come la presenza dei filtri ottici a banda stretta limiti questo effetto.

In maniera analoga a quanto fatto per i sistemi a singolo canale, nelle figg. 9 e 10 è riportata la massima capacità per sistemi ottici WDM in fibra DS e fibra SMF rispettivamente, considerando un fattore Q minimo pari a 6.

Nelle suddette figure, le curve a) e b) si riferiscono a sistemi singolo canale, riportate per confronto, mentre le curve da c) a g) a sistemi WDM. La curva a) rappresenta il limite teorico dovuto alla dispersione cromatica e b) si riferisce ai sistemi a solitoni con filtro a scivolamento di frequenza. La curva c) si riferisce a sistemi WDM-ASK-NRZ con spaziatura uniforme, d) a sistemi WDM-ASK-NRZ con spaziatura non uniforme, e) a sistemi WDM-IM-DD a solitoni con spaziatura uniforme in assenza di filtri, f) a sistemi WDM a solitoni con spaziatura uniforme in presenza di filtri a frequenza fissa e g) a sistemi WDM a solitoni con spaziatura uniforme in presenza di filtri a scivolamento di frequenza. Nel caso dei sistemi WDM, i singoli canali operano a 5 Gbit/s nei casi di fig. 9, mentre 1 Gbit/s nei casi di fig. 10.

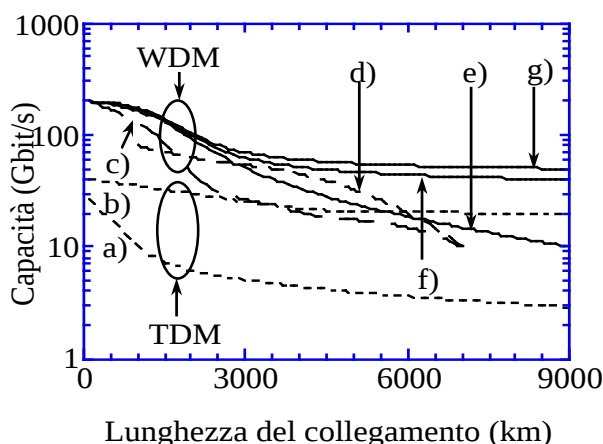


Figura 9 Massima capacità per sistemi ottici operanti in collegamenti con fibra DS. Le curve a) e b) si riferiscono ai sistemi singolo canale ed in particolare la curva a) rappresenta il limite teorico dovuto alla dispersione cromatica, mentre b) si riferisce ai sistemi a solitoni con filtro a scivolamento di frequenza. Le curve da c) a g) si riferiscono a sistemi WDM. La curva c) si riferisce ai sistemi WDM-ASK-NRZ con spaziatura uniforme, d) ai sistemi WDM-ASK-NRZ con spaziatura non uniforme, e) ai sistemi WDM-IM-DD a solitoni con spaziatura uniforme in assenza di filtri, f) ai sistemi WDM a solitoni con spaziatura uniforme in presenza di filtri a frequenza fissa e g) ai sistemi WDM a solitoni con spaziatura uniforme in presenza di filtri a scivolamento di frequenza. Nel caso dei sistemi WDM, i singoli canali operano a 5 Gbit/s

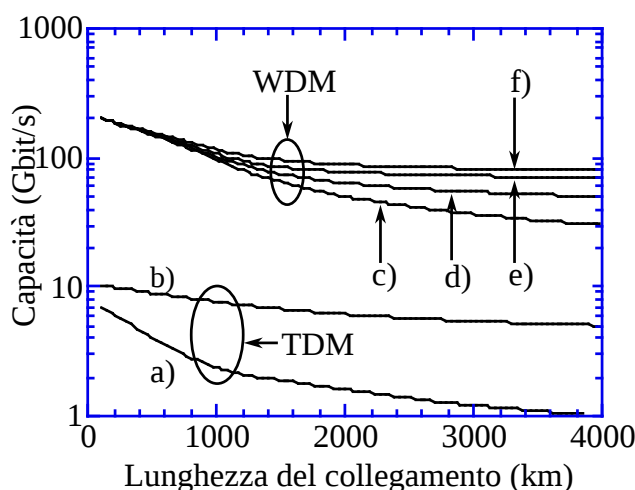


Figura 10 Massima capacità per sistemi ottici operanti in collegamenti con fibra SMF. I casi sono gli stessi della figura 9, tranne che nei sistemi WDM il bit rate dei canali è 1 Gbit/s

3.3 Esame dei risultati

I risultati presentati mostrano come nelle applicazioni singolo canale l'uso delle fibre DS permette di ottenere un fattore di merito capacità x lunghezza maggiore di quello ottenuto per fibre SMF; inoltre le migliori prestazioni sono ottenute con i solitoni. Nei sistemi WDM l'efficienza del FWM è ridotta in presenza di fibre SMF e per distanze transoceaniche i sistemi WDM a solitoni presentano le prestazioni migliori mentre, nel caso di medie distanze (2000-4000 km), i sistemi convenzionali WDM con spaziatura non uniforme possono fornire capacità confrontabili con quelle dei sistemi WDM a solitoni. Per distanze inferiori ai 1000 km, l'effetto del FWM risulta meno dannoso e la tecnica WDM convenzionale risulta più efficiente con spaziatura uniforme che con spaziatura non uniforme, poiché viene occupata meno banda ottica.

Un esame più completo deve comunque considerare la particolare applicazione cui si sta pensando: occorre distinguere tra i requisiti dei collegamenti transoceanici, realizzati con cavi sottomarini lunghi dai 4000 ai 9000 km, e i collegamenti in una futura rete di trasporto completamente ottica, che uniscono le principali città di una nazione o di un continente con lunghezze tra i 500 e i 4000 km.

Nel caso dei collegamenti transoceanici si tratta sempre di sistemi punto-punto e la richiesta principale è quella di trasmettere la massima capacità. Dai risultati già mostrati nelle figg. 5 e 9, si nota come capacità comprese tra 10 Gbit/s e 50 Gbit/s possano essere raggiunte con particolari accorgimenti tecnologici.

Nel caso dei collegamenti di terra possono esserci altri requisiti: uno scenario verosimile è quello nel quale nei nodi sono richieste operazioni di add/drop per flussi

compresi tra 155 Mbit/s e 10 Gbit/s, mentre il flusso complessivo tra due nodi potrebbe essere dell'ordine dei 40 Gbit/s. Da quanto mostrato nelle figg. 9 e 10 prestazioni adeguate possono essere ottenute mediante sistemi WDM con segnali NRZ e ricevitori ASK. Le tabb. 1 e 2 (rispettivamente per fibre DS e SMF) presentano dei risultati sulla possibilità di raggiungere capacità di 40 Gbit/s mediante sistemi WDM-ASK con segnali NRZ, considerando diverse granularità del sistema. Nelle tabelle sono riportati, per ogni distanza e per ogni configurazione del sistema, i massimi valori del fattore Q valutato in un intervallo di potenza media tra 0.1 e 2 mW. Le celle barrate indicano l'impossibilità di funzionamento del sistema, mentre le celle ombreggiate la necessità di una limitata tolleranza nella

Fibre DS

40 Gbit/s	500 km	1000 km	2000 km	3000 km	4000 km
256x155 Mbit/s	—	—	—	—	—
64x622 Mbit/s	—	—	—	—	—
16x2.5 Gbit/s	45.2	21.2	16.1	10.9	7.8
4x10 Gbit/s	16.1	6.8	—	—	—
40 Gbit/s	—	—	—	—	—

Tabella 1 Massimo fattore Q per differenti sistemi WDM-ASK-NRZ con capacità complessiva di 40 Gbit/s nel caso di fibre DS. Le celle barrate indicano l'impossibilità di funzionamento del sistema, mentre le celle ombreggiate la necessità di una limitata tolleranza nella potenza di lancio

Fibre standard

40 Gbit/s	500 km	1000 km	2000 km	3000 km	4000 km
256x155 Mbit/s	—	—	—	—	—
64x622 Mbit/s	39.9	19.8	13.1	10.0	7.6
16x2.5 Gbit/s	15.7	—	—	—	—
4x10 Gbit/s	—	—	—	—	—
40 Gbit/s	—	—	—	—	—

Tabella 2 Come in tabella 2 ma per fibra SMF

potenza di lancio: dai risultati si nota come i 40 Gbit/s su distanze di 4000 km possono essere raggiunti solo mediante sistemi del tipo 16x2.5 Gbit/s nel caso di fibre DS e 64x622 Mbit/s nel caso di fibre SMF [8].

Nelle tabb. 3 e 4 sono mostrate invece le diverse possibilità per raggiungere una capacità complessiva di 10 Gbit/s. Tali tabelle mostrano che i 10 Gbit/s su una distanza di 4000 km possono essere raggiunti con sistemi del tipo 4x2.5 Gbit/s o 16x622 Mbit/s nel caso di fibre DS e 16x622 Mbit/s o 64x155 Mbit/s nel caso di fibre SMF.

Fibre DS

10 Gbit/s	500 km	1000 km	2000 km	3000 km	4000 km
64x155 Mbit/s	>50	13.8	7.8	—	—
16x622 Mbit/s	>50	21.6	14.8	11.2	9.2
4x2.5 Gbit/s	>50	28.4	21.6	14.0	10.5
10 Gbit/s	>50	33.8	14.9	12.2	8.2 *

Tabella 3 Massimo fattore Q per differenti sistemi WDM-ASK-NRZ con capacità complessiva di 10 Gbit/s nel caso di fibre DS. Le celle barrate indicano l'impossibilità di funzionamento del sistema, mentre le celle ombreggiate la necessità di una limitata tolleranza nella potenza di lancio

Fibre standard

10 Gbit/s	500 km	1000 km	2000 km	3000 km	4000 km
64x155 Mbit/s	>50	18.3	13.4	10.1	8.6
16x622 Mbit/s	>50	25.5	16.8	13.2	10.4
4x2.5 Gbit/s	16.8	—	—	—	—
10 Gbit/s	—	—	—	—	—

Tabella 4 Come in tabella 3 ma per fibra SMF

4. Effetti dovuti alla dispersione di polarizzazione

Come anticipato nella sezione 2, un effetto che potrebbe limitare i sistemi ad alta capacità è la dispersione di polarizzazione (DP) [24].

Uno dei maggiori problemi connessi alla DP è la sua aleatorietà ed in particolare la sua fluttuazione temporale in funzione delle condizioni ambientali: una eventuale

compensazione risulta quindi molto complessa.

Recentemente, lavori teorici e soprattutto sperimentali hanno mostrato che la DP può essere compensata nella propagazione dei segnali solitonici, in quanto l'effetto Kerr induce una attrazione tra i due modi di birifrangenza della fibra [25] che si contrappone alla separazione che sarebbe indotta dalla DP.

Nel seguito si mostrerà che la DP può essere compensata mediante l'effetto Kerr anche nel caso di segnali NRZ [26].

Il principio su cui è basata la compensazione della DP si trova nella fig. 11, che mostra l'evoluzione delle

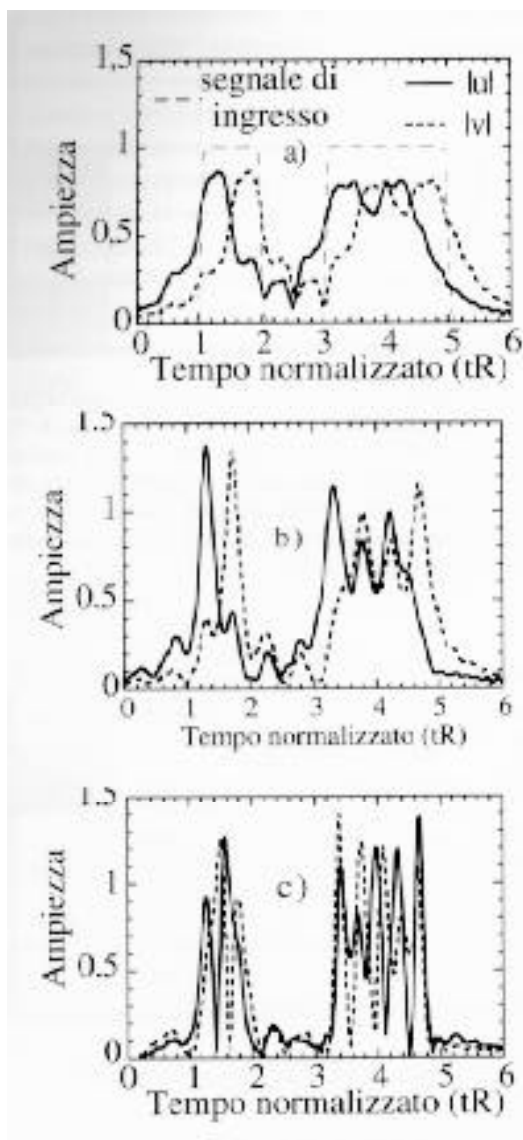


Figura 11 Evoluzione di un segnale NRZ a 5 Gbit/s in un collegamento con birifrangenza uniforme con $\Delta\beta'=0.025$ ps/km e $\beta_2=-1$ ps²/km. a) propagazione in condizioni lineari; b) propagazione in presenza di effetto Kerr per una potenza di ingresso pari a 0.54 mW; c) presenza di effetto Kerr per una potenza di ingresso pari a 0.75 mW

componenti del campo di un segnale NRZ con una velocità di cifra di 5 Gbit/s in una fibra con birifrangenza uniforme e $\Delta\beta'=0.025$ ps/km. Si considera propagazione lungo 5000 km in condizioni ideali di assenza di perdite e fibra SMF. Il caso a) rappresenta la propagazione in condizioni lineari, mentre i casi b) e c) la propagazione in presenza di effetto Kerr per una potenza di ingresso pari rispettivamente a 0.54 mW e 0.75 mW. Nel caso a) è chiaramente visibile l'effetto della separazione degli impulsi dovuto alla DP e l'allargamento dovuto alla dispersione cromatica; nel caso b) si vede come l'effetto Kerr produca un sensibile restringimento degli impulsi [27], ma solo una debole compensazione della DP, mentre nel caso c) un lieve innalzamento della potenza permette un'attrazione tra le componenti degli impulsi e quindi una buona compensazione della DP.

L'effetto Kerr compensa la DP anche nel caso reale di presenza di accoppiamento random e perdite nella fibra: nella fig. 12 le componenti del campo sono riportate all'uscita di un collegamento lungo 5000 km con $\Delta\beta'=0.025$ ps/km, perdite pari a 0.25 dB/km e un accoppiamento random con $L_c=500$ m (ciò corrisponde ad una DP di 0.22 ps/ $\sqrt{\text{km}}$). Per compensare le perdite

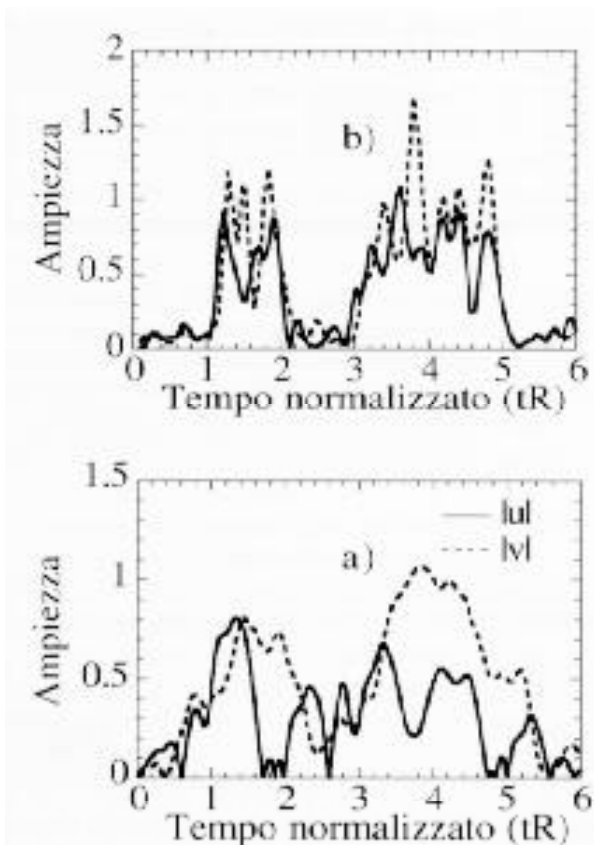


Figura 12 Evoluzione di un segnale NRZ a 5 Gbit/s in un collegamento lungo 5000 km con una DP pari a 0.22 ps/ $\sqrt{\text{km}}$. a) propagazione in condizioni lineari e in assenza di DVG; b) propagazione in presenza di effetto Kerr per una potenza di ingresso pari a 1 mW e con una DVG pari a $\beta_2=-1$ ps²/km

sono stati inseriti amplificatori ottici ogni 50 km, senza includere in questi primi risultati l'ASE. La fig. 2a mostra la propagazione in assenza di effetto Kerr e DVG rendendo evidente la degradazione prodotta dalla DP sul segnale, mentre la fig. 2b introduce l'effetto Kerr e una DVG pari a $-1 \text{ ps}^2/\text{km}$, per una potenza di ingresso pari a 1 mW. L'effetto della compensazione della DP è chiaramente visibile.

In fig. 13 il comportamento di un sistema IM-DD è riportato in termini del fattore Q in funzione della potenza di ingresso, in un collegamento con le stesse caratteristiche della fig. 12. La curva a) si riferisce ad un collegamento con fibre SMF e assenza di ASE, la curva b) alla assenza di DVG e di ASE, mentre la curva c) è come la curva a) ma in presenza di ASE. Nel caso a) l'aumento del fattore Q è dovuto all'effetto Kerr che compensa la degradazione indotta dalla DVG. Per una potenza compresa tra 0.8 e 1.1 mW la figura mostra un ulteriore miglioramento dovuto al fatto che l'interazione tra DVG e effetto Kerr permette di compensare la DP. La necessità della presenza della DVG per la compensazione della DP è confermata dalla curva b) che mostra come l'incremento della potenza non induce alcun miglioramento in termini di Q . La curva c) mostra come la presenza dell'ASE diminuisce ma non annulla la compensazione.

Nel caso dei solitoni si è visto che la DP può essere compensata quando è verificata la seguente relazione

$$\Delta\beta' \sqrt{L_c} < A \sqrt{\frac{\beta_2}{1.28}} \quad (15)$$

dove L_c è espressa in km e A è una funzione del rapporto L/L_D . Nel riferimento [28] si è trovato che A valeva 0.3 per $L/L_D = 57$. Le simulazioni numeriche hanno mostrato che la (15) è valida anche per segnali NRZ e la fig. 14

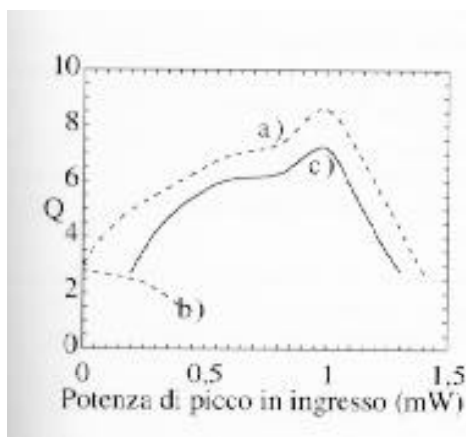


Figura 13 Fattore Q per un sistema IM-DD NRZ a 5 Gbit/s. Il collegamento è lungo 5000 km, con una distanza tra amplificatori pari a 50 km ed una DP pari a $0.22 \text{ ps}^2/\text{km}$. a) $\beta_2 = -1 \text{ ps}^2/\text{km}$ in assenza di ASE; b) in assenza di DVG e ASE; c) come la curva a) ma in presenza di ASE

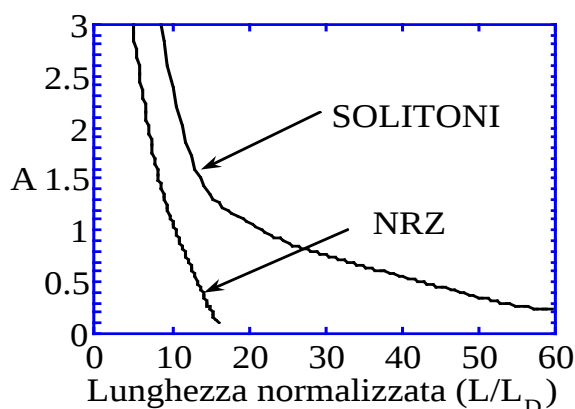


Figura 14 Andamento di A in funzione della lunghezza normalizzata L/L_D

riporta un confronto per la funzione A nei due casi. I valori sono stati trovati considerando il massimo valore del parametro $\Delta\beta' \sqrt{L_c}$ che permette di avere un Q minimo pari a 6. La figura mostra chiaramente come nel caso dei solitoni la compensazione non lineare della DP è più efficace rispetto ai segnali NRZ.

I risultati di questa sezione, che si riferiscono ad un collegamento lungo 5000 km, si possono facilmente estendere ad altri casi: in particolare, un impiego molto interessante potrebbe riguardare i sistemi a 10 Gbit/s in collegamenti con fibre DS con lunghezze comprese tra i 100 e i 300 km, in cui la DP può variare dai 10 ai 50 ps. Tali sistemi potrebbero funzionare senza altra tecnica di compensazione che una opportuna scelta della potenza di lancio, che rimane comunque nei valori degli attuali sistemi di trasmissione.

5. Conclusioni

In questo lavoro è stata analizzata la propagazione di impulsi in fibra tenendo conto degli effetti della dispersione cromatica, della dispersione di polarizzazione, della non linearità di tipo Kerr della fibra e del rumore ottico di amplificazione sulla distorsione del segnale.

Sebbene diversi sistemi di trasmissione abbiano raggiunto la maturità delle dimostrazioni sperimentali, è necessario un ulteriore sforzo per individuare il sistema migliore a seconda delle esigenze del traffico e proprio in questa ottica si è cercato di presentare un confronto tra alcuni dei più significativi sistemi proposti.

Per sistemi a singolo canale capacità elevate possono essere ottenute o con i solitoni in collegamenti con fibre operanti in regione anomala o con segnali NRZ in collegamenti con dispersione cromatica variabile con valore medio nullo uguale a zero.

I solitoni sono anche indispensabili per ottenere capacità elevate con sistemi WDM operanti su lunghissime distanze, mentre su medie distanze anche i

segnali NRZ permettono di ottenere capacità elevate, particolarmente utilizzando una spaziatura non uniforme. La differenza tra la capacità offerta dai sistemi WDM rispetto ai sistemi singolo canale è più evidente nelle fibre SMF.

Per ultimo si è visto come l'effetto Kerr possa compensare la dispersione di polarizzazione sia nel caso di segnali di tipo solitonico che NRZ.

Acronimi

ASE	Amplified Spontaneous Emission
ASK	Amplitude Shifting Keying
ASST	Azienda di Stato Servizi Telefonici
CPM	Cross Phase Modulation
DP	Dispersione di Polarizzazione
DS	Dispersion Shifted fiber
DVG	Dispersione della Velocità di Gruppo
FWM	Four Wave Mixing
IM-DD	Intensity Modulation-Direct Detection
NRZ	Non-Return-to-Zero
PCM	Pulse Code Modulation
SMF	Single Mode Fiber
SPM	Self Phase Modulation
WDM	Wavelength Division Multiplexing

Bibliografia

- [1] Taga, H.; Edagawa, N.; Tanaka, H.; Suzuki, M.; Yamamoto, S.; Watabayashi, H.; Bergano, N.S.; Davidson, C.R.; Homsey, G.M.; Kalmus, D.J.; Trischitta, P.R.; Gray, D.A.; Maybach, R.L.: *10 Gbit/s, 9000 km IM-DD Transmission Experiment using 274 Er-Doped Fibre Amplifiers Repeaters*. OFC/IOOC '93, San Jose (CA), February 21-26, PD1-1, pp. 9-12.
- [2] Mollenauer, L.F.; Mamyshev, P.V.; Neubelt, M.J.: *Measurements of timing jitter in filter-guided soliton transmission at 10 Gbit/s and achievement of 375 Gbit/s-Mm, error free, at 12.5 and 15 Gbit/s*. «Optics Letters», Vol. 19, 1994, pp. 704-706.
- [3] Agrawal, G.P.: *Non linear fibre optics*. Academic Press Inc., San Diego, 1989.
- [4] Elrefaie, A.F.; Wagner, R.E.; Atlas D.A.; Daut, D.G.: *Chromatic Dispersion Limitation in Coherent Lightwave Transmission Systems*. «J. of Lightwave Technology», Vol. 6, 1988, pp. 704-709.
- [5] Marcuse, D.: *Single-Channel Operation in Very Long Non linear Fibers With Optical Amplifiers at Zero Dispersion*. «J. of Lightwave Technology», Vol. 9, 1991, pp. 356-361.
- [6] De Marchis, G.; Matera, F.; Settembre M.: *Comparison of performance of long-haul high-capacity optical systems*. «El. Letters», Vol. 29, 1993, pp. 1777-1778.
- [7] Matera, F.; Settembre, M.: *Exploitation of optical fibre capacity in long links*. «Electronics Letters», Vol. 30, 1994, pp. 803-804.
- [8] Matera, F.; Settembre, M.; De Marchis, G.: *Study on the implementation of TDM and WDM optical systems up to 40 Gbit/s x 4000 km*. Proc. of ECOC '94, Firenze, September 25-29 1994, pp. 833-836.
- [9] Chraplyvy, A.R.: *Limitations on Lightwave communications imposed by optical-fiber non linearities*. «J. of Lightwave Technology», Vol. 8, 1990, pp. 1548-1557.
- [10] Shibata, N.; Braun R.P.; Waarts, R.G.: *Phase-Mismatch Dependence of Efficiency of Wave Generation Through Four-Wave Mixing in a Single-Mode Optical Fiber*. «IEEE J. of Quantum Electronics», Vol. QE-23, 1987, pp. 1205-1210.
- [11] Matera, F.; Settembre, M.: *Non linear compensation of chromatic dispersion for phase- and intensity-modulated signals in the presence of amplified spontaneous emission noise*. «Optics Letter», Vol. 19, 1994, pp. 1198-1200.
- [12] Gordon, J.P.; Haus, H.A.: *Random Walk of Coherently Amplified Solitons in Optical Fiber Links*. «Optics Letters», Vol. 11, 1986, pp. 665-667.
- [13] Marcuse, D.: *Simulations to Demonstrate Reduction of the Gordon-Haus Effect*. «Optics Letters», Vol. 17, 1992, pp. 34-36.
- [14] Mecozzi, A.; Moores, J.D.; Haus, H.A.; Lai, Y.: *Solitons Transmission Control*. «Optics Letters», Vol. 16, 1991, pp. 1841-1843.
- [15] Kodama, Y.; Hasegawa, A.: *Generation of Asymptotically Stable Optical Solitons and Suppression of the Gordon-Haus Effect*. «Optics Letters», Vol. 17, 1992, pp. 31-33.
- [16] Mollenauer, L.F.; Gordon, J.P.; Evangelides S.G.: *The sliding-frequency guiding filter: an improvement form of soliton jitter control*. «Optics Letters», Vol. 17, 1992, pp. 1575-1577.
- [17] Romagnoli, M.; Wabnitz, S.: *Bandwidth limits of soliton transmission with sliding filters*. «Optics Communications», Vol. 104, 1994, pp. 293-297.
- [18] Mecozzi, A.: *Long-distance transmission at zero dispersion: combined effect of Kerr non linearity and the noise of the in-line amplifiers*. «J. of Opt. Soc. of Am. B», Vol. 11, 1994, pp. 462-469.
- [19] Iannone, E.; Locati, F.S.; Matera, F.; Romagnoli, M.; Settembre, M.: *High Speed DPSK Coherent Systems in the Presence of Chromatic Dispersion and Kerr Effect*. «J. of Lightwave Technology», Vol. 11, 1993, pp. 1478-1485.

- [20] Anderson, C.J.; Lyle, J.A.: *Technique for evaluating system performance using Q factor in numerical simulations exhibiting intersymbol interference*. «Electronics Letters», Vol. 30, 1994, pp. 71-72.
- [21] Lichtman, E.; Evangelides, S.G.: *Reduction of the non linear impairment in ultralong lightwave systems by tailoring the fibre dispersion*. «Electronics Letters», Vol. 30, 1994, pp. 348-346.
- [22] Forghieri, F.; Tkach, R.W.; Chraplyvy, A.R.; Marcuse, D.: *Reduction of four-wave mixing cross talk in WDM systems using unequally spaced channels*. Proc. OFC '93, San Jose (Ca), February 21-26, pp. 252-253.
- [23] Mollenauer, L.F.; Evangelides, S.G.; Gordon, J.P.: *Wavelength Division Multiplexing with Solitons in Ultra-Long Distance Transmission Using Lumped Amplifiers*. «J. of Lightwave Technology», Vol. 9, 1991, pp. 362-367.
- [24] Curti, F.; Daino, B.; De Marchis, G.; Matera F.: *Statistical Treatment of the Polarization Mode Dispersion in Long Single-Mode Fibers*. «J. of Lightwave Technology», Vol.8, 1990, pp. 1162-1166.
- [25] Evangelides, S.G.; Mollenauer, L.F.; Gordon, J.P.; Bergano, N.S.: *Polarization Multiplexing with Solitons*. «J. of Lightwave Technology», Vol. 10, 1992, pp. 28-35.
- [26] Matera, F.; Settembre, M.: *Compensation of the Polarization Mode Dispersion by means of the Kerr effect for Non Return to Zero Signals*. «Optics Letters», Vol. 20, 1995, pp.28-30.
- [27] Iannone, E.; Locati, F.S.; Matera, F.; Romagnoli, M.; Settembre, M.: *Non linear Evolutions of ASK and PSK signals in repeaterless fibre links*. «Electronics Letters», Vol. 28, 1992, pp. 1902-1903.
- [28] Mollenauer, L.F.; Smith, K.; Gordon, J.P.; Menyuk, C.R. «Optics Letters», Vol. 14, 1989, pp. 1219-1221.
- [29] Di Vita, P. et al.: *Fibre ottiche per Telecomunicazioni: fibre monomodali*. "Notiziario Tecnico SIP", Vol. 2, n. 3, Dicembre 1993.

Tecnologie ottiche nello strato di trasporto della rete di Telecomunicazione⁽¹⁾

E. Iannone, R. Sabella (*)

Sotto la spinta dell'introduzione di nuovi servizi e dell'aumento della richiesta dei servizi tradizionali, si prevede che la capacità di trasporto richiesta alla rete di telecomunicazione debba crescere rapidamente nell'immediato futuro. Le modalità di questa crescita non sono facilmente prevedibili, per cui è necessario approntare una infrastruttura di rete flessibile che permetta di modulare gli investimenti in base alle reali esigenze. Una tale infrastruttura può essere ottenuta mediante l'introduzione nello strato di trasporto della rete di un livello di cammino ottico, per la cui realizzazione i ripartitori ottici sono elementi fondamentali. In questo lavoro si analizza un percorso evolutivo per l'introduzione dello strato di cammino ottico nella rete del trasporto, e si descrivono le tecnologie che, allo stato dell'arte, rendono questa evoluzione realistica. Infine si analizza la capacità dello strato di cammino ottico sotto un punto di vista trasmissivo.

1. Introduzione

Dall'introduzione delle prime fibre ottiche a bassa perdita negli anni '70, i sistemi di trasmissione ottica si sono rapidamente evoluti grazie all'introduzione continua di nuove tecnologie ed al raffinamento di quelle già esistenti. L'ultimo passo in avanti in ordine di tempo è stato costituito dalla introduzione di amplificatori ottici di elevate prestazioni e con caratteristiche di affidabilità tali da essere utilizzabili in sistemi di trasmissione in esercizio.

L'evoluzione dei sistemi di trasmissione in fibra ottica è schematicamente riportata in fig. 1 [1,2]. Tale evoluzione è chiaramente descritta dalla capacità di trasmissione, crescente nel tempo, di sistemi realizzati in versione prototipale e commerciale.

In seguito a questo sviluppo tecnologico i sistemi ottici di trasmissione sono penetrati sempre più profondamente nella rete di comunicazione, tanto che attualmente tutti i collegamenti trasmissivi nella rete di trasporto e di giunzione sono realizzati in fibra ottica e si sta attuando l'introduzione delle fibre nella rete di distribuzione. Di fatto, l'unico sistema trasmissivo nella rete di telecomunicazione che non sarà realizzato

interamente in fibra entro i prossimi dieci anni in tutti i paesi tecnologicamente avanzati è la linea d'utente.

Alla penetrazione delle tecnologie ottiche nella funzione trasmissiva della rete però, non è corrisposta una analoga penetrazione nelle funzioni di instradamento, dove l'elettronica digitale mette a disposizione sistemi di

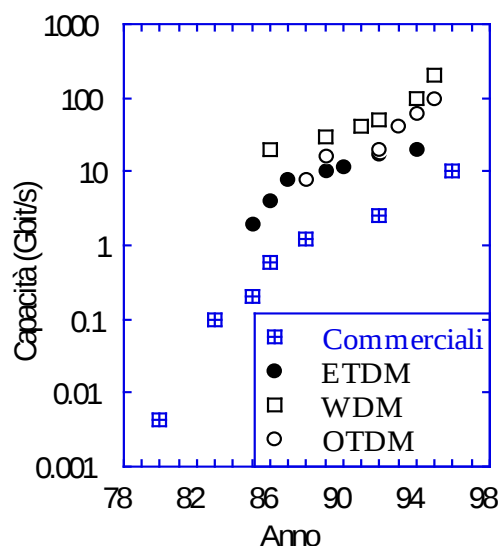


Figura 1

Evolutione della capacità trasmissiva dei sistemi in fibra ottica. ETDM=Sistemi singolo canale con moltiplicazione a divisione di tempo elettrica, OTDM=Sistemi con moltiplicazione a divisione di tempo ottica, WDM=Sistemi multicanale moltiplicati a divisione di frequenza

(*) Dott. Eugenio Iannone -Fondazione Ugo Bordon- Roma; ing. Roberto Sabella -Ericsson Telecomunicazioni- Roma

(1) Lavoro eseguito nell'ambito della convenzione tra la Fondazione Ugo Bordon e Telecom Italia.

elaborazione del segnale di alte prestazioni, affidabili ed economicamente convenienti. Ad esempio, nelle attuali reti del trasporto, mentre la trasmissione è interamente realizzata con sistemi ottici, l'instradamento e la commutazione sono effettuati da ripartitori digitali elettrici (DXC: Digital Cross-Connect) di grande capacità.

Sotto la spinta dell'introduzione di nuovi servizi, si prevede che la capacità della rete di trasporto debba crescere rapidamente nell'immediato futuro. I sistemi ottici di trasmissione assicurano una enorme potenzialità di crescita (ad esempio con le tecnologie attuali sono possibili collegamenti di 300-500 km senza ripetitori elettrici con capacità superiore a 100 Gbit/s [3,4]) cosicché il collo di bottiglia della rete di trasporto diviene il DXC. Al crescere della capacità della rete infatti, la complessità del nodo elettrico cresce rapidamente fino a rendere proibitivi sia i costi di sviluppo che quelli di manutenzione [5].

D'altra parte, negli ultimi anni, la ricerca nel campo dei dispositivi e dei sistemi ottici si è in parte spostata dalla trasmissione all'elaborazione di segnale, mettendo a disposizione nuovi dispositivi e nuove tecniche abbastanza affidabili da poterne prevedere un impiego nella rete. Si apre allora la possibilità di estendere le tecnologie ottiche dal semplice livello trasmissivo a strati logici superiori della rete. In questo modo si possono progettare ripartitori ottici di segnale (Optical Cross-Connect: OXC) di grande capacità e complessità molto minore dei loro analoghi elettronici [6,7,8].

In questo lavoro si analizza un percorso evolutivo possibile per l'introduzione di tecnologie ottiche nello strato di cammino della rete del trasporto, e si individuano le tecnologie chiave che possono permettere questa introduzione in modo affidabile ed economicamente conveniente.

Particolare attenzione saranno riservate alla struttura ed alle funzionalità del OXC ed alla conversione ottica di frequenza che gioca un ruolo centrale nell'elaborazione del segnale in una rete multiplata a divisione di frequenza.

2. Tecnologie Ottiche nello Strato di Cammino

La struttura delle architetture di rete nell'ambito delle reti pubbliche di telecomunicazione è stata discussa nelle sedi internazionali di standardizzazione ITU-T. In particolare è stata definita una struttura a strati e lo strato di trasporto è stato diviso in ulteriori tre strati [9]: di circuito, di cammino e di mezzo fisico di trasmissione (in breve, strato fisico), come illustrato in fig. 2. Lo strato di circuito è preposto a stabilire o rilasciare dinamicamente o attraverso "provisioning" a breve termine connessioni "end-to-end". Queste connessioni sono dedicate a servizi specifici come, ad esempio, la telefonia pubblica o la comunicazione dati a pacchetti. Lo strato fisico gestisce le interconnessioni fisiche tra i

nodi di rete, cioè i singoli sistemi trasmissivi che fisicamente trasferiscono l'informazione attraverso la rete. Lo strato di cammino fa da ponte tra i due strati precedenti e gioca un ruolo molto importante per quel che riguarda la realizzazione di reti flessibili ed affidabili, fornendo una piattaforma comune ai diversi servizi: nell'evoluzione della rete; i DXC costituiscono le principali entità appartenenti allo strato di cammino. Infine, l'accesso alla rete di trasporto avviene attraverso la rete di accesso, che è indicata in fig. 2 con una ellisse grigia.

La struttura a strati permette di semplificare la progettazione, lo sviluppo e le operazioni della rete, e ne consente l'evoluzione ed il potenziamento graduale. Ogni strato può evolvere indipendentemente, capitalizzando l'introduzione di nuove tecnologie indipendentemente dal modo in cui sono realizzate le entità di rete appartenenti ad altri strati.

Sino ad ora la tecnologia ottica è stata utilizzata per aumentare le capacità trasmissive dei collegamenti; quindi, nella schematizzazione di cui sopra, è stata confinata allo strato fisico. Inoltre l'affasciamento dei flussi di dati provenienti dalla rete di giunzione a formare i flussi ad alta velocità trasmessi attraverso la rete del trasporto è effettuato con tecniche TDM così da realizzare uno strato di cammino della rete integrato nelle tecniche (multiplicazione TDM e commutazione nel dominio del tempo).

Lo sfruttamento efficiente delle capacità trasmissive della fibra ottica però, con le tecnologie attuali, si può realizzare più facilmente utilizzando la moltiplicazione di frequenza (WDM: Wavelength Division Multiplexing) piuttosto che la moltiplicazione TDM. Inoltre la moltiplicazione WDM permette di utilizzare tecniche e dispositivi ottici anche per realizzare le operazioni di instradamento e di commutazione, così da introdurre l'uso delle tecnologie ottiche nello strato di cammino.

Poiché lo strato di circuito non può che restare essenzialmente elettrico allo stato attuale della tecnologia, l'impiego della tecnologia ottica nello strato

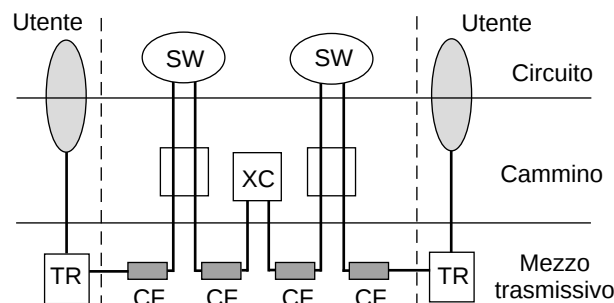


Figura 2

Struttura a strati della rete di trasporto e principali entità di ogni strato: SW=Software, XC=Ripartitore (Cross-Connect), TR=terminale di Rete, CF=Collegamento in Fibra ottica

di cammino porta naturalmente ad una ulteriore suddivisione di quest'ultimo in due livelli: uno elettrico, che offre servizi allo strato di circuito, ed uno ottico. Quest'ultimo, interfacciato direttamente con lo strato di mezzo trasmissivo, provvede all'instradamento di interi flussi di informazione ad alta velocità con tecnologie completamente ottiche. Quando questi flussi ad alta velocità devono essere demultiplati ed elaborati a velocità più bassa, essi sono trasferiti al livello elettrico a DXC in grado di realizzare tutte le necessarie funzionalità [6,8].

Perché lo strato ottico del cammino risulti integrato nelle tecniche sarà necessario realizzare, insieme alla multiplazione WDM, l'instradamento in base alla frequenza. Due delle tecniche più promettenti per realizzare l'instradamento in base alla frequenza sono il cammino fisico (WP: Wavelength Path) ed il cammino virtuale (VWP: Virtual Wavelength Path); queste due tecniche sono illustrate schematicamente in fig. 3.

Nel caso del WP il segnale viene instradato attraverso la rete in base alla sua frequenza senza che essa sia mai cambiata durante il cammino in rete. Anche nel caso del VWP l'instradamento avviene in base alla frequenza del segnale ma il cammino percorso in rete da un singolo flusso dati ad alta velocità è caratterizzato non da una singola frequenza ma da un vettore di frequenze, una per

ogni ramo di rete che costituisce il cammino. Questo implica che, durante l'attraversamento del nodo da parte del segnale avvenga conversione di frequenza.

Sia le strategie di instradamento che quelle di ripristino della rete in seguito ad eventuali guasti sono ovviamente influenzate dall'adozione del WP o del VWP. Queste problematiche sono studiate in dettaglio in [10] relazione a diversi modelli di rete, composti da 6, 15 o 50 nodi, che si immaginano sottoposti a diverse richieste di traffico. Dai risultati di questo studio, che non si può analizzare in dettaglio in questa sede per motivi di spazio, si vede che il numero di frequenze necessarie per realizzare il numero richiesto di cammini ottici è sempre minore nel caso si adotti il VWP. Questa differenza, ovviamente trascurabile in basse condizioni di traffico, arriva fino a circa il 10% in alte condizioni di traffico. La differenza tra WP e VWP diviene poi molto più evidente se si considera che, oltre al numero di frequenze necessarie per realizzare i cammini ottici richiesti, la rete deve anche disporre di un certo numero di frequenze per effettuare il ripristino dei cammini in caso di guasto. Tenendo conto di questo fattore si vede che, se nel caso del WP si vuole effettuare il ripristino mantenendo invariata la frequenza del cammino da ripristinare, il VWP richiede circa il 30% delle frequenze in meno per assicurare le stesse funzionalità di rete. Se invece, pur adottando il WP, si accetta di cambiare la frequenza ottica del cammino durante il ripristino, il risparmio di frequenze realizzato adottando il VWP scende circa al 15%. Questi ultimi dati dipendono dal traffico nella rete molto meno di quelli ottenuti non considerando le strategie di ripristino e dimostrano come, al crescere delle dimensioni della rete, il VWP consente una gestione più semplice ed economica. Ad esempio, in caso di guasto, il reinstradamento del traffico coinvolge una zona della rete assai più limitata che nel caso di uso del WP, poiché implica la riassegnazione di un numero molto minore di lunghezze d'onda.

Questi vantaggi ovviamente si ottengono grazie ad una maggior complessità del OXC, all'interno del quale deve essere effettuata la conversione di frequenza dei flussi ottici ad alta velocità.

Un possibile percorso evolutivo che porti l'attuale rete del trasporto verso una rete in cui compaia uno strato di cammino ottico è presentato in fig. 4. Il primo schema a sinistra in fig. 4, presenta una rete in cui le tecnologie ottiche sono limitate al mezzo trasmissivo e lo strato di cammino è basato in parte sulla tecnologia plesiocrona (PDH: Plesiochronous Digital Hierarchy) e in parte dalla tecnologia sincrona (SDH: Synchronous Digital Hierarchy). In una prima fase si introduce nello strato di cammino un livello ottico basato sul WP o sul VWP che fornisce uno strato di trasporto comune su cui sono basate le diverse strategie elettriche per il trasporto a livello superiore. In una seconda fase, a partire da questo strato di trasporto comune, può essere inserita la

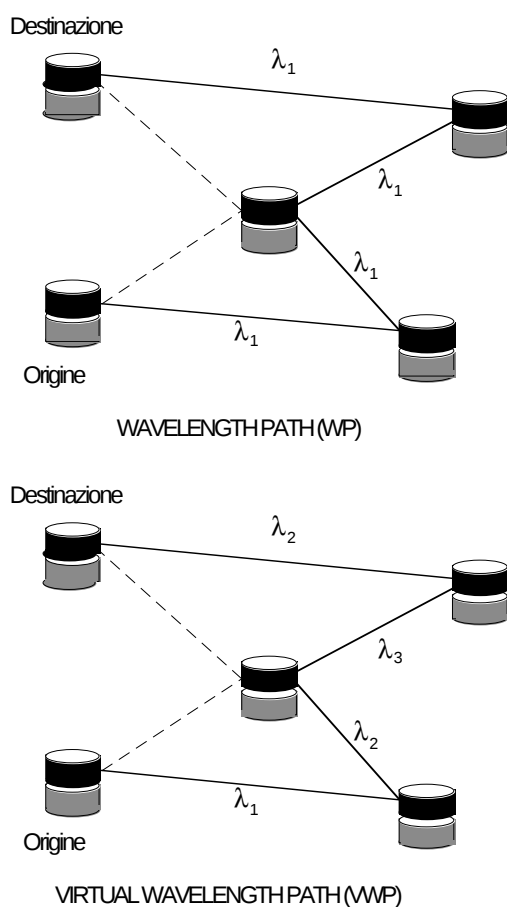


Figura 3 Diverse tecniche di instradamento in base alla frequenza

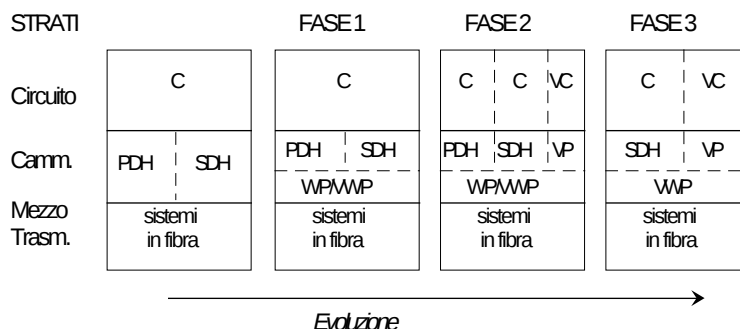


Figura 4 Evoluzione della struttura a strati della rete del trasporto; in questa ipotesi evolutiva si immagina che la rete ATM sia integrata con la rete a commutazione di circuito dopo l'introduzione dello strato di cammino ottico

tecnica ATM basata sul cammino virtuale (VP: Virtual Path) nello strato di cammino e sul circuito virtuale (VC: Virtual Circuit) nello strato di circuito. Infine, in una ulteriore fase dello sviluppo della rete, la tecnologia PDH nello strato elettrico ed il WP nello strato ottico tenderanno a scomparire dando luogo ad una rete in cui, su uno strato di cammino ottico comune basato sul VWP, convivono strategie di trasporto SDH e VP per diversi servizi. Ovviamente il cammino evolutivo mostrato in fig. 4 è da intendersi puramente indicativo. Ad esempio è possibile che le tecniche ATM siano introdotte nella rete di trasporto in modo esteso prima di uno strato ottico comune così come è possibile che i sistemi PDH spariscano di fatto dalla rete di trasporto prima di quanto illustrato in fig. 4.

Un esempio di evoluzione possibile della rete verso l'introduzione di uno strato ottico di cammino è mostrato

in fig. 5. A partire dalla rete elettrica attuale, in alcune zone geografiche in cui lo sviluppo socioeconomico lo richiede, verrà introdotto lo strato ottico della rete, così da arrivare ad una rete (fig. 5a) che presenti isole in cui si hanno cammini ottici immerse in una rete essenzialmente elettrica. All'aumentare della richiesta di servizio e quindi del traffico della rete, le isole ottiche diverranno sempre più estese e si fonderanno fino a giungere ad una rete essenzialmente ottica con delle isole elettriche (fig. 5b). Da qui con un'ulteriore crescita del traffico si potrà arrivare ad uno strato di cammino completamente ottico per i flussi informativi ad alta velocità.

Nell'ambito dello sviluppo della rete schematizzato in fig. 5 si può pensare che, nella prima fase dello sviluppo (fig. 5a) in ogni isola ottica all'interno della rete l'instradamento si effettui mediante il WP. Ovviamente nelle diverse isole sarà usato sempre lo

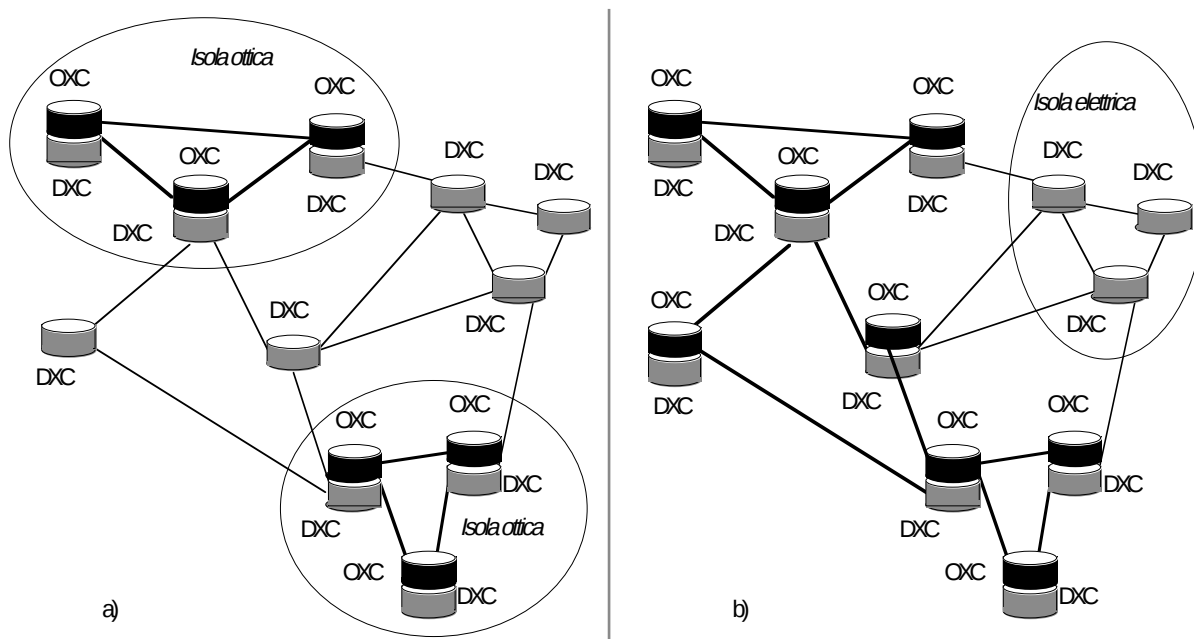


Figura 5 Evoluzione della rete del trasporto verso l'introduzione di uno strato ottico: a) isole ottiche, b) isole elettriche nella rete ottica

stesso insieme di frequenze, che dovranno essere standardizzate per ottenere la compatibilità dei diversi apparati. Quando si dovranno connettere tra loro diverse isole ottiche (fig. 5b) si passerà naturalmente al VWP. Questa operazione si può effettuare introducendo in un primo momento le funzionalità di conversione di frequenza solo in alcuni degli OXC, necessari per connettere diverse isole ottiche: a mano a mano che lo strato ottico si estende a tutta la rete si può introdurre la conversione di frequenza in tutti gli OXC per sfruttare a pieno le caratteristiche di flessibilità del VWP.

Dalla descrizione del cammino evolutivo della rete verso l'introduzione di uno strato di cammino ottico trasparente basato sul VWP emergono chiaramente quali sono i principali requisiti dello strato ottico di cammino [11,12]:

- *Trasparenza*: in senso stretto la trasparenza rappresenta la capacità di elaborare e trasmettere segnali ottici di qualsiasi velocità di cifra e formato di modulazione senza cambiare altro che trasmettitori e ricevitori. La trasparenza alla velocità di cifra è ovviamente un requisito molto importante per permettere l'aggiornamento della rete; la trasparenza al formato di modulazione non è altrettanto importante ma potrebbe risultare utile se emergessero nuove tecniche di trasmissione.
- *Modularità*: deve essere possibile realizzare la rete senza disporre fin dall'inizio di un progetto complessivo, ma partendo dalla realizzazione di alcune sue parti ed aggiungendone progressivamente altre senza dover alterare in modo radicale la struttura delle isole di rete già realizzate (in particolare non deve essere necessario rivoluzionare la struttura del nodo).
- *Espandibilità*: deve essere sempre possibile aggiungere nodi alla rete, cioè non deve mai succedere che la rete, in seguito al suo sviluppo, arrivi in uno stato in cui non sia più possibile aggiungere ulteriori nodi se non rivoluzionandone tutta la struttura.

3. Funzioni e Struttura del Cross-Connect Ottico

Nell'ambito dell'introduzione di uno strato di cammino ottico trasparente, un generico nodo della rete del trasporto può pensarsi composto da un OXC e da un DXC. L'OXC è preposto alla commutazione ed all'instradamento dei singoli canali ottici ad alta velocità ed alla loro distribuzione verso gli anelli ottici della rete di giunzione che effettuano le operazioni di add/drop direttamente al livello ottico. Quando i flussi dati ad alta velocità devono essere demultiplati in modo che i tributari che li compongono vengano separatamente elaborati, essi vengono traslati al livello elettrico dall'OXC e passati al DXC che provvede alla demultiplazione ed all'instradamento nel livello di cammino elettrico o all'invio verso la rete di giunzione elettrica. Questi meccanismi sono evidenziati in fig. 6.

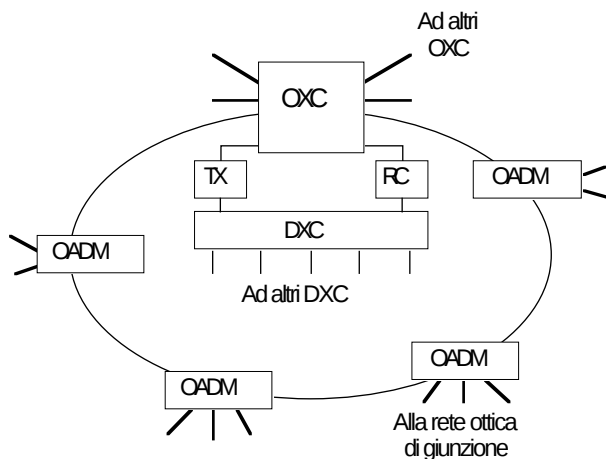


Figura 6 Accesso ad una rete di trasporto con livello di cammino ottico tramite dei ripartitori digitali (DXC: Digital Cross-Connect) e degli anelli ottici in area di giunzione con nodi ottico di Add/Drop (OADM). TX=trasmettitori, RC=Ricevitori

In questo contesto, canali ottici ad alta velocità possono essere inseriti nel livello ottico di cammino in due modi differenti: o attraverso un anello di add/drop ottico o attraverso il DXC che provvede ad affasciare diversi tributari in un flusso ad alta velocità, e lo invia ai trasmettitori dell'OXC che lo immettono nello strato ottico [13].

I requisiti funzionali dell'OXC si possono dedurre dai requisiti dello strato ottico di rete illustrati precedentemente. In particolare risulta importante la modularità dell'OXC, cioè la capacità dell'OXC di aumentare il traffico gestito senza dover cambiare completamente la struttura del commutatore, ma, idealmente, semplicemente aggiungendo alcune parti necessarie per elaborare il traffico aggiunto. In pratica, l'aggiunta di nuovi dispositivi può avvenire solo tramite adeguata predisposizione dell'OXC (es. con fibre in eccesso rispetto a quelle utilizzate nella fase iniziale, ecc.).

In particolare si possono distinguere due tipi di modularità dell'OXC [8]:

- *Modularità di canali*: si ha quando l'OXC è modulare rispetto all'aggiunta di nuove lunghezze d'onda nel pettine WDM sulle fibre di ingresso ed uscita;
- *Modularità di fibre*: si ha quando l'OXC è modulare rispetto all'aggiunta di nuove fibre in ingresso ed in uscita.

Per garantire l'espandibilità e la modularità della rete è importante che sia facile ed economico aggiungere fibre in ingresso ed in uscita al nodo, per attuare connessioni con nuovi nodi; l'OXC deve quindi essere modulare in fibre. La modularità in canali invece è utile per garantire la possibilità di adattare la rete a richieste superiori di traffico aumentando il numero di canali inviati su una determinata connessione fisica.

Un primo esempio di architettura dell'OXC è riportato in fig. 7 [6]: si tratta della struttura proposta nell'ambito del progetto RACE (2028: MWTN). In ingresso all'OXC, la demultiplazione dei canali WDM è realizzata attraverso divisori di potenza ottici passivi (splitter) e filtri accordabili. La funzione di accordabilità dei filtri permette di riconfigurare la demultiplazione, rendendo più flessibile la struttura del nodo stesso. Un commutatore ottico spaziale consente l'instradamento vero e proprio dei canali, mentre in uscita i canali sono multiplati tramite accoppiatori ottici passivi. L'inserimento di canali nel traffico avviene grazie ad un banco di trasmettitori, a lunghezze d'onda diverse tra loro, che consente di accedere al commutatore, mentre l'operazione di estrazione di canali avviene per mezzo di un banco di ricevitori. All'ingresso ed all'uscita del nodo sono posti amplificatori ottici in fibra drogata ad Erblio (EDFA) per compensare le perdite introdotte dalla fibra ottica e dagli elementi passivi del nodo.

Per quel che riguarda la tecnologia ottica impiegata per la realizzazione di questo tipo di nodo, è opportuno prendere in considerazione alcuni dispositivi chiave. Il cuore del commutatore è costituito da matrici ottiche di commutazione spaziale, realizzate utilizzando due differenti tecnologie basate su differenti materiali: LiNbO_3 [14] e InP [15]. Attualmente la tecnologia che appare più promettente è quella basata su InP : essa permette di realizzare matrici con perdite molto basse integrabili in circuiti monolitici con altri dispositivi basati sulla stessa tecnologia. Quest'ultima caratteristica è di una certa importanza poiché l'integrazione comporta generalmente una grande riduzione dei costi di realizzazione.

Allo stato dell'arte, sono disponibili due categorie di dispositivi in grado di realizzare il filtraggio ottico accordabile: i filtri Fabry-Perot [16] ed i filtri acustooptici [17]. Una terza alternativa, non ancora matura tecnologicamente, ma di sicuro interesse per le possibilità di integrazione che fornisce, è quella dei cosiddetti filtri attivi basati su reticoli di tipo DFB (Distributed FeedBack) in guide a semiconduttore [18].

L'architettura di nodo ottico presentata in fig. 7 si presenta modulare in canali, a patto di prevedere una ridondanza negli accoppiatori di ingresso ed uscita che realizzano la multiplazione e la demultiplazione, ma non è modulare in fibre. L'aggiunta di una nuova fibra infatti richiede la sostituzione di tutte le matrici di commutazione spaziale.

E' importante notare che l'architettura di fig. 7 non permette di realizzare il VWP. Inoltre, se si realizzano strategie di instradamento diverse dal WP, il commutatore ottico risulta bloccante. Infatti, quando diversi canali alla stessa lunghezza d'onda e provenienti da fibre di ingresso diverse devono essere instradati verso la stessa uscita, si verifica una situazione di blocco. Questo blocco può essere rimosso solo tramite l'indirizzamento dei canali che sono in conflitto verso i

ricevitori locali e la loro ritrasmissione a frequenza diversa.

Ambedue i problemi si possono risolvere introducendo nell'architettura di fig. 7 alcuni convertitori di frequenza, come mostrato in fig. 8 [19]. In questo modo, la conversione di frequenza necessaria per la soluzione dei conflitti si può operare direttamente al livello elettrico ed il commutatore ottico, che ora è di tipo misto spazio-frequenza, risulta non bloccante in senso stretto. Inoltre la disponibilità di commutatori di frequenza permette la realizzazione del VWP.

Un secondo esempio di struttura dell'OXC è presentato in fig. 9 [20]. All'ingresso dell'OXC sono presenti M linee di ingresso ($M-1$ fibre esterne ed una che proviene dalla batteria dei trasmettitori locali) e su ogni linea di ingresso sono multiplati n_c canali WDM.

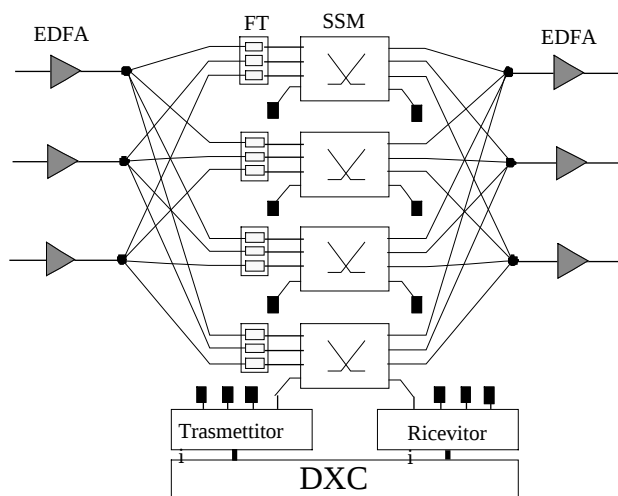


Figura 7 Ripartitore ottico (OXC: Optical Cross Connect) elaborato nell'ambito del RACE 2028 (MWTN). SSM=Matrice di commutazione ottica (Space Switch Matrix), FT=Filtri Sintonizzabili

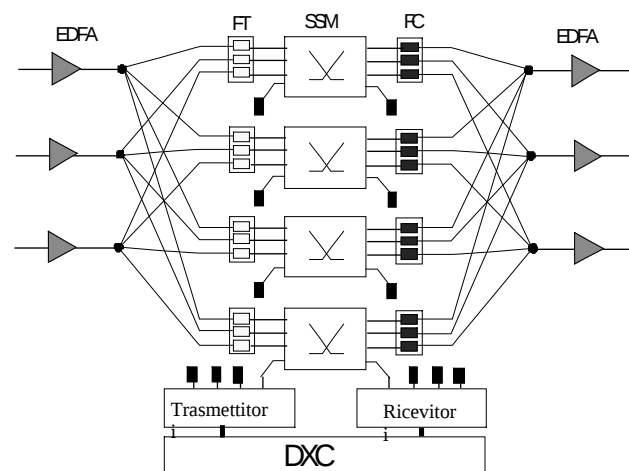


Figura 8 Ripartitore ottico trasparente a commutazione spaziale. SSM=Matrice di commutazione ottica (Space Switch Matrix),

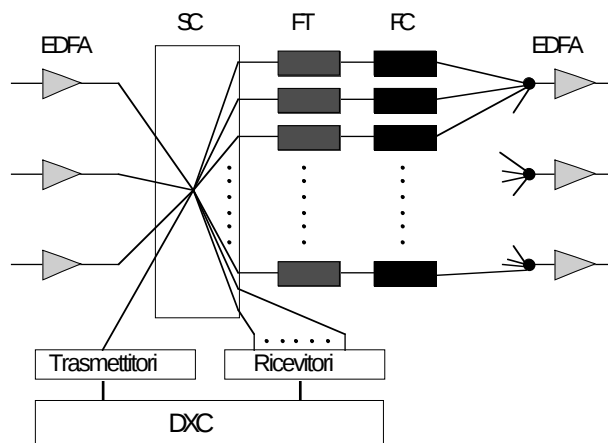


Figura 9 Ripartitore ottico trasparente a commutazione di frequenza. SC=Accoppiatore a stella, FT=Filtri Sintonizzabili, FC=Convertitori ottici di frequenza

Gli M pettini WDM in ingresso al commutatore sono disposti in regioni contigue dello spettro, come mostrato in fig. 10. All'ingresso del commutatore i pettini WDM vengono multiplati tutti insieme in una stella passiva le cui uscite sono in numero pari ai canali ottici in ingresso (M_{nc}). Su ogni uscita un filtro accordabile seleziona uno dei canali ed un traslatore di frequenza trasforma la frequenza del canale selezionato in una delle frequenze associate al pettine della corrispondente fibra in uscita. Infine i canali sono multiplati sulle fibre di uscita mediante accoppiatori passivi. Anche in questo caso le perdite dei componenti passivi del nodo ed eventualmente dei tratti trasmissivi sono compensate da una coppia di EDFA.

Lo schema di OXC proposto in fig. 9 presenta modularità di fibre se è predisposta una ridondanza nella stella che effettua la multiplazione in ingresso al commutatore ed anche modularità di canali se tra i pettini WDM delle fibre in ingresso sono previsti intervalli spettrali per l'allocazione di eventuali altri canali aggiunti. La configurazione del commutatore è

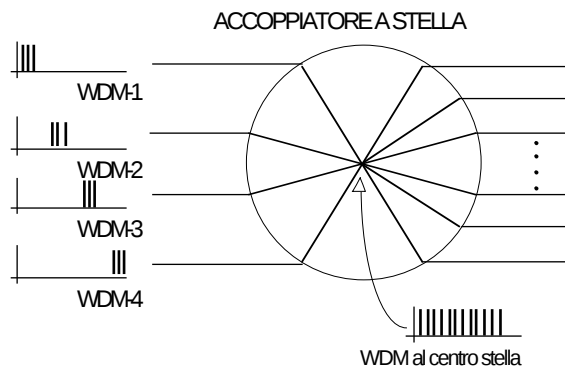


Figura 10 Ripartizione spettrale dei canali WDM sulle linee di ingresso di un ripartitore del tipo mostrato in figura 9

intrinsecamente non bloccante e permette, per il modo stesso in cui è realizzato l'instradamento, di effettuare il VWP. Bisogna però osservare che i vantaggi considerati precedentemente sono ottenuti a prezzo di una disposizione delle frequenze rigidamente fissata sulle fibre di ingresso e di uscita del commutatore. Questo fatto riduce di fatto la banda ottica disponibile in fibra di tante volte quante sono le fibre in ingresso al commutatore e soprattutto rende più rigide le strategie di instradamento dei flussi informativi nello strato ottico della rete.

La necessità di allocare in intervalli spettrali contigui i pettini WDM nelle fibre in ingresso ed in uscita all'OXC si può eliminare prevedendo un ulteriore stadio di conversione di frequenza in ingresso all'OXC, che permetta di passare da una configurazione in cui si ha lo stesso pettine WDM su tutte le fibre in ingresso, alla situazione descritta precedentemente, e pilotando lo stadio di conversione di frequenza in uscita in modo da riprodurre lo stesso pettine WDM su tutte le fibre in uscita. Questa nuova architettura dell'OXC, che presenta due stadi di conversione di frequenza e permette alta modularità e flessibilità, è ancora nella fase preliminare di studio e le sue prestazioni non verranno analizzate in questo lavoro.

Altre strutture possibili per l'OXC sono considerate in [8], in cui si propone anche uno schema di OXC basato sulla commutazione spaziale che risulta modulare in fibre ed un diverso schema di OXC basato sulla commutazione di frequenza.

4. Convertitori Ottici di Frequenza

La conversione di frequenza è una funzionalità essenziale in una struttura di rete che possa evolvere verso l'introduzione di uno strato di cammino completamente ottico: senza la conversione di frequenza non si può realizzare l'instradamento mediante VWP né, allo stato attuale della tecnologia, si possono progettare OXC completamente trasparenti a causa della necessità di risolvere situazioni di contesa all'interno del commutatore ottico mediante trasduzione ottico-elettrica.

Negli ultimi anni sono stati proposti diversi dispositivi in grado di effettuare conversione di frequenza interamente ottica. Alcuni di questi dispositivi sono basati su tecnologie consolidate (la tecnologia dei dispositivi in fibra o dei dispositivi attivi a semiconduttore) e sono già in fase di ingegnerizzazione per permetterne la produzione industriale. Altri dispositivi sono basati su tecniche più avanzate e le loro modalità di funzionamento sono ancora oggetto di ricerca.

In questo paragrafo verrà presentata una breve rassegna dei convertitori di frequenza basati su tecnologia consolidata, mentre per gli altri si rimanda alla letteratura specialistica.

Nella tab. 1 si considerano cinque diversi convertitori; uno di essi è basato sulla tecnologia delle fibre, gli altri su dispositivi a semiconduttore: laser DFB o amplificatori ad onda viaggiante (SOA: Semiconductor Optical Amplifiers). I primi tre dispositivi considerati in tabella sono basati sulla miscelazione a quattro onde (FWM: Four Wave Mixing) in diversi tipi di dispositivi, in particolare fibre ottiche [21], amplificatori a semiconduttore [22,23] e laser a semiconduttore [24]. Gli altri due tipi di dispositivi sono basati sulla saturazione di guadagno (XGM: Cross Gain Modulation) [25] e sulla saturazione di indice (XPM: Cross Phase Modulation) [26,27] in amplificatori a semiconduttore. Per una descrizione dettagliata della struttura e dei principi di funzionamento dei diversi convertitori si rimanda alla letteratura specialistica mentre, in questa sede, è interessante confrontare tra loro le caratteristiche dei diversi dispositivi, per evidenziare quali di essi sono adatti ad essere inseriti nell'OXC. Per far questo è opportuno commentare una ad una le colonne di tab. 1.

- 1) I convertitori basati sulla tecnologia dei semiconduttori si possono integrare, nell'ambito di circuiti optoelettronici più complessi, con altri dispositivi come laser, filtri e così via. Questa proprietà è importante per la realizzazione economica ed affidabile di un sistema così complesso come un OXC.
- 2) Come massimo intervallo di conversione di frequenza si è indicato l'intervallo su cui si può ottenere da un convertitore ottimizzato una efficienza di conversione non inferiore a circa -15 dB, valore che ancora permette un utilizzo pratico del dispositivo. Ovviamente i dati riportati in tabella devono essere intesi come indicativi, dipendendo i valori precisi dal dispositivo e dalle condizioni di impiego. Dalla tabella è evidente che i dispositivi basati sui SOA raggiungono ampi intervalli di conversione, confrontabili con la banda disponibile negli EDFA. Questi dispositivi quindi permettono, di fatto, di traslare la frequenza di un segnale

portandola in qualsiasi punto della banda disponibile. Gli altri convertitori invece hanno un intervallo di conversione più limitato.

- 3) La massima velocità di cifra del segnale convertibile dipende in maniera critica da quale fenomeno genera la conversione di frequenza, ma come si vede dalla tabella, tranne che nel caso dei convertitori basati sul FWM nei laser DFB, possono essere convertiti segnali fino ad almeno 20 Gbit/s.
- 4) La possibilità di mettere in cascata un numero elevato di convertitori di frequenza lungo un collegamento ottico, senza pregiudicare la qualità della trasmissione è ovviamente un requisito importante per dispositivi che devono essere applicati nelle reti ottiche. Questo è possibile per tutti i dispositivi analizzati, tranne che per quelli basati sull'XGM negli amplificatori a semiconduttore: il motivo va ricercato nella degradazione del rapporto di estinzione del segnale che si verifica durante la conversione.
- 5) L'indipendenza dalla polarizzazione del segnale in ingresso è molto utile in un convertitore di frequenza, anche se nell'ambito di un OXC probabilmente la polarizzazione dei segnali ottici dovrà essere comunque controllata rendendo questo requisito meno critico. I convertitori basati sul FWM non sono di per sé indipendenti dalla polarizzazione ma, tranne il convertitore basato sui laser DFB, si possono rendere tali con particolari accorgimenti: questi accorgimenti comunque comportano almeno l'utilizzo di due laser di pompa invece che uno [28,29]. I convertitori basati sull'XPM e sull'XGM nei SOA sono invece intrinsecamente indipendenti dalla polarizzazione se è tale il dispositivo che si utilizza per la conversione.
- 6) La capacità di convertire simultaneamente in frequenza un intero pettine WDM [30], caratteristica dei convertitori basati sul FWM, potrebbe essere utile in alcune architetture di OXC basate sulla commutazione di frequenza per diminuire le perdite

	Integrabile	Max. $\Delta\lambda$ (nm)	Max. R (Gbit/s)	Cascata	Indip. Polariz.	WDM Convers.	FSK PSK	Efficienza $f(\Delta\lambda)$	Coniug. di fase
FWM in fibra	NO	≈ 4	40	SI	NO (SI)	SI	SI	SI	SI
FWM nei SOA	SI	≈ 20	>100	SI	NO (SI)	SI	SI	SI	SI
FWM nei DFB	SI	$\approx 0.3-0.5$	2.5	SI	NO	SI	SI	SI	SI
XGM nei SOA	SI	≈ 30	20	NO	SI	NO	NO	NO	NO
XPM nei SOA	SI	≈ 30	20	SI	SI	NO	NO	NO	NO

Tabella 1 Confronto tra vari tipi di convertitori ottici di frequenza

all'interno del nodo mediante l'eliminazione di una operazione di demultiplazione. Un esempio di OXC, in cui questa proprietà è utile, è il secondo tipo di OXC basato sulla commutazione di frequenza descritto nella sezione precedente.

- 7) I convertitori basati sul FWM, poiché preservano l'andamento temporale della fase del segnale convertito possono convertire segnali basati su qualsiasi formato di modulazione mentre i convertitori basati sull'XPM o sull'XGM possono convertire solo segnali modulati in intensità o in ampiezza. La capacità di convertire segnali modulati in fase o in frequenza rende la rete trasparente al formato di modulazione.
- 8) Nei convertitori basati sul FWM l'efficienza di conversione dipende dall'intervallo di conversione, mentre ciò non è vero per i convertitori basati sull'XPM o sull'XGM. Questa dipendenza è tanto meno pronunciata all'aumentare dell'intervallo di conversione. A causa di questo fenomeno, all'aumentare della velocità di cifra R del segnale convertito, per evitare distorsioni, è necessario aumentare l'intervallo di conversione $\Delta\nu$ in modo che sia $\Delta\nu \gg R$. Inoltre canali che vengono spostati in frequenza di intervalli diversi sono attenuati in modo diverso dal convertitore ed in seguito dovranno essere equalizzati.
- 8) I convertitori basati sul FWM effettuano la coniugazione ottica di fase del segnale convertito. Questa proprietà è molto utile in reti di grande estensione geografica. Infatti un coniugatore di fase posto a metà di un collegamento permette di compensare completamente le distorsioni introdotte dalla dispersione cromatica della fibra [30] e parzialmente quelle introdotte dalla non linearità Kerr [31,32]. Nel caso che all'interno dell'OXC si utilizzino convertitori di frequenza basati sul FWM, essi non saranno ovviamente distribuiti in modo regolare lungo il cammino in rete del segnale e quindi l'effetto di compensazione non sarà completo. Comunque ci si può aspettare una compensazione parziale degli effetti della dispersione e della non linearità Kerr con un conseguente miglioramento delle prestazioni trasmissive della rete.

Da questo confronto si può concludere che alcuni dei dispositivi proposti non sono adatti all'uso in un OXC come dispositivi di commutazione. Il convertitore in fibra non è integrabile e comunque ha un intervallo massimo di conversione troppo limitato, il convertitore basato sul FWM nei laser ha un intervallo di conversione estremamente limitato ed il convertitore basato sull'XGM non permette di mettere molti dispositivi in cascata. In conclusione i dispositivi che sembrano più promettenti per l'applicazione di interesse sono quelli basati sul FWM e sull'XPM negli amplificatori a semiconduttore.

5. Aspetti Trasmissivi dello Strato Ottico di Rete

L'introduzione di uno strato di cammino ottico nella rete del trasporto introduce una serie di nuovi problemi di natura trasmissiva che devono essere considerati per giudicare l'adeguatezza delle tecnologie scelte. Infatti, poiché il nodo ottico di rete si presenta ora come un sistema trasparente, il cammino generico di un flusso dati nella rete si può schematizzare come in fig. 11: un OXC origine, dove è collocato il trasmettitore, una catena di OXC che figurano come dispositivi in linea, ed un OXC destinazione in cui è localizzato il ricevitore. Nei collegamenti in fibra tra due OXC in linea possono eventualmente essere presenti degli amplificatori EDFA per compensare le perdite della fibra.

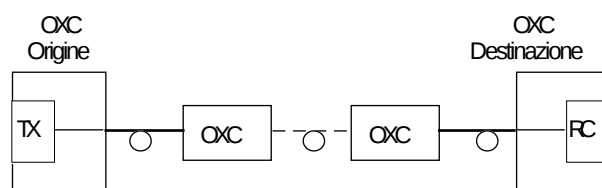


Figura 11 Schema trasmissivo di un cammino ottico attraverso la rete del trasporto

In questo lavoro, tra tutte le possibili strutture di OXC, si prenderanno in considerazione per la valutazione delle prestazioni trasmissive della rete, le strutture mostrate in fig. 8 ed in fig. 9, sia nel caso che impieghino convertitori basati sul FWM nei SOA sia nel caso che impieghino convertitori basati sull'XPM. In particolare tra i convertitori basati sull'XPM si prenderà in considerazione quello basato su una configurazione di tipo Mach-Zehnder.

Diversi fenomeni limitano la massima capacità trasmissiva di un collegamento come quello mostrato in fig. 10:

- *Rumore Ottico*: ogni elemento attivo lungo il collegamento introduce un rumore ottico (ASE: Amplified Spontaneous Emission) inevitabile perché legato alla natura quantistica del processo di amplificazione del campo ottico [33]. Questo rumore, che si può schematizzare come un processo gaussiano, si accumula lungo il collegamento e costituisce uno dei limiti fondamentali delle prestazioni trasmissive di un sistema ottico che utilizzi dispositivi attivi.
- *Dispersione*: la dispersione cromatica della fibra introduce, in collegamenti lunghi ad alta velocità, un allargamento degli impulsi trasmessi con conseguente interferenza intersimbolica al ricevitore.
- *Effetto Kerr*: quando si propagano in fibra per lunghe distanze segnali di grande velocità di cifra, la propagazione in fibra non può più essere descritta come un processo lineare [34]. Tra i fenomeni non

lineari, il più importante per i sistemi di trasmissione che si considerano in questo lavoro è l'effetto Kerr. Questo effetto consiste in una dipendenza dell'indice di rifrazione della fibra dall'intensità del campo trasmesso, provocando così un accoppiamento non lineare tra ampiezza e fase del campo che va sotto il nome di automodulazione di fase (SPM: Self Phase Modulation). Poiché la banda ottica deve essere ristretta per mezzo di filtri per limitare la potenza di ASE che arriva al ricevitore, si genera un effetto di conversione della modulazione spuria di fase generata dall'SPM in una sorta di rumore di ampiezza che degrada le prestazioni del sistema. Inoltre l'interazione dell'effetto Kerr con la dispersione genera anche altre fonti di distorsione del segnale come la instabilità di modulazione (MI: Modulation Instability).

- *Rumore di fase*: i laser a semiconduttore presentano rumore di fase che determina essenzialmente la loro larghezza di riga. Nel caso che usino convertitori di frequenza basati sul FWM, il rumore di fase del laser di pompa ottica del convertitore si riflette amplificato sul segnale convertito [35]. Questo genera accumulo di rumore di fase lungo il collegamento. Poiché a causa dei filtri ottici si ha conversione di rumore di fase in rumore di intensità questo effetto può limitare le prestazioni del sistema.
- *Interferenza lineare*: in ogni sistema WDM si ha interferenza lineare. Questa interferenza è dovuta al fatto che lo spettro di tutti i canali trasmessi si estende su tutto l'asse delle frequenze e quindi nella banda del canale desiderato entrano componenti frequenziali dei canali adiacenti [36,37]. Questo effetto di interferenza, detto eterointerferenza (HEC: Heterowavelength Crosstalk), è tanto più evidente quanto minore è la selettività del filtro ottico che seleziona il canale al ricevitore. Inoltre, in una rete WDM commutata, poiché il pettine WDM dei canali contigui al canale desiderato cambia in ogni transito attraverso un OXC, l'HEC si accumula in modo incoerente lungo la tratta. Oltre all'HEC, la presenza degli OXC introduce un altro tipo di interferenza lineare: l'omointerferenza (HOC: Homowavelength Crosstalk) [38]. Questo fenomeno si verifica quando una certa parte della potenza relativa ad un canale attraversa l'OXC lungo un cammino ottico diverso da quello previsto ed, all'uscita dell'OXC, interferisce in modo coerente con il segnale relativo a quello stesso canale. Questo tipo di interferenza, che pure si accumula lungo tutto il sistema trasmissivo, è particolarmente deleteria per il suo effetto coerente. Riguardo all'HOC si può dimostrare che, se si adotta la conversione di frequenza ed opportune strategie di instradamento delle frequenze, l'effetto dell'HOC può essere grandemente ridotto nel caso che si adottino i VWP [20,39]. Questo aspetto non può essere approfondito in questa sede per motivi di spazio ma,

nell'analizzare le prestazioni trasmissive della rete, supporremo sempre che queste ipotesi siano soddisfatte.

- *Interferenza non lineare*: la presenza dell'effetto Kerr in fibra non genera solo distorsione del singolo canale trasmesso ma anche un battimento non lineare tra diversi canali a causa del FWM in fibra [40,41]. Questo battimento crea delle componenti spettrali che in parte si sovrappongono ai canali trasmessi generando interferenza non lineare. Poiché su ogni tratto del cammino di un flusso dati, all'interno di una rete commutata, gli altri canali del pettine WDM cambiano, le componenti di interferenza non lineare che si generano nei diversi tratti di fibra si cumulano in modo incoerente nell'intero collegamento.

Se si vogliono valutare in modo teorico le prestazioni trasmissive di una rete di trasporto con uno strato ottico trasparente, bisogna tenere in conto contemporaneamente tutti questi fenomeni. Per fare questo è stato messo a punto un modello semi-analitico di valutazione delle prestazioni in cui la propagazione del segnale in fibra è simulata accuratamente, e la probabilità di errore è valutata a partire dall'involuppo analitico simulato del segnale e dalla potenza del rumore ASE al ricevitore, che viene calcolata a partire dalle caratteristiche dei dispositivi in linea. Nel calcolo della probabilità di errore sono tenuti in conto la caratteristica non gaussiana della variabile di decisione e l'influenza del rumore di fase [19,42,43].

Per tutti i dispositivi ottici presenti nel sistema sono adottati, sia nella simulazione della propagazione del segnale sia nel calcolo della potenza di rumore ottico al ricevitore, dei modelli fisici accurati mutuati dalla letteratura tecnica. Nelle simulazioni riportate in questo lavoro si sono utilizzati per i principali dispositivi ottici i modelli ed i parametri riportati nei riferimenti riportati in tab. 2.

Oltre ai parametri dei dispositivi fisici impiegati nella rete, sono importanti i seguenti parametri di sistema

- Numero di tratti di fibra del collegamento (N);
- Numero di porte fisiche dell'OXC (M), $M-1$ sono fibre di ingresso ed una porta è occupata dai trasmettitori locali;
- Numero di canali WDM per porta fisica (n_c);
- Velocità di trasmissione (R);
- Spaziatura tra portanti nei pettini WDM (Δf);
- Vettore delle lunghezze dei collegamenti in fibra (L_k), o la lunghezza del generico collegamento (L) se tutti i collegamenti in fibra sono della stessa lunghezza;
- Rapporto tra banda ottica del filtro di ricezione (B) e larghezza di banda del canale trasmesso ($\rho=B/2R$).

Infine, le prestazioni trasmissive su un cammino all'interno della rete, dipendono da quali traslazioni di frequenza sono subite in ogni nodo dal canale considerato: questo è vero soprattutto quando si utilizza il convertitore basato sul FWM nei SOA, la cui efficienza dipende dall'intervallo di conversione. Per effettuare

Dispositivo	Modello fisico	Parametri
Single mode, dispersion shifted fibre	[44]	[20]
EDFA nell'OXC	[45]	[20]
EDFA in linea	[45]	[20]
Filtri Fabry-Perot	[46]	[46]
Filtri acustooptici	[16]	[16]
Matrici di commutazione 4x4 (tecnologia LiNbO ₃)	[14]	[6]
Matrici di commutazione 4x4 (tecnologia InP)	[15]	[6]
Convertitori basati sul FWM nei SOA	[46,47]	[20]
Convertitori basati sull'XPM nei SOA di tipo Mach-Zehnder	[49]	[20]
Laser sintonizzabili	[50]	[20]
Modulatore ad elettroassorbimento al trasmettitore	[51]	[51]
Ricevitori IM-DD a 2.5 e 10 Gbit/s	[52]	[20]

Tabella 2 Riferimenti bibliografici per i modelli ed i parametri dei dispositivi utilizzati nelle simulazioni

un calcolo di caso peggiore si è supposto che, in ogni nodo, il canale sia traslato del massimo intervallo di frequenza possibile sulla banda più sfavorevole del convertitore, cioè verso le frequenze più alte.

Un primo confronto, significativo sotto il punto di vista delle proprietà trasmissive, è quello tra strutture di OXC che utilizzino convertitori di frequenza basati sul FWM o sull'XPM nei SOA. In fig. 12 si mostra la probabilità di errore al ricevitore in funzione della potenza di picco per canale (P_{max}) nel collegamento, cioè la potenza all'uscita dell'OXC quando viene

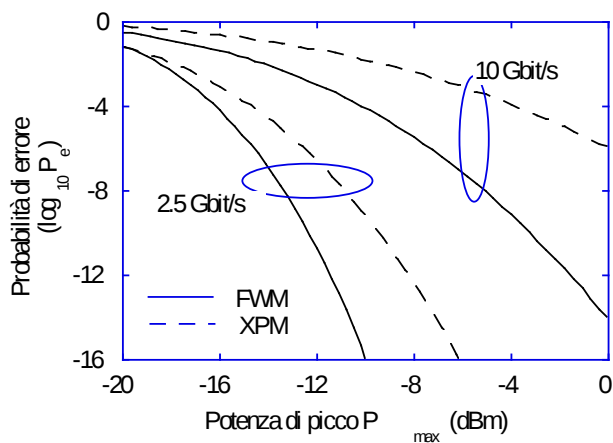


Figura 12 Probabilità di errore in funzione della potenza ottica di picco all'uscita di un OXC. I dati della rete esaminata sono riportati nel testo

trasmissione un "uno". Si considera una rete con OXC equispaziati del tipo di quelli considerati in fig. 7, $N=10$, $M=4$, $L=100$ km, $n_c=4$, $\rho=2$. La spaziatura tra canali è di 2 nm (244 GHz intorno ad una lunghezza d'onda di $1.55 \mu\text{m}$) cosicché gli effetti di interferenza sono trascurabili; la velocità di cifra è 2.5 o 10 Gbit/s. Dalla figura si può notare come, per una velocità di cifra di 2.5 Gbit/s, ambedue i convertitori considerati forniscono buone prestazioni, mentre a 10 Gbit/s la penalità dovuta alla dispersione cromatica nel caso dei convertitori basati sull'XPM impedisce un funzionamento corretto del sistema. Nel caso invece dei convertitori basati sul FWM, poiché gli OXC sono uniformemente spazati lungo il collegamento, questa penalità è perfettamente compensata dalla coniugazione di fase che avviene durante la conversione di frequenza. Poiché la dispersione cromatica rappresenta una limitazione molto severa delle prestazioni trasmissive, nel caso si usino convertitori basati sull'XPM o i nodi siano spazati in modo molto irregolare lungo il collegamento, si deve compensare questo effetto in qualche modo. Da ora in poi si assume che la dispersione cromatica di ogni collegamento in fibra sia compensata tramite una fibra passiva di compensazione posta all'ingresso di ogni OXC che assicuri un coefficiente di dispersione residuo del collegamento inferiore a 0.4 ps/nm/km.

L'effetto dell'interferenza nel caso di una rete che adotti la multiplazione WDM a piccola spaziatura è mostrato in fig. 13 mediante le curve che danno la potenza P_{max} necessaria per raggiungere una probabilità

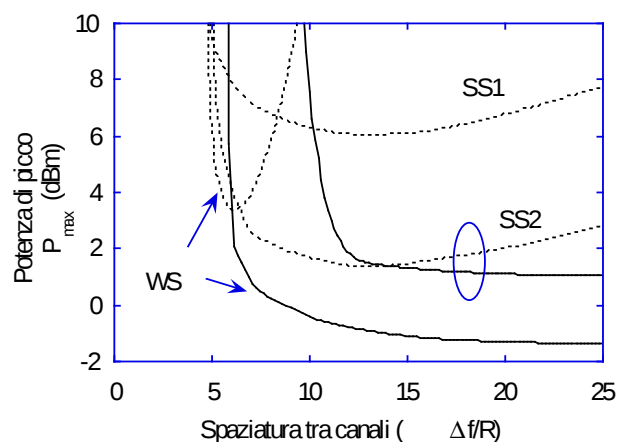


Figura 13 Potenza di picco necessaria per assicurare una probabilità di errore di 10^{-10} in funzione della spaziatura in frequenza dei canali. WS=OXC a commutazione di frequenza, SS1=OXC a commutazione spaziale con matrici in LiNbO₃, SS2=OXC a commutazione spaziale con matrici in InP. Le linee tratteggiate si riferiscono ad OXC che usano convertitori basati sul FWM nei SOA, le linee continue ad OXC che usano convertitori basati sull'XPM. In quest'ultimo caso nella rete è adottata la compensazione passiva della dispersione delle fibre

di errore di 10^{-10} in funzione della spaziatura tra canali. Si considera una rete con quattro OXC equispaziati ad una distanza $L=80$ km, $R=10$ Gbit/s, $M=4$, $p=2$ ed $n_c=4$; dalla figura è evidente come, nel caso in cui si adottino convertitori basati sul FWM nei SOA esiste una spaziatura ottima dei canali. Infatti, all'aumentare della spaziatura tra canali da una parte diminuisce l'effetto dell'interferenza, dall'altra diminuisce l'efficienza di conversione, che in questi convertitori dipende dall'intervallo di conversione: l'equilibrio tra questi due fenomeni determina la presenza di un valore ottimo di Δf . Questo valore ottimo è molto più critico nel caso dei convertitori basati sulla commutazione di frequenza dove i canali sperimentano delle traslazioni di frequenza molto più ampie. Dalla figura si vede pure come, tra le reti che usano la commutazione spaziale, quelle i cui OXC sono basati su matrici ad InP (SS2) hanno prestazioni migliori poiché le perdite della matrice di commutazione spaziale sono inferiori.

Nel caso invece di OXC che usino convertitori basati sull'XPM, la minima distanza cui si possono mettere i canali è superiore per la maggior sensibilità di questi convertitori alle fluttuazioni di ampiezza del segnale, ma le prestazioni migliorano monotonamente al crescere della spaziatura tra canali, almeno fino al punto in cui qualcuno dei canali esce dalla banda a 3 dB del SOA che effettua la conversione. Questo è dovuto al fatto che, in queste condizioni, l'efficienza del convertitore in pratica

è indipendente dall'intervallo di conversione.

Per evidenziare le potenzialità di una rete ad alta capacità con uno strato di cammino ottico, in questo lavoro si considera un caso concreto, quello di una possibile topologia della rete di trasporto italiana come appare in fig. 14 in cui è anche indicata la lunghezza di ogni tratta tra due OXC. Per valutare le prestazioni trasmissive della rete si è considerato il cammino di segnale evidenziato in figura, che è il più lungo cammino possibile all'interno della rete (1673 km). Si suppone che nella rete siano utilizzati OXC basati su matrici a commutazione spaziale in InP con una spaziatura tra canali di 2 nm ($\Delta f=244$ GHz), $M=4$, $p=2$ ed $n_c=4$. Lungo i collegamenti, le perdite della fibra sono compensate da amplificatori EDFA in linea disposti ogni 50 km.

In fig. 15 è riportata la probabilità di errore in funzione della potenza ottica di picco (P_{max}) nel caso che si utilizzino convertitori basati sull'XPM e si adotti compensazione passiva della dispersione della fibra, mentre in fig. 16 si riporta un grafico simile nel caso in cui si utilizzino convertitori basati sul FWM nei SOA. E' importante notare che, in quest'ultimo caso, poiché i nodi della rete non sono uniformemente spazati, la coniugazione ottica di fase che avviene al momento della conversione di frequenza non permette una completa compensazione degli effetti dispersivi e del SPM. Per questo in figura sono considerati anche casi in cui si adotta, come nel caso dell'XPM, una compensazione passiva della dispersione dei tratti di fibra.

Dalle figg. 15 e 16 si può vedere come la rete presenti buone prestazioni trasmissive con ambedue i convertitori sia ad una velocità di cifra di 2.5 Gbit/s che ad una velocità di cifra di 10 Gbit/s. Mentre nel caso dell'XPM il collegamento non può funzionare correttamente in assenza

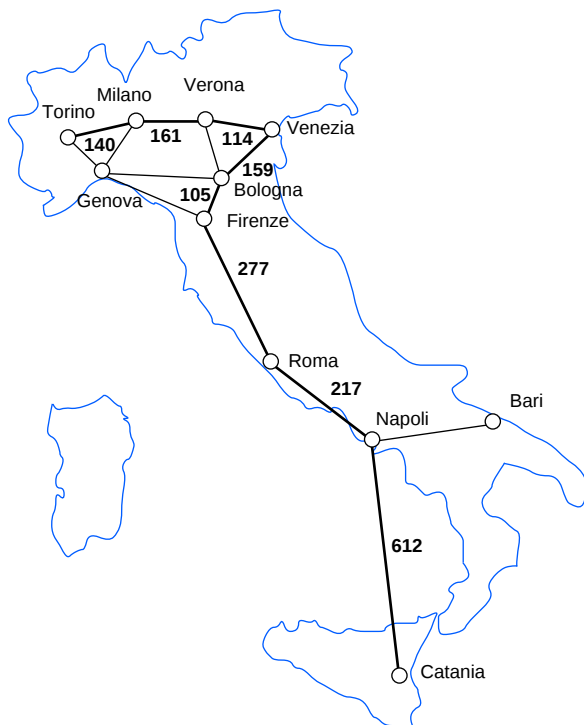


Figura 14 Possibile topologia dello strato di cammino ottico della rete italiana del trasporto. In figura sono indicate le lunghezze in km dei vari rami della rete

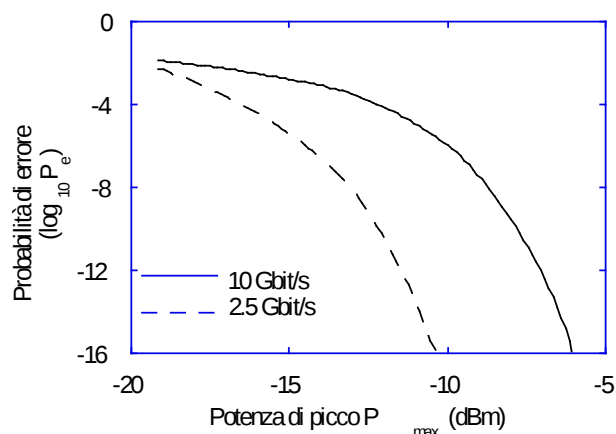


Figura 15 Probabilità di errore in funzione della potenza ottica di picco all'uscita di un OXC nel caso del cammino evidenziato in figura 15 attraverso una possibile rete italiana del trasporto. Si considerano OXC che usano su convertitori di frequenza basati sull'XPM. Nella rete è adottata la compensazione passiva della dispersione delle fibre

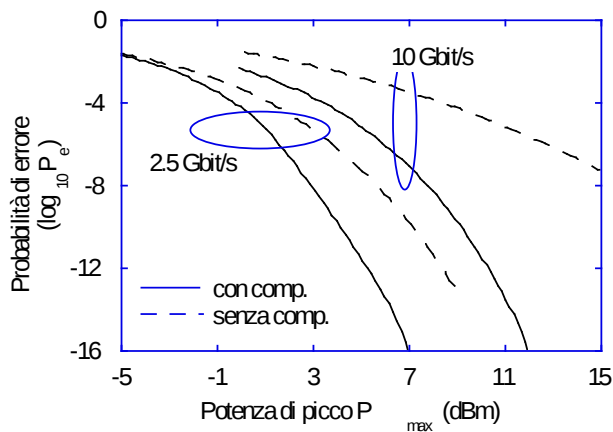


Figura 16 Probabilità di errore in funzione della potenza ottica di picco all'uscita di un OXC nel caso del cammino evidenziato in figura 15 attraverso una possibile rete italiana del trasporto. Si considerano OXC che usano su convertitori di frequenza basati sul FWM nei SOA. Nella rete è adottata la compensazione passiva della dispersione delle fibre nel caso delle linee a tratto pieno

di compensazione della dispersione, nel caso del FWM una parziale compensazione è assicurata dai convertitori anche con una spaziatura tra i nodi non uniforme.

6. Conclusioni e Prospettive Future

L'introduzione di uno strato di cammino ottico trasparente nello strato di trasporto della rete di telecomunicazione permette, in prospettiva, di integrare diverse tecniche di trasmissione e diversi servizi sulla stessa piattaforma di rete. Poiché si possono progettare i ripartitori ottici (OXC) in modo che lo strato ottico di rete risulti flessibile e modulare, è possibile un cammino evolutivo che porti gradualmente, sotto la spinta delle esigenze di servizio, all'introduzione di tale strato ottico di rete.

I dispositivi ottici fondamentali per realizzare l'OXC (multiplex e demultiplex WDM, filtri accordabili, matrici di commutazione spaziale e, soprattutto, convertitori di frequenza) sono ad uno stadio di maturazione tecnologica sufficiente da poter prevedere, in tempi brevi, la loro applicazione nella rete di telecomunicazione.

Uno strato ottico di rete così concepito, allo stato attuale della tecnologia, può essere basato su OXC capaci di elaborare decine di canali a 2.5 o 10 Gbit/s per una capacità totale nell'ordine dei 100-200 Gbit/s e può coprire distanze nell'ordine dei 1000-2000 km.

L'evoluzione della tecnologia in questo settore è comunque molto rapida, soprattutto nel campo dei convertitori di frequenza. Recenti risultati hanno segnalato la possibilità di ottenere convertitori di frequenza basati sul FWM negli amplificatori a semiconduttore con efficienza fino a 10-15 dB superiori a quelle dei dispositivi correnti

attuando una accurata ottimizzazione dell'amplificatore che viene progettato specificatamente per la conversione di frequenza [53]. Inoltre, mediante l'uso del FWM negli amplificatori a semiconduttore all'interno di anelli non lineari di fibra, si è dimostrata sperimentalmente la possibilità sia di ridurre drasticamente il rumore introdotto dai convertitori di frequenza basati sul FWM [54], sia di realizzare veri e propri rigeneratori ottici che effettuino conversione di frequenza [55].

L'applicazione di queste nuove tecniche permetterà, in prospettiva, sia di realizzare OXC più complessi, in grado di elaborare un numero di canali più alto, sia di incrementare la lunghezza del massimo cammino di rete percorribile. Inoltre la capacità trasmissiva della rete potrà essere anche aumentata utilizzando le tecniche messe a punto per le trasmissioni a lunghissima distanza [56,57,58].

In conclusione, nel medio periodo, si può pensare che la capacità dello strato ottico di rete si possa ulteriormente incrementare fino a prevedere tratte nell'ordine dei 4000 km con capacità totali elaborate dall'OXC nell'ordine del 300-400 Gbit/s.

Bibliografia

- [1] Miki, T.: *The potential of photonic networks*. «IEEE-Communications Magazine», Vol. 32, n. 12, 1994, pp. 23-27.
- [2] Jopson, R.M.: *Dispersion control in long high-dispersion fibre link*. ECOC'95, 1995, pp. 675-680.
- [3] Gnauck, A.H. et al.: *160 Gbit/s (8x20 Gbit/s WDM) 300 km transmission with 50 km amplifier spacing and span by span dispersion reversal*. «Electronics Letters», Vol. 30, n. 15, 1994, pp. 1241-1243.
- [4] Iannone, E.; Amoroso, N.: *Very high capacity FDM optical transmission system based on direct and coherent detection*. «Journal of Optical Communications», Vol. 16, 1995.
- [5] Kaiser, P.: *Optical communication network trends*. «Proceedings of the IEEE», Vol. 81, n. 11, 1993, pp. 1543-1547.
- [6] Hill, G.R. et al.: *A transport network layer based on optical network elements*. «IEEE-Journal of Lightwave Technol.», Vol. 11, n. 5/6, 1993, pp. 667-679.
- [7] Alexander, S.B. et al.: *A precompetitive consortium on wide-band all optical networks*. «IEEE-Journal of Lightwave Technol.», Vol. 11, n. 5/6, 1993, pp. 714-735.
- [8] Watanabe, A. et al.: *Optical path cross-connect node architecture with high modularity for photonic transport networks*. «IEICE Transaction on Communications», Vol. E77-B, n. 10, 1994, pp. 1220-1229.
- [9] *Architectures of transport networks based on the synchronous digital hierarchy (SDH)*. Raccomandazione ITU-T G 803, 03/93, 1993.
- [10] Sato, K.: *Optical Path and Realisation Technologies*. Workshop of COST 239, University of Essex (UK), June 1995, pp. 95-139.
- [11] Brackett, C.A.: *The principles of scalability and*

- modularity in multiwavelength optical networks*. OFC/IOOC '93, lavoro TuJ, 1993.
- [12] Brackett, C.A.: *Dense wavelength division multiplexing networks: principles and applications*. «IEEE Journal of Selected Areas in Communications», Vol. 8, n. 6, 1990, pp. 948-964.
- [13] Fioretti, A. et al.: *Application of optical transparency to the telecommunications core network*. ISSS'95, lavoro B1.4, 1995, pp. 67-71.
- [14] Frangen, J. et al.: *Integrated optical acoustically tunable wavelength filter*. «Electronics letters», Vol. 25, n. 23, 1989, pp.1583-1584.
- [15] Gustavsson, M. et al.: *Monolithically integrated 4x4 InGaAsP/InP laser amplifier gate switch arrays*. Topical Meeting on Optical Amplifiers and their Applications, post-deadline paper, 1992, pp. 38-42.
- [16] Humblet, P.A.; Hamdy, W.M.: *Crosstalk analysis and filter optimisation of single-and-double cavity Fabry-Perot filters*. «IEEE-Journal of Selected Areas in Communications», Vol. 8, n. 6, 1990, pp. 1095-1107.
- [17] Boyd, G.D.; Heismann, F.: *Tunable Acoustooptic reflection filters in LiNbO₃ without a doppler shift*. «IEEE-Journal of Lightwave Technology», Vol. 7, n. 4, 1989.
- [18] Sahlén, O.: *Active DBR filters for 2.5 Gbit/s operation: linewidth, crosstalk, noise, and saturation properties*. «IEEE-Journal of Lightwave Technology», Vol. 10, n. 11, 1992, pp. 1631-1643.
- [19] Iannone, E.; Sabella, R.: *Performance Evaluation of an Optical Multi-Carrier Network using wavelength Converters based on FWM in Semiconductor Optical Amplifiers*. «IEEE-Journal of Lightwave Technology», Vol. 13, n. 2, 1995.
- [20] Iannone, E.; Sabella, R.: *Analysis of wavelength-switched high-density WDM networks employing wavelength conversion by four-wave-mixing in semiconductor amplifiers*. «IEEE-Journal of Lightwave Technology», Vol. 13, n. 7, 1995.
- [21] Inoue, K. et al.: *Multichannel frequency conversion experiment using fibre four-wave mixing*. «Electronics Letters», Vol. 29, 1993, pp. 1708-1710.
- [22] Tatham, M.C.; Sherlock, G.: *20 nm wavelength conversion using ultrafast highly nondegenerate four-wave mixing*. Integrated Photonics Research, Topical Meeting, Palm Springs, CA, Paper PD1, March 1993.
- [23] D'Ottavi, A. et al.: *Frequency conversion by four-wave mixing on a frequency range of 8.6 THz*. ECOC'95, 1995, pp. 737-740.
- [24] Mecozzi, A. et al.: *Nearly degenerate four-wave mixing in distributed feedback semiconductor lasers operating above threshold*. «IEEE-Journal of quantum electronics», Vol. 29, n. 6, 1993, pp. 1477-1487.
- [25] Glance, B. et al.: *High performance optical wavelength shifters*. «Electronics Letters», Vol. 28, 1992, pp. 1714-1715.
- [26] Durhuus, T. et al.: *Penalty free all-optical wavelength conversion by SOA's in Mach-Zehnder configuration*. ECOC '93, TuC5.2, 1993.
- [27] Schilling, M. et al.: *Monolithic Mach-Zehnder interferometer based optical wavelength converter operated at 2.5 Gb/s with extinction ratio improvement and low penalty*. ECOC '95, 1995, pp. 647-650.
- [28] Jopson, R.M.; Tench, R.E.: *Polarisation-independent phase conjugation of lightwave signals*. «Electronics Letters», Vol. 29, n. 25, 1993, pp. 2216-2217.
- [29] Schnabel, R. et al.: *Polarisation insensitive frequency conversion of a 10 channels OFDM signal using four-wave mixing in a semiconductor laser amplifier*. «IEEE-Photonics Technology Letters», Vol. 6, n. 1, 1994, pp. 56-58.
- [30] Tatham, M.C.; Sherlock, G.; Westbrook, L.D.: *Compensation of fiber chromatic dispersion by mid-way spectral inversion in a semiconductor laser amplifier*. ECOC '93, lavoro ThP12.3, 1993.
- [31] Kikuchi, K.: *Compensation for pulse waveform distortion in ultra-long distance optical communication systems by using midway optical phase conjugator*. «IEEE Photonics Technology Letters», Vol. 6, n. 1, 1994, pp. 104-105.
- [32] Sabella, R. et al.: *Impact of wavelength conversion by FWM in semiconductor amplifiers in long distance transmissions*. CLEO/Pacific Rim'95, lavoro P102, 1995.
- [33] Yamamoto, Y.; Haus, H.A.: *Preparation, measurement and information capacity of optical quantum states*. «Review of Modern Physics», Vol. 58, n. 4, 1986, pp. 1001-1020.
- [34] Chraplyvy, A.R.: *Limitations on lightwave Communications imposed by Optical Fiber Nonlinearities*. «IEEE-Journal of Lightwave Technology», Vol. 8, n. 10, 1990, pp. 1548-1557.
- [35] Hui, R.; Mecozzi, A.: *Phase Noise of Four Wave Mixing of Semiconductor Lasers*. «Appl. Phys. Lett.», Vol. 60, 1992, pp. 2454-2456.
- [36] Jacobsen, G.: *Multichannel system design using optical preamplifiers and accounting for the effects of phase noise, amplifier noise and receiver noise*. «IEEE-Journal of Lightwave Technol.», Vol. LT-10, n. 3, 1992, pp.367-376.
- [37] Kazovsky, L.G.; Gimlett, J.L.: *Sensitivity Penalty in Multichannel Coherent Optical Communications*. «IEEE-Journal of Lightwave Technology», Vol. 6, n.9, 1988, pp. 1353-1365.
- [38] Zhou, J. et al.: *Analysis of optical crosstalk effects in multi-wavelength switched networks*. «IEEE Photonics Technology Letters», Vol. 6, n. 2, 1994, pp.302-305.
- [39] Iannone, E.; Sabella, R.: *Crosstalk in WDM optical networks adopting wavelength conversion*. Da presentare ad ECOC'95, Bruxelles, Belgio, 17-21 Settembre 1996.
- [40] Lichtman, E.: *Performance Degradation due to Four-Wave Mixing in Multichannel Coherent Optical Communications Systems*. «Journal of Optical Communications», Vol. 12, n. 2, 1991, pp. 53-58.
- [41] Inoue, K.; Toba, H.; Oda, K.: *Influence of Fiber Four-Wave Mixing on Multichannel FSK Direct Detection Transmission Systems*. «IEEE-Journal of Lightwave Technology», Vol. 10, n. 3, 1992, pp. 350-360.
- [42] Mazo, J.E.; Salz, J.: *Probability of Error for Quadratic Detectors*. «The Bell System Technical Journal», Vol. 44, n. 11, 1965, pp. 2165-2187.

- [43] Foschini, G.J. et al.: *Envelope Statistics for Filtered Optical Signals Corrupted by Phase Noise*. «IEEE-Trans. on Communications», Vol. 37, n. 12, 1989, pp. 1293-1302.
- [44] Agrawal, P.: *Nonlinear Fibre Optics*. Academic press, New York, 1989, capitolo 2.
- [45] Desurvire, E.; Simpson, R.J.: *Amplification of spontaneous emission in erbium-doped single-mode fiber amplifiers*. «IEEE-Journal of Lightwave Technology», Vol.7, n. 5, 1989, pp. 835-845.
- [46] Mecozzi, A.: *Analytical theory of four-wave mixing in semiconductor amplifiers*. «Optics Letters», Vol. 19, 1994, pp. 892-894.
- [47] D'Ottavi, A. et al.: *Efficiency and Noise Performances of Wavelength Converters Based on FWM in Semiconductor Optical Amplifiers*. «IEEE-Photonics Technology Letters», Vol. 7, n. 4, 1995, pp. 357-359.
- [48] Stubkjaer, K.E. et al.: *Optical wavelength converters*. ECOC'95, 1995, pp. 635-642.
- [50] Kobayashi, K.; Mito, I.: *Single frequency and tunable laser diodes*. «IEEE-Journal of Lightwave Technology», Vol. 6, 1988, pp. 1621-1633.
- [51] Sabella, R. et al.: *Numerical evaluation of bulk electroabsorption modulators transmission performances*. In Proceedings of EFOC & Networks '94, 1994, pp. 179-183.
- [52] Betti, S.; De Marchis, G.; Iannone, E.: *Coherent Optical Communication Systems*. J. Wiley&sons, New York, 1995, sezione 5.2.
- [53] D'Ottavi, A. et al.: *Note on frequency translators based upon four wave mixing in semiconductor amplifiers*. Rapporto interno FUB, da pubblicare.
- [54] Zhou, J.; Vahala, K.J.: *Noise reduction in four-wave-mixing wavelength converters*. CLEO'95, lavoro CThT1, 1995, p. 367.
- [55] Mori, K. et al.: *Novel of four-wave-mixing configuration: the optical parametric loop mirror*. CLEO'95, lavoro CThT2, 1995, pp. 367-368.
- [56] Taga, H. et al.: *10 Gbit/s, 9000 km IM-DD transmission experiment using 274 Er-doped fiber amplifiers repeaters*. OFC/IOOC'93, lavoro postdeadline PD1-1, 1993, pp. 9-12.
- [57] Mollenauer, L.F. et al.: *Demonstration, using sliding frequency guiding filters, of error free soliton transmission over more than 20000 km at 10 Gbit/s single channel and over more than 13000 km at 20 Gbit/s in two channel WDM*. OFC/IOOC'93, lavoro postdeadline PD8-1, 1993, pp. 37-40.
- [58] Mecozzi, A.: *Long distance transmission at zero dispersion: the combined effect of Kerr nonlinearity and noise of the in-line amplifiers*. «Journal of Optical Society of America B», Vol. 11, n. 3, 1994, pp. 462-469.
- DFB Laser a Semiconduttore a Controreazione Distribuita (Distributed Feedback Laser)
- DXC Ripartitore Elettronico (Digital Cross-Connect)
- EDFA Amplificatore Ottico in Fibra Drogata ad Erblio (Erbium Doped Fiber Amplifier)
- FWM Miscelazione a Quattro Onde (Four-Wave Mixing)
- HEC Interferenza Lineare Fuori Banda (Heterowavelength Crosstalk)
- HOC Interferenza Lineare in Banda (Homowavelength Crosstalk)
- OXC Ripartitore Ottico (Optical Cross-Connect)
- PDH Gerarchia per la Trasmissione Plesiocrona (Plesiochronous Digital Hierarchy)
- SDH Gerarchia per la Trasmissione Sincrona (Synchronous Digital Hierarchy)
- SOA Amplificatore Ottico a Semiconduttore (Semiconductor Optical Amplifier)
- TDM Multiplazione a Divisione di Tempo (Time Division Multiplexing)
- VC Circuito Virtuale (Virtual Circuit)
- VP Cammino Virtuale (Virtual Path)
- VWP Cammino Ottico Virtuale in Frequenza (Virtual Wavelength Path)
- WDM Multiplazione a Divisione di Frequenza (Wavelength Division Multiplexing)
- WP Cammino Ottico in Frequenza (Wavelength Path)
- XGM Modulazione Incrociata di Guadagno (Cross Gain Modulation)
- XPM Modulazione Incrociata di Fase (Cross Phase Modulation)

Acronimi

- ASE Rumore di Emissione Spontanea Amplificata (Amplified Spontaneous Emission Noise)
- ATM Modo di Trasferimento Asincrono (Asynchronous Transfer Mode)

"Spazi mediali": stili di interazione, opportunità tecnologiche

F. Tisato (*)

Questo lavoro si propone di fornire una chiave di lettura del fenomeno "multimedia", assumendo come obiettivo principale quello di aumentare l'efficacia della comunicazione fra utenti. Vengono introdotti quattro livelli logici corrispondenti a diversi aspetti del problema: quello degli "stili di interazione, quello dei "modelli di comunicazione", quello dei "media" e, infine, quello delle "piattaforme tecnologiche". Vengono discusse in particolare le correlazioni fra i diversi livelli, e vengono formulate alcune ipotesi sullo scenario di evoluzione dell'area.

1. Introduzione

Si assiste periodicamente alla comparsa di "parole chiave" che acquisiscono rapidamente una larghissima diffusione e vengono usate ed abusate, nei contesti più diversi e -spesso- a sproposito. Le parole "media" e "multimedia" appartengono indubbiamente a questa categoria, e sono state recentemente affiancate dal termine "spazio mediale" (traduzione forse discutibile di "media space").

Quasi tutto è "multimediale": editoria, comunicazioni interpersonali, computer, televisione, periferiche, sistemi trasmissivi, didattica, e via scorrendo. Il risultato è che "media", "multimedia", "spazio mediale", come è destino delle parole chiave onnicomprensive, tendono a svuotarsi di contenuto e si prestano a fraintendimenti e ad usi impropri.

Ciò non significa tuttavia che dietro queste parole non si nascondano contenuti assai significativi. Certamente la loro rapida diffusione corrisponde all'esigenza reale di categorizzare e qualificare un insieme di fenomeni tecnologici, produttivi, commerciali, culturali, sociali e -perché no- politici che assumono un rilievo via via crescente.

Non si tratta quindi di liquidare con sufficienza le parole chiave solo perché abusate. Occorre invece comprendere e razionalizzare i contenuti che esse celano, e compiere uno sforzo per analizzare problematiche e opportunità.

Questo lavoro si propone di fornire un contributo, sia pur limitato, alla comprensione del fenomeno "multimedia", focalizzandosi principalmente sulle problematiche legate all'interazione fra persone e gruppi. L'assunzione fondamentale è che, in ultima analisi, si vuole aumentare l'efficacia della comunicazione, utilizzando le opportunità tecnologiche in accordo alle esigenze specifiche degli utenti.

La trattazione sarà volutamente informale e potrà essere criticata perché superficiale e imprecisa. Si tenga presente che l'obiettivo primo è di introdurre uno "stile" di approccio ai problemi e di evidenziare questioni aperte, non di fornire improbabili soluzioni generali o informazioni tecniche di dettaglio.

2. Le parole chiave

2.1 Media

Nell'accezione a cui si fa riferimento in questo contesto, un *medium* è un mezzo che supporta il trattamento, la memorizzazione e la trasmissione di informazione. Questa definizione generica merita una breve discussione, soprattutto per sgombrare il terreno da un possibile equivoco: la tendenza a confondere il concetto di medium con quello di supporto fisico e tecnologico.

Ciò che qualifica un medium e lo caratterizza rispetto ad un altro è il tipo di informazione assieme alle operazioni che possono essere compiute sull'informazione stessa.

(*) Prof. ing. Francesco Tisato -Dipartimento di Scienze dell'Informazione - Università degli Studi di Milano- Milano

Il supporto fisico e tecnologico è relativamente poco rilevante, purché fornisca funzionalità adeguate alle operazioni da eseguire sull'informazione.

La distinzione è di particolare rilievo, in quanto le caratteristiche logiche e funzionali del medium ne costituiscono l'aspetto fondamentale in termini di efficacia comunicativa per l'utente, e sono quindi il principale elemento per la scelta di un medium rispetto ad un altro. La scelta di un particolare supporto dipende invece dalla facilità e immediatezza di utilizzo (che costituiscono ovviamente un prerequisito chiave per l'efficacia), dal rapporto prestazioni/costo, da considerazioni di mercato, e così via.

Ad esempio: un testo è un medium che supporta caratteri alfanumerici, in diversi formati, su cui possono essere eseguite operazioni base di lettura e (non sempre) di scrittura; un testo può inoltre consentire operazioni più complesse, come la ricerca di particolari sequenze di caratteri, il posizionamento in base a un sistema di coordinate (pagina, capitolo, eccetera), l'indicizzazione. Che poi il testo sia fisicamente supportato da un foglio di carta o dallo schermo di un personal computer è poco rilevante in termini di efficacia comunicativa, se si prescinde da fattori come, ad esempio, quelli ergonomici (la lettura di un testo stampato è, allo stato dell'arte, più gradevole della lettura di uno schermo; d'altra parte, la modifica di un testo è probabilmente più agevole su un supporto elettronico che su un supporto cartaceo).

Una categorizzazione dei media in base a specifiche caratteristiche verrà introdotta nel seguito. Per il momento, basti citare i media più diffusi:

- testo;
- immagini statiche;
- immagini in movimento;
- grafica statica;
- grafica in movimento;
- suono;
- voce.

2.2 Multimedialità

Il termine "multimediale" non è difficile da definire, almeno in modo informale. Esso identifica la coesistenza di flussi di informazione supportati da media diversi e logicamente correlati.

L'approccio multimediale non è certo una invenzione recente. L'uso di figure in un libro, piuttosto che quello di una lavagna in una lezione, sono casi emblematici di multimedialità "ante litteram". E' chiaro che l'evoluzione tecnologica consente di integrare media con modalità fino a ieri impensabili, ed apre quindi prospettive affascinanti purché non si perda di vista il fatto che l'obiettivo fondamentale è sempre quello di aumentare l'efficacia della comunicazione.

L'integrazione di flussi informativi supportati da media diversi pone problemi non banali, a livello sia concettuale che tecnologico, su cui si tornerà in maggior dettaglio nel seguito. Si può intuire che, se l'integrazione di testo e immagini statiche è cosa agevole (come è dimostrato da parecchi secoli), non lo è altrettanto quella fra testo e immagini in movimento. Le difficoltà sono sia di ordine concettuale, legate alle diverse modalità di fruizione dell'informazione, che di ordine tecnologico. In questo senso, esse derivano in gran parte dal fatto che i supporti tecnologici sono stati storicamente sviluppati con l'obiettivo di ottimizzare il rapporto prestazioni/costo rispetto a media specifici, e non in un'ottica di integrazione.

2.3 Spazio Mediale

Il termine "spazio mediale" si riferisce in modo specifico all'utilizzo di flussi informativi finalizzato alla comunicazione fra attori, siano essi individui o gruppi.

Uno spazio mediale è, in ultima analisi, uno spazio fisico o logico accessibile a una molteplicità di attori, che possono interagire utilizzando media diversi.

Una sala di riunione è un ovvio esempio di spazio mediale "fisico". Ma anche la stampa quotidiana è, in senso lato, uno spazio mediale "virtuale". Sebbene apparentemente banali, questi esempi si prestano ad alcune considerazioni di un certo interesse.

A prima vista, la differenza principale fra i due esempi sembra risiedere nella localizzazione fisica - concentrata nel primo caso, distribuita nel secondo. In realtà la differenza è più sottile, e risiede essenzialmente nelle modalità di interazione fra utenti, che sono a loro volta legate alle caratteristiche dei media utilizzati.

Nella sala conferenze i media principali sono la voce e il contatto visivo diretto, entrambi caratterizzati dal fatto che le azioni di un attore vengono percepite "in tempo reale" dagli altri. Nel caso della stampa i media sono testi e immagini, e le informazioni generate da un attore vengono percepite dagli altri in tempo differito. Come si vedrà nel seguito, le caratteristiche dei flussi informativi influiscono pesantemente sugli stili di interazione fra utenti e, in ultima analisi, sull'efficacia della comunicazione.

L'errore di valutazione deriva evidentemente da limitazioni tecnologiche, che hanno creato storicamente l'abitudine a considerare la contiguità fisica come condizione necessaria per l'interazione in tempo reale (sola eccezione significativa, la comunicazione vocale). Nel momento in cui l'evoluzione tecnologica, e in particolare quella dei sistemi di comunicazione, tende a rendere sempre meno rilevante l'aspetto "spaziale" del problema, non si deve dimenticare che le caratteristiche dei flussi informativi costituiscono comunque un fattore fondamentale ai fini dell'efficacia

della comunicazione, e che quindi la costruzione di spazi mediali innovativi deve sfruttare le tecnologie, ma deve anche e soprattutto tener conto delle esigenze degli attori. Il rischio è quello, tipico, di sviluppare "soluzioni alla ricerca di un problema".

3. L'architettura degli spazi mediali

Si è già evidenziato come l'uso diffuso e dispersivo dei termini renda difficile l'identificazione corretta dei problemi. E' quindi utile definire una "architettura" degli spazi mediali, sottolineando che essa non va vista come una architettura "concreta" direttamente traducibile in termini tecnico-progettuali (anche se su ciò sarebbe interessante discutere). In questo contesto, si tratta di una architettura "di riferimento" finalizzata alla identificazione e alla classificazione dei problemi.

Lo spazio mediale può essere visto come articolato in quattro livelli principali di astrazione (fig. 1). Al livello più alto sono evidenziati gli attori (cioè gli utenti che "agiscono" nello spazio mediale) e gli stili di interazione, legati ad obiettivi di efficacia comunicativa.

Al secondo livello si evidenziano i modelli di comunicazione (come avviene la comunicazione fra attori, come è strutturata l'informazione), finalizzandoli alla realizzazione degli stili di interazione desiderati.

Al terzo livello si considerano in maggior dettaglio i *media* utilizzati, esaminandone le caratteristiche in funzione dei requisiti dei livelli superiori.

Infine, al quarto livello si evidenziano le caratteristiche delle *piattaforme*, cioè dei supporti tecnologici usati per realizzare in concreto gli spazi mediali.

Nelle sezioni che seguono si discutono brevemente le problematiche relative a ciascun livello e le correlazioni fra esse.

4. Gli stili di interazione

Parlare di "stili di interazione" significa esaminare gli aspetti contenutistici, comportamentali e motivazionali delle interazioni fra attori. Un esame approfondito di questi aspetti esula dagli obiettivi di questo lavoro e dalle competenze dello scrivente. E' tuttavia possibile tentare una classificazione che, per quanto "ingenua", fornisce alcuni criteri di valutazione utili ai fini della discussione successiva.

La classificazione si riferisce in modo specifico ad attività che coinvolgono più attori, di cui alcuni svolgono un ruolo di pilotaggio ("leadership"). Si tratta di una situazione non del tutto generale, ma sicuramente ricorrente in numerose strutture organizzative e sociali. Nel seguito si farà riferimento frequente ad esempi tipici dell'area della didattica, che il lettore potrà facilmente estrapolare ad altre situazioni.

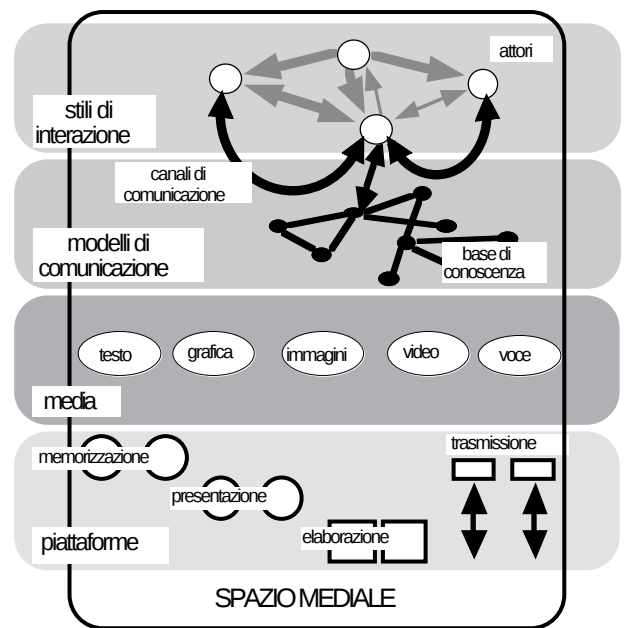


Figura 1 Architettura di uno spazio mediale

Lo "stile" di interazione fra attori può essere classificato in base a due caratteristiche fondamentali, come schematizzato in fig. 2:

- il livello di *guida* metodologica e contenutistica;
- il livello di *supporto* psicologico e motivazionale.

In funzione di quanto sia accentuata una o l'altra delle due caratteristiche, si possono identificare quattro classi fondamentali di stili di interazione:

- stile di *prescrizione*: è riassumibile in un atteggiamento del tipo "ti spiego in dettaglio quanto bisogna fare". Il livello di guida è elevato, mentre non lo è quello di supporto. Si pensi ad una tipica lezione "ex cathedra", in cui metodologie e contenuti vengono spiegati in dettaglio prestando (a volte) poca o nessuna attenzione

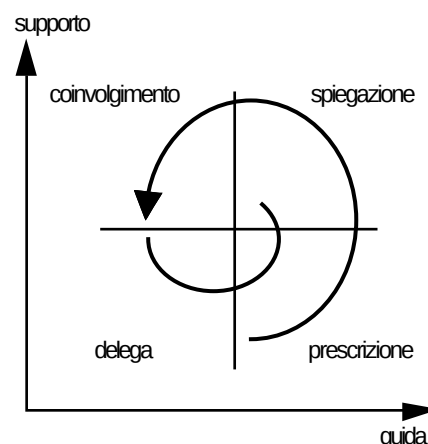


Figura 2 Classificazione degli stili di interazione

al coinvolgimento degli ascoltatori;

- stile di *spiegazione*: è riassumibile in un atteggiamento del tipo "facciamolo insieme". I livelli sia di guida che di supporto sono elevati. Il caso tipico è quello delle attività didattiche individuali o per piccoli gruppi, svolte con la presenza di un docente o di un tutor, attenti sia agli aspetti contenutistici che a quelli motivazionali;
- stile di *coinvolgimento*: l'atteggiamento è questa volta "sei in grado di operare in modo autonomo, ma sappi che sono sempre a disposizione se necessario". Il livello di guida è basso, mentre è alto quello di supporto. In ambiente didattico, è la situazione di studenti che vengono stimolati a svolgere autonomamente ricerche su temi di loro interesse, fornendo adeguata motivazione e disponibilità a dare supporto su richiesta;
- stile di *delega*: siamo qui all'atteggiamento "questo è quanto bisogna fare, e lo devi fare da solo". Guida e supporto sono entrambi a basso livello. E' (purtroppo) la situazione di molti studenti che svolgono la tesi di laurea; si tratta peraltro di uno stile che in molti casi presenta notevole efficacia.

Non è possibile definire quale sia lo stile ottimale; in effetti, come schematizzato in fig. 2, l'iter tipico di un processo di comunicazione efficace passa spesso attraverso i quattro stili: dalla impostazione iniziale di un nuovo problema (prescrizione), alla spiegazione per affrontare difficoltà sia di contenuti che psicologiche, al coinvolgimento quando è stata raggiunta una buona padronanza dei contenuti ma sussiste l'esigenza di supporto motivazionale, fino alla delega in presenza di autonomia sotto ogni punto di vista. Dopodiché il ciclo riprende per affrontare un nuovo problema, tipicamente di difficoltà superiore.

5. Requisiti degli stili

La costruzione di uno spazio mediale comporta dunque, in prima istanza, l'identificazione del, o degli, stili di interazione più efficaci in relazione alle esigenze degli attori. Si tratta ora di evidenziare quali siano i requisiti principali dei diversi stili, per poter effettuare le scelte più opportune ai livelli inferiori dell'architettura di riferimento.

E' possibile identificare tre requisiti principali:

- interattività della comunicazione;
- strutturazione dell'informazione;
- espressività della presentazione.

La rilevanza dei requisiti può essere messa in relazione allo "spazio degli stili" discusso in precedenza, in modo da identificare, in funzione dello stile desiderato, quali siano i requisiti richiesti (fig. 3).

La *interattività della comunicazione* è essenziale soprattutto dal punto di vista del supporto motivazionale. Con "interattività della comunicazione" si indica la

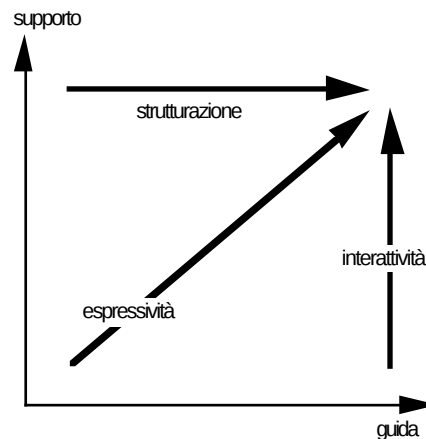


Figura 3 Requisiti degli stili

granularità e la rapidità con cui un attore percepisce gli effetti delle azioni compiute da un altro attore, e con cui può a sua volta reagire.

Ad esempio, una comunicazione telefonica è fortemente interattiva, in quanto l'azione (cioè quello che viene detto) è immediatamente percepita e la reazione può essere immediata (si può interloquire in qualsiasi istante). E' quindi una interazione "a grana fine" e "in tempo reale". La posta elettronica è moderatamente interattiva, poiché la trasmissione di informazione avviene per blocchi (messaggi) e con un certo ritardo; il ricevente esamina i messaggi ricevuti e invia le risposte in un tempo successivo da lui controllato. Si tratta quindi di una interazione "a grana grossa" e "differita".

Una videoconferenza consente interazione a grana fine e in tempo reale, anche se talvolta il sistema di videoconferenza impone alcune "regole di protocollo" che disciplinano la possibilità di interloquire. Viceversa, una trasmissione televisiva "classica" è del tutto priva di interattività. Questo esempio è utile per anticipare un concetto che verrà ripreso nel seguito: non bisogna confondere le caratteristiche dei media (che in entrambi i casi sono immagini in movimento e voce) con l'interattività a livello di utente finale.

La *strutturazione dell'informazione* riveste un ruolo fondamentale in termini di guida sul metodo e sui contenuti. Con "strutturazione dell'informazione" si intende il modo in cui i singoli frammenti di informazione (immagini, parti di testo...) sono correlati in una struttura logica che guida il reperimento e l'utilizzo delle informazioni. Ad esempio, una banca dati e un pacchetto per autoistruzione guidata trattano informazione fortemente strutturata. Viceversa, una conversazione telefonica tratta informazione non strutturata (se non nella mente degli interlocutori).

E' intuitivo che un colloquio telefonico (forte interattività, scarsa strutturazione dell'informazione) potrebbe essere molto efficace in termini di motivazione per uno studente con buone basi conoscitive, mentre

una lezione di inquadramento tenuta per telefono avrebbe una efficacia quanto meno discutibile. Anche qui si tratta di un caso apparentemente banale, ma facilmente estrapolabile a problematiche più complesse. Ad esempio: esiste una reale motivazione per utilizzare una (costosa) tecnologia di video conferenza per lo svolgimento di lezioni "ex cathedra" nello stile "prescrizione"?

I requisiti di interattività e di strutturazione comportano, come si vedrà, vincoli sui modelli di comunicazione utilizzati.

L'espressività è un requisito fondamentale sia dal punto di vista della guida che da quello del supporto. Con "espressività" si indica l'efficacia nella trasmissione dei contenuti. Si tratta di un requisito che ha forti impatti sia in termini di guida che in termini di supporto: la comunicazione di informazione attraverso immagini e video favorisce sia la trasmissione dei contenuti che la motivazione.

Il requisito di espressività comporta vincoli sulla scelta dei media.

Vale la pena di rilevare che l'aumento di livello di un requisito non costituisce necessariamente un fattore positivo. Ad esempio, una forte interattività e/o una forte strutturazione dell'informazione possono costituire, nello stile di delega, un fattore di disturbo. Analogamente, l'uso di media fortemente espressivi (grafica, immagini...) può essere inefficace per utenti motivati ed esperti: un caso tipico è quello delle interfacce interattive di tipo grafico, a cui gli utenti esperti preferiscono i più efficienti, anche se meno espressivi, comandi da tastiera.

6. I modelli di comunicazione

Una volta identificati gli stili di interazione più opportuni ed i corrispondenti requisiti, si tratta di scegliere i modelli di comunicazione che lo spazio mediale deve supportare. In fig. 1 si è evidenziata in modo intuitivo una distinzione "classica" fra comunicazione ottenuta attraverso una "base di conoscenza" ben strutturata (o, in parole povere, un archivio di informazione) accessibile a tutti gli attori, e comunicazione ottenuta attraverso "canali" che consentono la trasmissione diretta di informazione fra gli attori stessi. Casi tipici sono, rispettivamente, l'accesso ad una banca dati pubblica e il colloquio telefonico.

Si tratta però di una distinzione grossolana, che trae le sue origini da una situazione storica in cui tecnologie, media e fornitori di prodotti e servizi erano chiaramente diversificati. Da una parte, il mondo dell'informatica, centrato sulla memorizzazione dell'informazione in archivi. Dall'altra parte, il mondo delle comunicazioni, centrato sulla trasmissione diretta dell'informazione senza memorizzazione.

L'evoluzione tecnologica e di mercato (se si vuole, la

"rivoluzione multimediale"...) ha reso la situazione molto più complessa; è quindi opportuno compiere, anche qui, uno sforzo di categorizzazione. In uno spazio mediale la comunicazione fra attori prevede comunque la condivisione di informazioni e/o risorse di qualche tipo. I modelli di comunicazione possono essere categorizzati in base a due criteri fondamentali, che si riconducono in ultima analisi alle caratteristiche di interattività e di *strutturazione dell'informazione*.

In base al primo criterio, si possono identificare quattro modelli di comunicazione, caratterizzati in funzione di *cosa viene condiviso* fra gli attori nello spazio mediale:

- *condivisione di dati*: gli utenti accedono a informazioni comuni organizzate come basi dati, ipertesti distribuiti (è il caso di World Wide Web), bollettini informativi ("bulletin board"). Un caso limite è quello delle informazioni distribuite su CD-ROM. Si tratta di soluzioni fortemente diversificate, ma caratterizzate comunque dalla bassa interattività: l'interazione fra utenti è infatti possibile in quanto un utente modifica informazioni che vengono successivamente esaminate da altri utenti su loro iniziativa. Si noti che ciò richiede un alto livello di motivazione da parte degli utenti destinatari dell'informazione modificata. In molti casi (tipico quello del CD-ROM!) l'informazione è scarsamente movimentata. Il modello architetturale tipico è il "client-server";
- *condivisione di messaggi*: gli utenti generano messaggi che vengono temporaneamente memorizzati per essere poi recuperati dal, o dai, destinatari finali (modello architetturale "store-and-forward", tipico della posta elettronica). Il livello di interattività fra utenti è medio, nel senso che la presenza di un messaggio può venire notificata al, o ai, destinatari finali; l'interazione fra utenti avviene comunque in tempo differito, con ritardi che possono andare da alcuni secondi a alcune ore;
- *condivisione di applicazioni*: gli utenti accedono allo stesso programma applicativo modificando l'informazione gestita dal programma e, ciò che è essenziale, percepiscono "in tempo reale", cioè con un ritardo praticamente trascurabile, gli effetti delle azioni eseguite dagli altri utenti. Il modello architetturale è quello cosiddetto a "blackboard", facendo riferimento alla metafora della lavagna su cui tutti scrivono e da cui tutti leggono (si noti peraltro che il termine "blackboard" viene talvolta usato per sistemi a condivisione di dati, quindi non in tempo reale). Sistemi di supporto al lavoro cooperativo, e in particolare programmi di elaborazione di testi in cui più utenti accedono contemporaneamente allo stesso documento, ricadono in questa categoria. Il livello di interattività è ovviamente molto alto;

- *condivisione di canali*: gli utenti accedono a canali di comunicazione che trasmettono direttamente l'informazione fra due o più utenti, senza alcuna elaborazione (almeno dal punto di vista logico). E' ovvio che tutti i sistemi di comunicazione "classici" rientrano in questa categoria, caratterizzata da flussi informativi in tempo reale e, per canali bidirezionali, da elevata interattività.

In base alla strutturazione dell'informazione, si possono identificare tre categorie principali:

- *informale*: l'informazione condivisa non è strutturata (più precisamente, la strutturazione è interamente lasciata all'utente finale). E' il caso della posta elettronica (in cui i singoli messaggi non sono logicamente correlati), dei sistemi di "chatting" (comunicazione interattiva testuale in rete), e in generale dei canali di comunicazione diretta;
- *strutturazione dei dati*: l'informazione condivisa è strutturata, come nel caso delle banche dati, degli ipertesti, dei bollettini, eccetera;
- *strutturazione del processo*: la comunicazione e la strutturazione dell'informazione è pilotata da un modello del processo comunicativo o della struttura organizzativa. E' il caso dei sistemi di supporto al lavoro cooperativo, e in particolare del "workflow" in cui i flussi informativi sono gestiti in accordo con il flusso delle attività in una organizzazione.

La categorizzazione dei modelli di comunicazione fornisce le indicazioni essenziali per sceglierli in funzione degli stili di interazione desiderati e dei loro requisiti.

7. I media

Dei media si è già detto, ma il discorso va ripreso per evidenziare che, in linea di principio e alla luce delle prospettive di evoluzione tecnologica, la scelta dei media è largamente indipendente dai modelli di comunicazione e, a maggior ragione, dagli stili di interazione. In altri termini, una volta identificati prima gli stili, poi i modelli in funzione dei requisiti degli stili, i media possono essere scelti in funzione della loro espressività rispetto ai contenuti. La possibilità concreta di utilizzare uno specifico medium con uno o l'altro modello di comunicazione dipende ovviamente dalle caratteristiche funzionali e dal rapporto prestazioni/costo dei supporti tecnologici disponibili.

Proseguendo nella logica della classificazione, è utile evidenziare quattro caratteristiche fondamentali che individuano un medium:

- il tipo di informazione;
- le *operazioni* che l'utente può compiere;
- il *sistema di coordinate* che consente di identificare particolari frammenti di informazione;
- le caratteristiche *temporali*.

L'identificazione del *tipo* di informazione (testo, immagine, suono...) è sufficientemente intuitiva. E' utile sottolineare che la scelta del tipo di informazione è fortemente legata, da un lato, al requisito di espressività; va fatta quindi in funzione dei contenuti e dello stile di interazione desiderato. Dall'altro lato, la scelta del tipo di informazione può essere pesantemente condizionata dalle caratteristiche delle piattaforme tecnologiche.

L'identificazione delle *operazioni* che l'utente può compiere è forse meno immediata, ma non meno rilevante. Le operazioni fondamentali sulle informazioni sono: *generazione, manipolazione, trasmissione, memorizzazione, presentazione*. Non tutti i media presentano, da questo punto di vista, caratteristiche omogenee. In particolare, le modalità di generazione e manipolazione hanno forti impatti sulla possibilità di utilizzare un particolare medium nel contesto di uno specifico modello di comunicazione.

Ad esempio, la differenziazione fra immagine e grafica è motivabile se si definisce come "immagine" un medium per il quale la generazione di informazione può avvenire solo per "cattura" dall'ambiente esterno (tipicamente usando uno scanner o una telecamera), mentre "grafica" viene definito come un medium che consente di sintetizzare e manipolare informazione (tipicamente attraverso un editor grafico). E' chiaro che la distinzione è fondamentale quando, ad esempio, lo stile di interazione richieda elevata interattività e si voglia utilizzare un modello di comunicazione a condivisione di applicazione ("whiteboard") in cui gli utenti interagiscono manipolando informazione condivisa. E' altrettanto chiaro che si tratta di classificazioni opinabili e mutevoli: proprio in questo caso, l'evoluzione tecnologica tende a far svanire la differenziazione, che può essere utile tenere per ora in considerazione dei fini pratici.

Il sistema di *coordinate* consente di identificare particolari frammenti di informazione. Ad esempio, nel caso già citato di un testo, il sistema di coordinate può essere costituito dal semplice numero di carattere, o dal numero di pagina, o dal numero di capitolo e di sezione. Nel caso di un filmato, l'identificazione di un frammento può avvenire mediante una coordinata temporale.

Le caratteristiche del sistema di coordinate hanno forti correlazioni con il requisito di strutturazione dell'informazione: per un testo, la strutturazione è molto bassa nel caso di sistema di coordinate "lineare" (numero di carattere), è invece alta nel caso di sistema di coordinate "gerarchico" (capitolo, sezione...). Per inciso, si fa notare che queste considerazioni portano a identificare, per ciascun medium, un certo numero di "sottoclassi" che hanno in comune le proprietà base, ma sono diversificate per alcuni aspetti specifici (testo non strutturato, testo strutturato per capitoli, eccetera).

Si tratta in sostanza di utilizzare i principi fondamentali della analisi e della progettazione orientate agli oggetti.

Le caratteristiche dei sistemi di coordinate sono particolarmente critiche, in relazione alla strutturazione dell'informazione, quando si entra in un'ottica multimediale: occorre allora stabilire correlazioni fra informazioni supportate da media diversi, che possono essere caratterizzati da sistemi di coordinate diversi.

Si può intuire che stabilire e gestire una relazione fra due frammenti di testo è più semplice che non stabilire e gestire una relazione tra un frammento di testo e un frammento di filmato. Si tratta di un problema tipico dei sistemi ipermediali, che sono appunto costruiti come reti di relazioni tra frammenti di informazione supportati da media diversi. Occorre garantire (e questo è un problema di piattaforma tecnologica) che i riferimenti tra frammenti di informazione possano essere definiti in modo generale o, più precisamente, in modo indipendente dai media ("media-independent").

L'aspetto temporale porta ad operare una distinzione fra media intrinsecamente *asincroni* (testo, immagini statiche...) e media intrinsecamente *sincroni* (voce, suoni, video...). Nel primo caso le operazioni sull'informazione (in particolare, la fruizione) possono avvenire secondo una tempificazione scelta arbitrariamente dall'utente. Nel secondo caso, le operazioni sono vincolate ad una tempificazione strettamente associata all'informazione stessa (si può scegliere liberamente la velocità con cui leggere un testo, non quella con cui ascoltare un parlato).

Per evitare equivoci, è opportuno sottolineare che le caratteristiche temporali del medium non sono direttamente correlate alle caratteristiche di interattività del modello di comunicazione. Come già evidenziato, esiste una tradizione storica che porta a considerare i media sincroni (voce, video...) candidati naturali per essere usati nell'ambito di un modello di comunicazione fortemente interattivo, come quello basato su canali diretti. Viceversa, i media asincroni (testo, immagini statiche...) ben si prestano ad essere utilizzati nell'ambito di modelli scarsamente interattivi, come quello a condivisione di dati.

Ciò è indubbiamente vero, ma non necessario. Ad esempio, in un sistema di messaggistica vocale un medium intrinsecamente sincrono (voce) è utilizzato nell'ambito di un modello di comunicazione a interattività medio-bassa. Ovviamente, la fruizione dell'informazione da parte del destinatario non potrà non tener conto delle caratteristiche di sincronicità del medium; ma ciò non è in relazione con il livello di interattività fra mittente e destinatario. Viceversa, in un sistema di editing cooperativo di testo (modello "whiteboard") un medium intrinsecamente asincrono viene utilizzato nell'ambito di un modello di comunicazione fortemente interattivo.

8. Le piattaforme

Il termine "piattaforma" rientra anch'esso nel novero delle parole chiave usate e abusate. In questo contesto, si intende per "piattaforma" un insieme di componenti hardware e software che forniscono le funzionalità di base su cui costruire le funzionalità applicative disponibili all'utente dello spazio mediale.

Si parla normalmente di piattaforme "standard" o "industry standard". Usualmente, in effetti, le piattaforme utilizzate non vengono sviluppate "ad hoc"; si utilizzano invece prodotti (hardware e software) e servizi disponibili sul mercato. Il problema è, per chi utilizza le piattaforme, quello di scegliere la più opportuna; per chi le costruisce, quello di identificare le funzionalità di base utili per supportare le funzionalità applicative che soddisfano le esigenze dell'utente finale, sulla falsariga di quanto discusso in precedenza.

Una discussione anche sommaria delle caratteristiche tecnologiche delle diverse piattaforme esula dagli obiettivi di questo lavoro. Come al solito, è possibile delineare una sommaria classificazione, per passare poi a una breve discussione dei problemi aperti e delle linee di sviluppo.

Le piattaforme tecnologiche possono essere raggruppate in base alle quattro funzionalità principali che esse supportano, come delineato in fig. 1:

- *trasmissione* dell'informazione;
- *elaborazione* dell'informazione;
- *presentazione* dell'informazione;
- *memorizzazione* dell'informazione.

Fin qui si tratta di concetti evidenti. Ciò che sarebbe desiderabile è la disponibilità o di una piattaforma integrata, in grado di supportare le quattro funzionalità principali; o di quattro piattaforme distinte, ognuna delle quali supporta una funzionalità specifica.

Le cose non sono così semplici nella pratica.

In primo luogo, la distinzione tra funzionalità non è così chiara. Se ci si pone in un'ottica "informatica", un Sistema Operativo (ad esempio, Unix o Window) tende a supportare in modo integrato tutte le quattro classi di funzionalità. Tuttavia, se ci si pone in un'ottica "comunicazionistica", un sistema di commutazione fornisce un insieme di servizi di trasmissione che si integrano (o si sovrappongono) in modo non chiaro con quelli forniti da un Sistema Operativo. In generale, è difficile se non impossibile identificare una piattaforma in grado di supportare tutte le funzionalità necessarie per un sistema complesso; il problema è dunque quello della coesistenza e dell'integrazione di piattaforme diverse, in un'ottica di "interoperabilità".

In secondo luogo, e con riferimento specifico a quanto discusso finora, le piattaforme esistenti non sono "indipendenti dai media". Senza pretendere di discutere il problema in modo esaustivo, alcuni esempi possono essere di qualche utilità.

Per quanto riguarda la trasmissione, non sempre media diversi possono essere trasmessi in modo omogeneo. Il caso limite (e ormai in via di superamento) è forse quello della trasmissione voce analogica e della trasmissione digitale di dati, che ha portato storicamente allo sviluppo di sistemi di comunicazione completamente indipendenti, e radicalmente diversi per le funzionalità, per le modalità di gestione, per la codifica dell'informazione, per le caratteristiche di tempo reale, per la banda trasmissiva, e così via.

In particolare, gli aspetti di tempo reale e di banda trasmissiva meritano alcune brevi considerazioni. Si consideri ad esempio il caso in cui nello spazio mediale debba coesistere un modello di comunicazione a canale per la voce, e un modello di comunicazione a messaggi per dati e immagini (che è esattamente quanto avviene in un gran numero di situazioni reali). Purtroppo la voce è un medium intrinsecamente sincrono che richiede una banda trasmissiva limitata; il supporto trasmissivo "classico" (rete telefonica voce) garantisce in effetti la trasmissione con ritardo costante su banda vocale. Viceversa, per dati e immagini è essenziale una elevata banda media (e soprattutto di picco) che può essere superiore di alcuni ordini di grandezza a quella necessaria per la voce, mentre la costanza del ritardo di trasmissione non è rilevante. Il supporto trasmissivo "classico" (rete dati) possiede infatti queste caratteristiche. Il risultato finale è che si è reso necessario l'utilizzo di piattaforme di comunicazione separate, con funzionalità del tutto disgiunte, rendendo di fatto impossibile la realizzazione di spazi medial integrati - non a caso è normale che su una scrivania coesistano un apparecchio telefonico e un personal computer connesso in rete, e che l'integrazione tra i flussi informativi sia interamente a carico dell'utente.

L'evoluzione in atto (ma non completata) verso sistemi di trasmissione digitali (in particolare, ATM), a banda flessibile e con la possibilità di calibrare la qualità di servizio, apre la strada verso la possibilità di scegliere liberamente media e modelli di comunicazione, e di costruire quindi spazi medial ritagliati in base alle esigenze degli utenti finali.

Per quanto riguarda l'elaborazione e la *presentazione* dell'informazione la situazione è ormai meno critica, una volta che le informazioni siano comunque codificate in forma digitale secondo standard più o meno affermati (e sono quindi manipolabili), e che siano disponibili piattaforme hardware e software a larghissima diffusione che consentono di presentare media diversi in modo integrato.

Qualche problema si pone ancora per la cattura "intelligente" dell'informazione (riconoscimento del parlato e delle immagini), nel senso che possono essere necessari supporti hardware specializzati, ma comunque facilmente integrabili in piattaforme standard.

Per quanto riguarda la memorizzazione della informazione, è ormai in atto la tendenza a sviluppare piattaforme largamente indipendenti dai media: è il caso dei sistemi di gestione dati orientati a oggetti, delle estensioni multimediali dei sistemi relazionali, e dei sistemi ipertestuali distribuiti.

9. Conclusioni: uno scenario possibile

Quanto sarebbe desiderabile -ed a cui occorre comunque tendere in una visione strategica- è uno scenario di integrazione architetturale sintetizzato in fig. 4.

Alla base dell'architettura, le piattaforme tecnologiche standard gestiscono oggetti (informazioni, risorse di calcolo, canali trasmissivi, periferiche...) realizzati con tecnologie specifiche. Il punto chiave è che le piattaforme, anche se diverse in termini di struttura interna, devono mettere a disposizione dei livelli superiori un insieme omogeneo e standardizzato di funzionalità per il trattamento, la trasmissione, l'elaborazione di oggetti supportati da media diversi (testi, immagini, video, voce...). E' quindi possibile strutturare e gestire le informazioni secondo criteri e metodologie che dipendono dai media utilizzati, ma non dalle particolari tecnologie utilizzate dalle piattaforme.

A un livello più alto, l'informazione viene organizzata come un insieme di oggetti "astratti" che possono essere correlati, memorizzati e trasmessi indipendentemente dalle caratteristiche dei media. E' quindi possibile organizzare "spazi medial" orientati alle esigenze dell'utente finale, e quindi agli stili di interazione ed ai

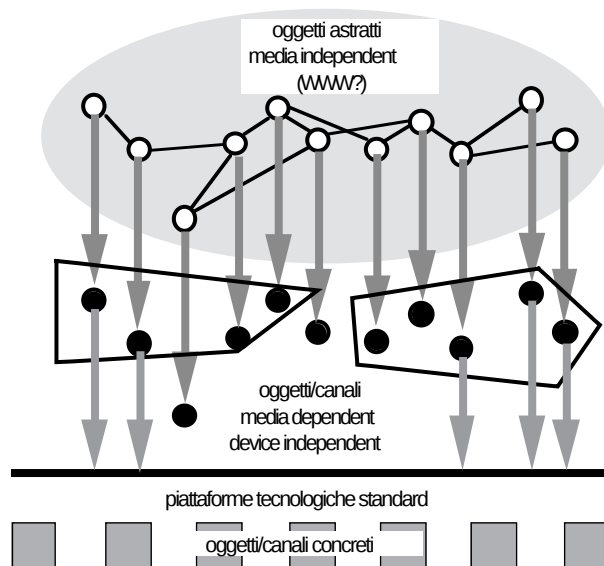


Figura 4 Un possibile scenario di integrazione

modelli di comunicazione più adeguati, scegliendo, e integrando caso per caso, le piattaforme più opportune in base a considerazioni tecnologiche, di mercato e di rapporto prestazioni/costo.

Uno scenario di questo tipo può sembrare utopico, ma in realtà si assiste ad una significativa evoluzione in questa direzione.

Ad esempio, il fenomeno Internet e l'evoluzione dei servizi disponibili in rete (in primis, come già evidenziato, World Wide Web) forniscono già una piattaforma ad ampia diffusione per la definizione, la gestione e la trasmissione di informazioni organizzate come reti di oggetti largamente indipendenti da media e tecnologie.

I problemi aperti derivano da diversi fattori, riconducibili in sostanza alla distanza fra "culture" corrispondenti a livelli differenti dello spazio mediale ed a realtà tecnologiche e di mercato fortemente diversificate e, spesso, contrapposte.

Un elemento di notevole complessità deriva dalla eterogeneità del mercato, in cui intervengono operatori provenienti da segmenti storicamente diversificati: fornitori di servizi informatici, di servizi di comunicazione, di hardware, di software. Poiché esiste ormai una chiara percezione del fatto che l'area degli "spazi mediali" offre ampie possibilità di sviluppare prodotti a forte valore aggiunto, e poiché tale area costituisce di fatto l'intersezione dei diversi segmenti di mercato, tutti gli operatori tendono ad "occuparla" fornendo soluzioni coerenti con la propria filosofia di prodotto e, spesso, in contrapposizione. Ad esempio, il problema di quanta "intelligenza" risieda sulle stazioni di lavoro piuttosto che nel sistema di comunicazione non è una disputa accademica, ma corrisponde a un preciso problema di mercato: il valore aggiunto è di competenza nel primo caso dei produttori di informatica, nel secondo dei fornitori di servizi di comunicazione.

Non è certo possibile prevedere chi risulterà vincente; è tuttavia possibile affermare che la capacità di supportare e realizzare spazi mediali evoluti, sfruttando le opportunità tecnologiche in modo da soddisfare efficacemente le esigenze dell'utente finale, costituisce un fattore chiave di competitività.

La Videotelefonía a basso bitrate nella normativa ITU-T⁽¹⁾

G. Russo, P. Talone (*)

Un fattore di primaria importanza per lo sviluppo dei servizi di videocomunicazione, è costituito dai sistemi videotelefonici in grado di operare con frequenze di cifra inferiori a 64 kbit/s.

Questi trovano applicazione immediata per servizi su rete telefonica commutata (anche con accesso analogico) e, in misura ancora da definire, su rete radiomobile numerica.

Questi sistemi, inoltre, costituiscono un efficiente strumento tecnologico intorno al quale definire svariati ambienti multimediali.

Nel presente lavoro vengono illustrati i recenti progressi della normativa in merito, prodotta in ambito ITU-T. Particolare attenzione viene dedicata alle codifiche video di nuova generazione che costituiscono la novità tecnologica di maggior rilievo.

1. Il terminale videotelefonico

Le problematiche riguardanti le codifiche videotelefoniche sono state affrontate, congiuntamente a quelle di videoconferenza, in ambito ITU-T, a partire dai primi anni 80, all'interno dello Study Group 15 – Transmission Systems and Equipements.

Come è tradizione in ambito ITU, la definizione di un servizio implica sia la scelta di una rete di comunicazione, sia la specifica di una serie di aspetti funzionali. Nel caso della generica videocomunicazione, il servizio è stato definito, in primo luogo, per la rete ISDN, coinvolgendo le funzionalità, anche multimediali, riportate nello schema di fig. 1, tratto da [1].

Oltre alla struttura generale del terminale [1], i principali blocchi funzionali previsti sono relativi a:

- interfacce relative alla rete di comunicazione [2];
- segnalazione di rete [2];
- moltiplicazione e demoltiplicazione dei flussi informativi inerenti ai media coinvolti nella comunicazione [3];
- gestione dei media coinvolti nella comunicazione e precisamente:
 - co-decodifiche audio AV.250 in [4], ovvero [5], [6], [7];
 - co-decodifiche video [8];
 - gestione di altri media (grafica, immagini statiche, ecc.) [9];
- segnalazione tra terminali [10], [11].

Lo schema di fig. 1 può rappresentare sia un terminale videotelefonico che per videoconferenza. In particolare la serie di Raccomandazioni T.120 [9], in corso di approvazione presso lo Study Group 8, definisce un sistema di gestione di immagini statiche, grafica, testi ed altri servizi adatti a realizzare una comunicazione multimediale condividendo le risorse di trasmissione con l'audio ed il video.

Il sistema H.320 prevede comunicazioni multipunto tramite unità di rete per videoconferenza (MCU) anch'esse in corso di definizione.

Il sistema H.320 è inoltre previsto per operare a bitrate (velocità di cifra) diversi (multipli di 64 kbit/s) ottenendo diverse qualità audio e video che costituiscono il principale elemento discriminante tra un sistema videotelefonico ed uno per videoconferenza.

In particolare, per quanto riguarda la codifica video, che risulta sempre il punto debole di questi sistemi, l'impiego della ormai storica Raccomandazione H.261, esclude il soddisfacimento dei requisiti di un servizio videotelefonico su una rete con prestazioni inferiori ai 64 kbit/s offerti da ISDN. D'altro canto però, è disponibile una connettività numerica, su rete telefonica analogica, di circa 28 kbit/s, con possibilità di compressione e correzione d'errore, offerta dai modem di recente generazione [12], [13], [14], [15]. Inoltre da qualche anno esiste la disponibilità di prodotti videotelefonici commerciali che offrono, in qualche modo, un servizio

(*) Ing. Giuseppe Russo, ing. Paolo Talone -Fondazione Ugo Bordoni- Roma

(1) Lavoro eseguito nell'ambito della convenzione tra la Fondazione Ugo Bordoni e Telecom Italia.

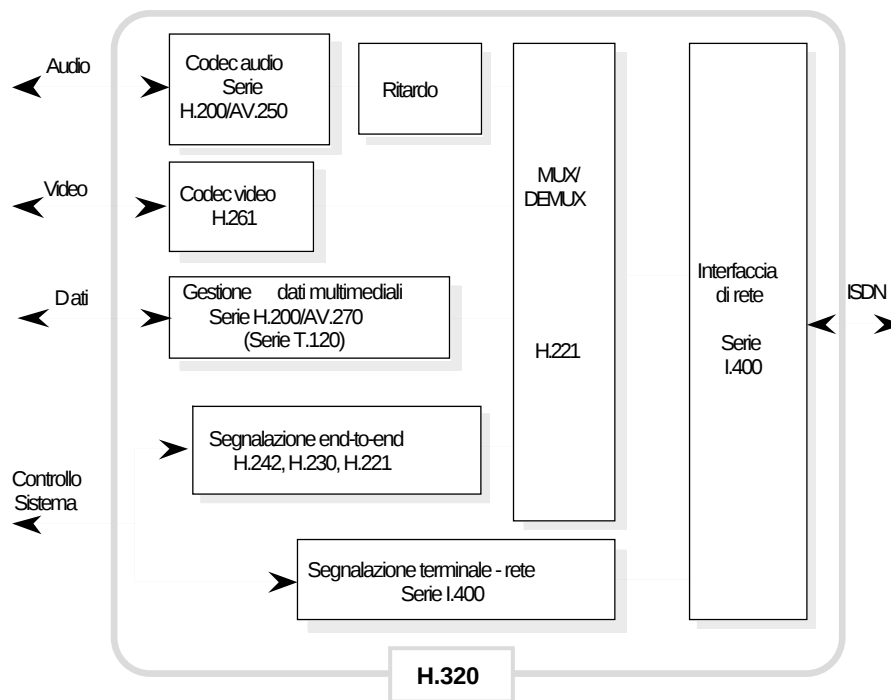


Figura 1 Schema funzionale di un generico terminale H.320 "Visual Telephone System"

su rete telefonica analogica. Questi elementi hanno spinto l'ITU-T a definire un nuovo servizio, adatto a portanti a bassa velocità.

Nel Gruppo di Studio 15 è stato pertanto costituito, all'inizio dell'attuale quadriennio di studio (1993-1996), un gruppo di lavoro sul tema "Very Low Bitrate Visual Telephony" (Question 15/2).

Il servizio è previsto operare su rete telefonica pubblica (General Switched Telephone Network - GSTN) e, in prospettiva, su reti radiomobili numeriche.

Anche in questo caso la definizione del servizio implica la scelta della rete di comunicazione e la specifica delle sue differenti componenti funzionali. Conseguentemente, nell'ambito del gruppo di lavoro 15/2 sono operanti i seguenti sottogruppi ciascuno dei quali affronta un aspetto particolare del problema:

- struttura generale del terminale: gruppo "Video H.324";
 - interfacce relative alla rete di comunicazione: gruppi "DTE/DCE interface", "V.8/V.34";
 - moltiplicazione e demoltiplicazione dei flussi informativi inerenti ai media coinvolti nella comunicazione: gruppo "Multiplex, System Control";
 - gestione dei media coinvolti nella comunicazione e precisamente:
 - co-decodifiche audio: gruppo "Speech";
 - co-decodifiche video: gruppi "General Video", "Audio visual tests", "Video bitstream verification";
 - segnalazione tra terminali: gruppo "H.246".
- Sono inoltre operanti:
- un gruppo "Requirements" relativo a requisiti generali

di servizio;

- un gruppo "Video H.320 interoperability" per l'interoperabilità con terminali videotelefonici tipo H.320 su ISDN;
- un gruppo "Mobile audio-visual terminal" che si occupa dell'adattamento dell'intero sistema ad una rete radiomobile numerica;
- un gruppo "H.263/L" relativo a codifiche video innovative per il lungo termine.

Lo schema a blocchi funzionali del sistema, tratto da [16], è riportato in fig. 2 ed è relativo ad una soluzione "a breve termine" per terminali su rete GSTN.

In tale figura sono anche indicate le sigle delle Raccomandazioni che specificano le sezioni del sistema [16], [17], [18], [19], [20].

Anche in questo caso si tratta dello schema generale di un terminale dotato di codificatori audio e video e di funzioni multimediali, affidate, quest'ultime, alla serie di Raccomandazioni T.120. Le principali differenze rispetto allo schema di fig. 1, risiedono nell'uso di nuovi codec audio e video e di un modem a 28,8 kbit/s al posto di un'interfaccia ISDN. In questo caso comunque, il bitrate previsto fa presagire, nonostante la maggior efficienza del codec, una qualità video adatta per un impiego prevalentemente videotelefonico.

Entro il quadriennio di studio dell'ITU, si prevede la definizione di analoghi blocchi funzionali e delle relative Raccomandazioni, ove differiscano da quelle di fig. 2, per un terminale operante su rete radiomobile numerica.

Sia per la GSTN sia per la rete radiomobile numerica, è inoltre allo studio una soluzione a lungo termine basata su tecniche di codifica video innovative.

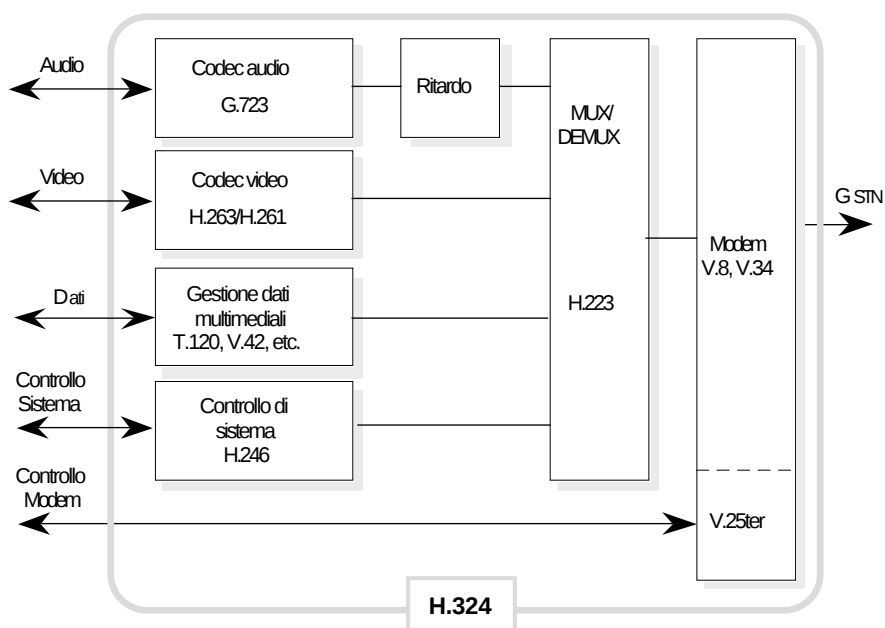


Figura 2 Schema funzionale di un terminale videotelefonico H.324

2. Interoperabilità tra terminali

L'interoperabilità tra terminali è ovviamente tema centrale per un servizio di videotelefonazione. A questo proposito il problema principale riguarda la salvaguardia del "presunto" parco di terminali H.320 già operanti.

L'interoperabilità tra terminali H.324 connessi a GSTN e terminali H.320 connessi a ISDN, è a tutt'oggi prevista attraverso due modalità (cfr fig. 3):

- La futura probabile definizione di un nuovo terminale su ISDN, che permetta la cosiddetta modalità "dual-mode". Questa consente la coesistenza dei blocchi funzionali H.320/H.261 e H.324/H.263. L'interoperabilità è prevista su base H.324/H.263.
- L'interoperabilità tra i "vecchi" terminali H.320 ed i terminali H.324 tramite una unità di adattamento che si può supporre essere apparato di rete (H.320 IWA: InterWorking Adapter). La struttura di questa unità non è a tutt'oggi definita. La funzione di base è costituita dal trasferimento della codifica H.261 tra i due ambienti convertendo i protocolli e la codifica audio. Infatti la co-decodifica H.261, come si dirà nel seguito, è necessaria anche per i terminali H.324. Si possono comunque prefigurare anche IWA in grado di convertire le codifiche video H.261/H.263.

Per quanto riguarda l'interoperabilità H.320/H.324 tramite IWA, la soluzione adottata, come è tradizione dell'ITU, tende a rendere i "nuovi" terminali H.324 compatibili con i "vecchi" terminali H.320. Ricordando che i terminali H.320/H.261 devono operare sia in codifica che in decodifica con il formato QCIF e che il formato CIF è opzionale, è previsto che:

- nei terminali H.324 siano obbligatoriamente presenti entrambi gli algoritmi di co-decodifica H.261 ed H.263;
- valgano, per i terminali H.324, le seguenti regole per il formato delle immagini (cfr. Appendice):
 - I codificatori devono essere in grado di trattare uno dei due formati QCIF o sub-QCIF (più adatto al basso bitrate) ma non sono obbligati a trattarli entrambi. Il trattamento del formato CIF è, invece, sempre opzionale. L'informazione sul formato è trasferita dal codificatore al decodificatore in fase di inizializzazione della comunicazione.
 - I decodificatori devono obbligatoriamente essere in grado di decodificare entrambi i formati QCIF e sub-QCIF. Il trattamento del formato CIF è opzionale.
 - Per ciò che riguarda la presentazione su schermo, una prima ipotesi di regole, è descritta in [21], per i terminali che non implementano il formato CIF, sarebbero ammessi schermo e memoria di quadro sia a risoluzione QCIF sia sub-QCIF. I terminali QCIF possono presentare le sequenze sovracampionando il formato sub-QCIF oppure presentare il formato sub-QCIF circondato da un bordo nero. I terminali sub-QCIF possono sottocampionare il formato QCIF oppure possono presentarne solamente una porzione.

Il vantaggio di tale soluzione è che la co-decodifica del formato QCIF è obbligatoria ma è consentito l'impiego di schermi e memorie di quadro di formato sub-QCIF per terminali a basso costo [22].

In fig. 3 sono illustrate graficamente le modalità di interconnessione previste.

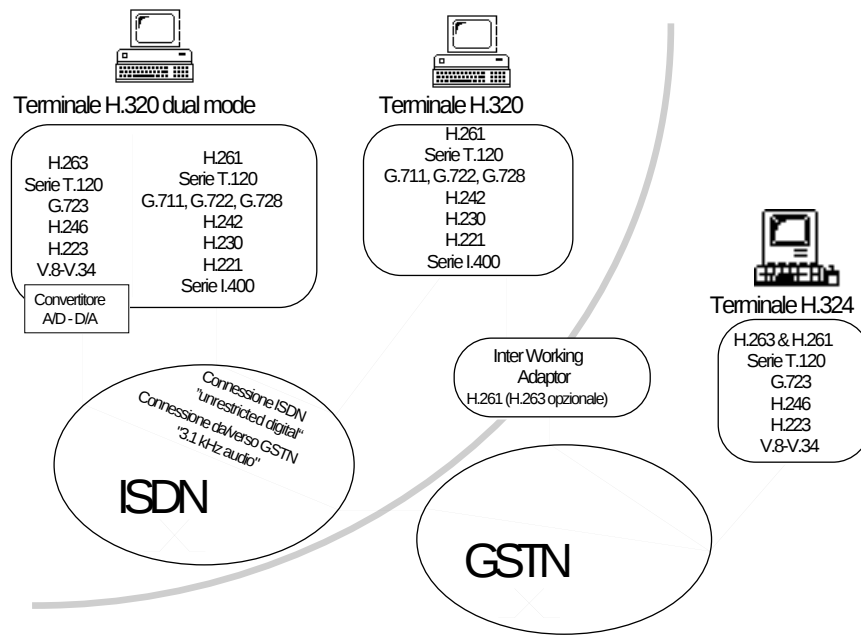


Figura 3 Interoperabilità tra terminali videotelefonici

3. Algoritmi di codifica video

La parte maggiormente critica di un terminale videotelefonico, è costituita dal co-decodificatore video. In questa sezione pertanto, dopo un breve richiamo al

codec H.261, sono illustrate le principali innovazioni introdotte nell'algoritmo H.263.

In fig. 4 è riportato uno schema a blocchi del codificatore video H.261 ricavato sulla base di quanto specificato in [8].

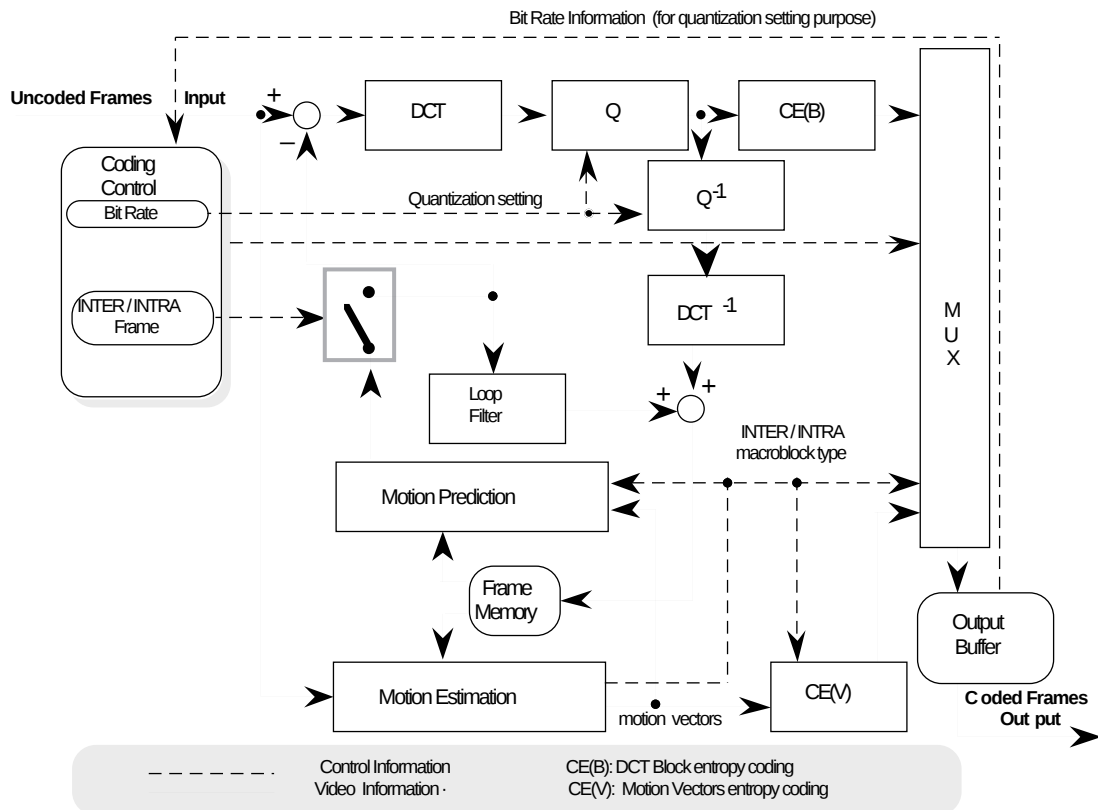


Figura 4 Schema a blocchi del codificatore H.261

L'algoritmo H.261 è basato sull'impiego della trasformata DCT su blocchi 8x8 e compensazione del movimento su macroblocchi 16x16 con vettori di spostamento di lunghezza multipla di un pixel.

Sono inoltre introdotte le due modalità fondamentali di codifica di quadro: intraframe ed interframe.

In fig. 5 viene riportata la sequenza dei quadri video.

Nella modalità intra (Quadri I) la codifica avviene

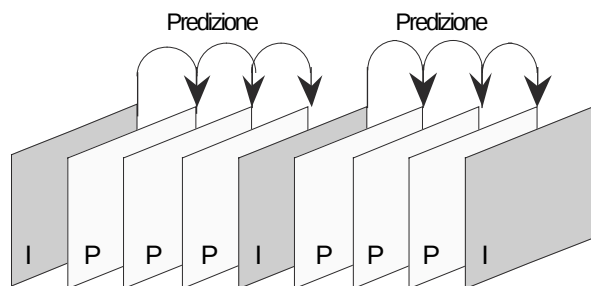


Figura 5 Sequenza di quadri

sostanzialmente come un'immagine statica (DCT / quantizzazione senza predizione temporale). Il raggiungimento di elevati tassi di compressione è legato alla perdita irreversibile di informazione visuale operata dalla quantizzazione dei coefficienti DCT che, come è noto, risulta operazione irreversibile. Nello schema di fig. 4 pertanto, con il blocco Q^{-1} si intende un'operazione di ripristino di dinamica (moltiplicazione dei coefficienti quantizzati per i relativi passi di quantizzazione) e non una (impossibile) inversione matematica dell'operazione di quantizzazione. La codifica entropica (metodo di Huffman) consente poi di codificare i coefficienti quantizzati, in maniera efficiente.

Nella modalità inter (quadri P) il quadro viene predetto da quello immediatamente precedente.

A ciascun macroblocco viene associato un vettore di spostamento che specifica la posizione, nel quadro precedente decodificato, del macroblocco di predizione. L'individuazione di questo è realizzata mediante tecniche di stima di moto basate su criteri di similitudine spaziale (correlazione). Il processo di predizione può essere modificato da un filtro spaziale bidimensionale (filtro d'anello) che opera su ciascun macroblocco predetto.

Oggetto della codifica sono l'errore di predizione (differenza tra il macroblocco originale ed il macroblocco predetto) codificata con DCT / quantizzazione / codifica entropica) e le informazioni sulla predizione medesima (vettori di spostamento dei macroblocchi con codifica entropica).

Anche all'interno dei quadri P è possibile che alcuni macroblocchi siano codificati Intra e non Inter nel caso in cui risulti inefficiente la procedura di predizione.

La frequenza di quadro e conseguentemente il numero di quadri P interposti tra due quadri I consecutivi non

risultano fissati ma possono variare nel tempo (al pari del tipo di quantizzazione) secondo opportune strategie di codifica. Tale possibilità è di particolare importanza quando il requisito primario della codifica è la generazione di un bit-rate costante nel tempo.

È noto che l'H.261 è, storicamente, il primo algoritmo standard di codifica video del tipo a trasformata. Le tecniche previste (predizioni temporali) risultano quindi meno sofisticate delle corrispondenti negli standard successivi quali MPEG [23] e, soprattutto H.263 [17].

Questo è particolarmente avvertibile nelle condizioni critiche provocate dalla disponibilità di bassi bitrate.

L'ITU-T pertanto, dopo aver acquisito dall'ISO l'esperienza delle codifiche MPEG, ha ritenuto, anche sotto la spinta di prodotti commerciali, di compiere un ulteriore passo definendo una nuova codifica mirata a bitrate inferiori a 64 kbit/s.

Gli obiettivi sono distinti nella definizione di una soluzione a lungo termine ed una a breve già definita in [17] che fa riferimento, per l'intero sistema, allo schema a blocchi di fig. 2 [16] ed al relativo elenco di Raccomandazioni.

La soluzione a breve termine è basata sull'impiego di tecniche di codifica di tipo DCT a blocchi e compensazione del movimento. Queste tecniche, pur mantenendosi in un ambito tradizionale, sfruttano appieno un decennio di esperienza in materia, introducendo opzioni sofisticate soprattutto nella compensazione del moto e nella predizione temporale.

Nel lungo termine saranno probabilmente utilizzate tecniche di codifica video più avanzate (codifiche ad "oggetti" basate su metodi di analisi e sintesi dell'immagine) ed i lavori saranno presumibilmente condotti in stretta cooperazione con il progetto MPEG4 dell'ISO.

Per quanto riguarda l'algoritmo di codifica video H.263, è stata definita una procedura di base, presente obbligatoriamente in tutti i terminali H.324, che ricalca sostanzialmente la Raccomandazione H.261.

Sono previste poi alcune opzioni di codifica, negoziabili nella fase di set-up del colloquio tra terminali, che consentono di incrementare le prestazioni della codifica base in termini di rapporto tra qualità visuale e bitrate. L'algoritmo di base e tutte le opzioni previste sono descritte in [17] ed illustrate in fig. 6, ricavata sulla base di quanto specificato in [24] e [17].

Nel seguito verranno brevemente illustrati alcuni degli aspetti più significativi relativamente all'algoritmo.

In codifica si prevedono, come in H.261, quadri intra e quadri inter, con codifica di tipo DCT a blocchi 8x8 e compensazione di movimento su macroblocchi 16x16. La differenza principale rispetto ad H.261 consiste nella compensazione, non più prevista a pixel interi ma a frazioni di mezzo pixel. Non è inoltre previsto nessun filtro nell'anello di compensazione.

Le opzioni negoziabili, che costituiscono l'innovazione rispetto ad H.261, sono le seguenti:

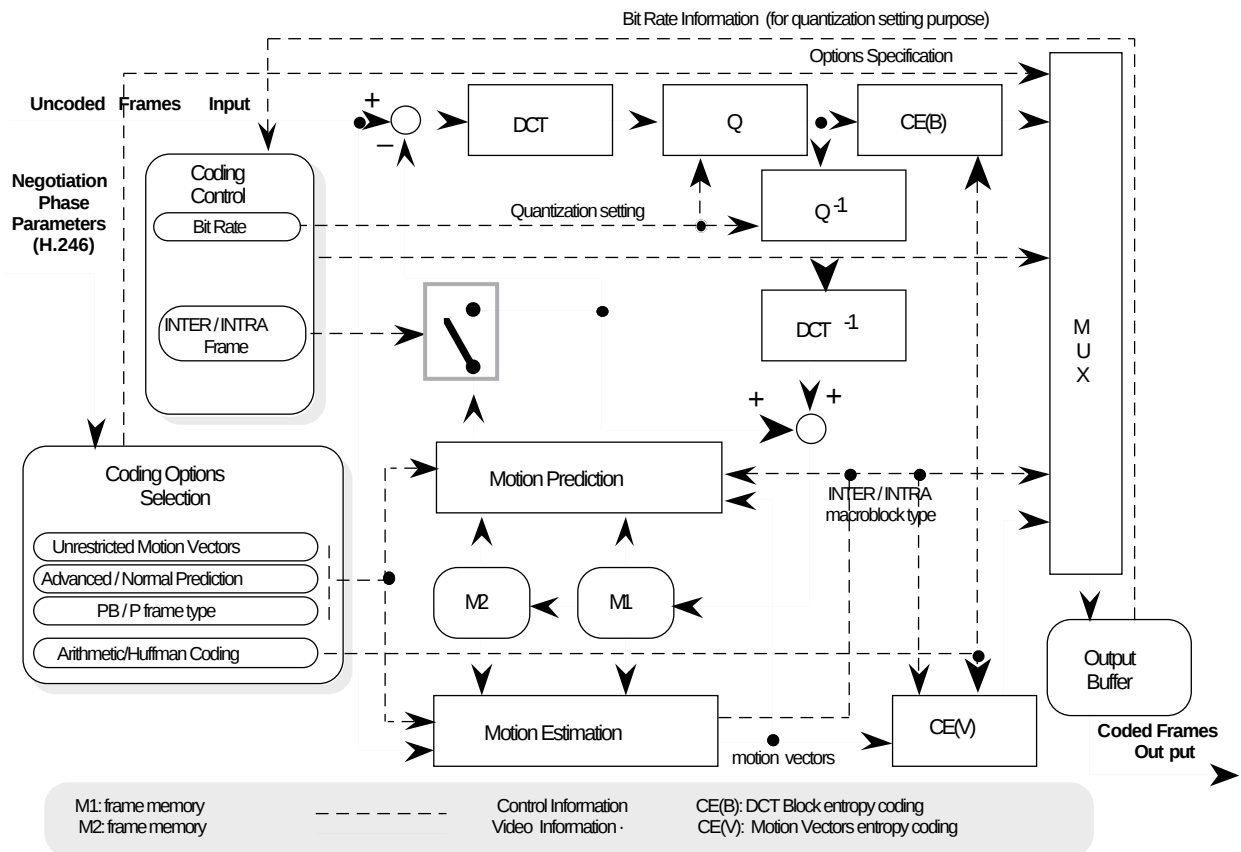


Figura 6 Schema a blocchi del codificatore H.263

- “Unrestricted Motion Vector mode” che prevede vettori di spostamento che possano puntare anche fuori del quadro.
- “Advanced Prediction mode” che prevede la compensazione di moto su blocchi 8x8 e con modalità "overlapped".
- “Syntax-based Arithmetic Coding mode” che prevede l'impiego per la codifica entropica, al posto della tecnica di Huffman, la più recente codifica aritmetica.
- “PB-frames mode” che prevede l'uso di frames (PB) con predizione temporale bidirezionale.

In fig. 7 sono illustrati esempi di “Unrestricted Motion Vectors”. La possibilità di considerare vettori di movimento che puntano parzialmente fuori dell'area d'immagine, consente di ottenere significativi incrementi di presentazioni nel trattamento dei bordi, in sequenze in cui siano presenti movimenti dovuti a panoramiche della telecamera (panning) o ad oggetti che attraversano la scena. In questi casi infatti, risulta più efficiente effettuare predizione temporale con l'uso di blocchi a cavallo dei bordi e quindi solo parzialmente definiti. Il completamento di tali blocchi avviene tramite semplice replica dei pixel di bordo. Nel caso di movimenti traslatori, tale replica costituisce una efficiente predizione dei pixel del quadro successivo, evitando inoltre che nei blocchi completati si abbiano discontinuità spaziali. La presenza di queste ultime

infatti, generando alte frequenze DCT oggetto di forti quantizzazioni, darebbe origine ad artefatti notevoli nei blocchetti ricostruiti.

La seconda opzione definita prevede, in primo luogo, sia la possibilità di inviare un unico vettore di movimento



Figura 7 Esempi di "Unrestricted Motion Vectors"

per ciascun macroblocco (16x16) sia la possibilità di associare un vettore di movimento per ciascuno dei quattro blocchi 8x8 componenti il macroblocco.

In secondo luogo l'opzione prevede l'impiego di una modalità di predizione denominata "overlapped".

Le due modalità, illustrate in fig. 8, sono state previste solo per un uso congiunto; sono stati infatti prodotti documenti atti a dimostrare che l'uso di una sola delle due modalità risulta inefficace per il rapporto qualità/bitrate ottenibile.

La modalità "overlapped" prevede un cambiamento sostanziale nella predizione dei pixel. Nelle tecniche classiche (H.261 ed MPEG) per ciascun macroblocco viene inviata la differenza rispetto al macroblocco (16x16) del quadro precedente individuato dal vettore di movimento (predizione). Utilizzando una notazione simile a quella definita in fig. 8 il macroblocco predetto in H.261 può essere espresso semplicemente da:

$$Mpp_n(i, j) = Mp_{n-1}(MV, i, j)$$

La differenza viene calcolata tra pixel omologhi senza particolari accorgimenti.

Nell'opzione in oggetto, invece, il calcolo della predizione per ciascun blocco 8x8 avviene tenendo conto della posizione di ciascun pixel rispetto ai bordi del blocco. Per ciascun pixel del blocco predetto viene calcolata una media pesata dei pixel omologhi dei seguenti blocchi:

- il blocco predetto a partire dal vettore di movimento del blocco a cui il pixel appartiene (tecnica usuale);
- i due blocchi predetti a partire dai vettori di movimento dei due blocchi più prossimi al pixel, tra i quattro adiacenti (sopra, sotto, destra, sinistra).

In [17] sono specificate le matrici dei coefficienti di

pesatura dei pixel.

Nel caso in cui un macroblocco possieda un unico vettore di movimento, la procedura "overlapped" si applica invariata associando il vettore medesimo ad ognuno dei blocchi 8x8 componenti.

Con riferimento alla fig. 8, l'impiego di MV_0 al posto di MV_4 trova una giustificazione pratica dal fatto che in decodifica il vettore MV_4 risulta disponibile solo dopo la ricezione di un'ulteriore striscia di blocchetti 8x8. Questo fatto comporterebbe nel decodificatore l'aggiunta di una memoria ed un ulteriore ritardo.

La terza opzione prevede l'impiego, nel codificatore entropico, di un algoritmo di codifica aritmetica al posto della usuale codifica di Huffman come già avviene in alcune opzioni dello standard JPEG per codifica di immagini statiche [25]. Un codificatore aritmetico riesce in genere ad ottenere un guadagno tra il 5 ed il 10% rispetto ad un codificatore di Huffman.

La quarta opzione, illustrata in fig. 9, prevede nella codifica la presenza di quadri cosiddetti B, in analogia a quanto avviene per gli standard MPEG. I quadri B sono un particolare tipo di quadri inter nei quali la predizione, su base stima di moto, avviene sia rispetto al quadro precedente (I o P) sia rispetto al quadro successivo (P).

Tuttavia, a differenza di MPEG, è previsto un unico quadro B tra coppie adiacenti di quadri P. Inoltre al fine di evitare l'overhead necessario per codificare il quadro B come unità dati indipendente, i dati di questo quadro sono moltiplicati con quelli del successivo quadro P; si parla quindi di quadri PB. Infine, le coppie di vettori di movimento associate ai quadri B (predizione bidirezionale) non vengono trasmesse ma sono ricavate semplicemente dimezzando i corrispondenti vettori spostamento dei due quadri rispetto ai quali si effettua la predizione.

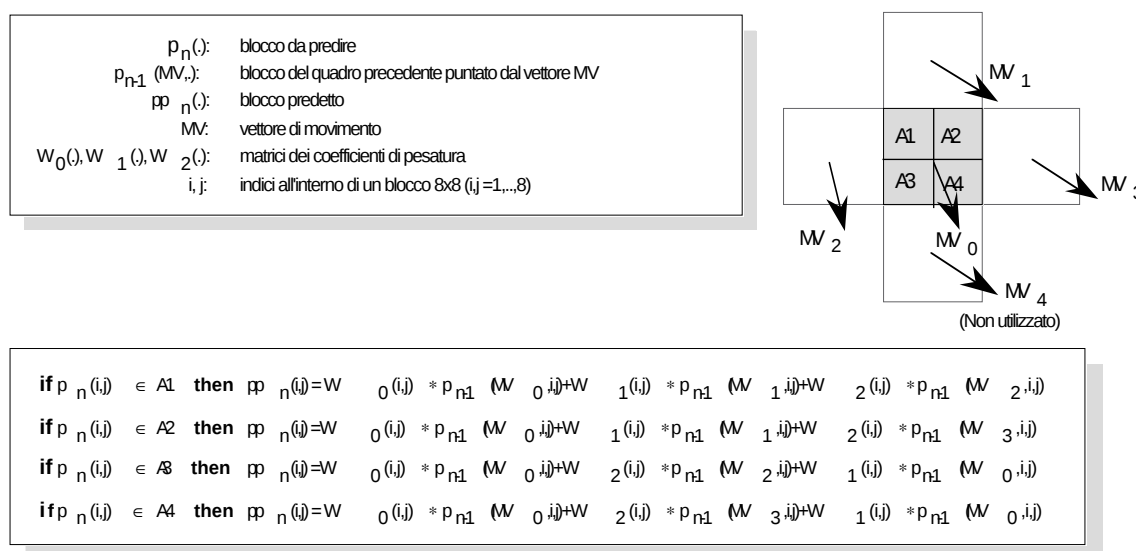


Figura 8 Principio dell'"Overlapped Motion Compensation"

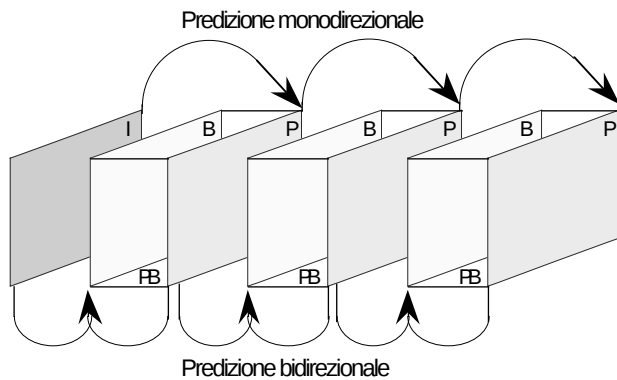


Figura 9 Sequenza quadri PB

Il modello di codificatore, utilizzato per la messa a punto finale della raccomandazione, è denominato TM (Test Model) ed è evoluto a partire da una prima versione, prodotta alla fine del 1993, indicata come TMN 1 (mutuata dal modello SIM3 prodotto dal progetto COST 211 ter) [26], [27]. Questa presentava, rispetto alla codifica H.261, le seguenti differenze sostanziali:

- la compensazione del moto con vettori di spostamento di lunghezza multipla di 1/2 pixel e con l'impiego di tecniche di interpolazione;
- la possibilità di suddividere la compensazione del moto di un macroblocco 16x16 nella compensazione di ciascuno dei quattro blocchetti 8x8 componenti.

Successivamente è stata dimostrata l'inefficienza, in termini di rapporto tra qualità psicovisuale e complessità aggiuntiva, della codifica ottenuta prevedendo l'opzione b) con quella in cui tale opzione non era prevista. Conseguentemente è stato deciso di semplificare il modello eliminando la suddetta opzione (TMN 2) [28].

Inoltre, una nuova tavola per la codifica entropica (Huffman) delle informazioni relative ai macroblocchi codificati è stata adottata durante la riunione tenutasi in Norvegia nel mese di Luglio 1994. Al termine di detto incontro è stata quindi prodotta la terza revisione del Test Model indicata come TMN 3 che è strutturalmente identica al TMN 2 tenendo però in conto la variante suddetta [29], [22].

Al termine della riunione di Piscataway nell'ottobre 1994, è stata prodotta la quarta revisione del modello, TMN 4 [30].

Durante la riunione di Leidschendam nel mese di Gennaio 1995 sono stati definiti i testi dell'intera serie di Raccomandazioni da portare all'approvazione entro l'anno corrente (Novembre 1995). È stata inoltre prodotta un'altra versione del modello di codificatore, TMN5 [24], che definisce la procedura base della codifica H.263 e fa riferimento alle appendici della Raccomandazione H.263 per ciò che riguarda le specifiche delle opzioni.

Va infine ricordato che per l'uso dell'algoritmo H.263 su rete radiomobile numerica sono allo studio tecniche di codifica combinate sorgente-canale al fine di rendere la codifica compatibile con i tassi d'errore caratteristici

dei canali a radiofrequenza. Un riassunto dello stato dei lavori ed una bibliografia sono reperibili in [31].

4. Conclusioni

L'H.263 costituisce probabilmente un punto di arrivo per le codifiche video a trasformata, riassumendo l'esperienza acquisita in numerosi anni di studi e realizzazioni. Attualmente il lavoro è essenzialmente rivolto alla valutazione delle prestazioni, in relazione alle differenti opzioni, in termini di qualità psicovisuale e bitrate, rapportati alla complessità aggiuntiva introdotta nel codificatore.

Le attività su questo tema, svolte in Fondazione Bordini, prevedono l'impiego di codec software realizzati su workstation e test di qualità soggettiva. Le prestazioni finora rilevate, mettono in luce una qualità accettabile per servizi videotelefonici già attorno a 20 kbit/s. Considerando quindi l'efficienza della codifica, potrebbe sorgere il dubbio che le innovazioni apportate possano comportare l'obsolescenza degli Standard precedenti. Il problema non si pone per MPEG le cui applicazioni, rivolte alle alte qualità di immagine, non trarrebbero sostanziale beneficio dall'uso di opzioni simili a quelle introdotte in H.263. Per quanto riguarda, invece, H.261, la diffusione dei nuovi terminali H.320 Dual Mode (H.261 - H.263 compatibili) fa prevedere alcune conseguenze.

La qualità superiore dell'H.263, a parità di bitrate, e la possibilità di impiego anche di risoluzioni adatte alla videoconferenza (CIF) prefigurano una diffusione di questo standard anche per applicazioni su ISDN e reti ad alta velocità. Questa scelta, oltre a comportare un generale miglioramento della qualità di videocomunicazione a bitrate medio bassi, può consentire, con l'impiego delle procedure T.120, la realizzazione di applicazioni multimediali avanzate, quali teledidattica, telelavoro o altre, con soli 64 kbit/s. Ciò potrebbe comportare per l'H.261, una sorte analoga a quella della codifica FAX G2, inclusa e funzionante a tutt'oggi nei FAX G3, ma inutilizzata a causa della scomparsa del parco di terminali G2.

Un sostanziale passo in avanti nelle codifiche video è atteso dalla cosiddetta soluzione a lungo termine, con uso di tecniche innovative, prevista anche dall'ITU, ma sostanzialmente studiata a confronto in MPEG 4 ed in diversi progetti di ricerca europei.

Le prestazioni attese, riguardano sia le basse qualità, per un video accettabile sotto i 10 kbit/s, sia le alte qualità per il superamento dei problemi strutturali delle tecniche che impiegano la trasformata a blocchi. Allo stato attuale della ricerca, le principali tecniche innovative fanno capo alla rappresentazione della scena mediante una serie di oggetti o regioni dei quali possono essere descritte forma bidimensionale ed evoluzione temporale, oppure essere generate rappresentazioni

mediante modelli tridimensionali.

La Fondazione Bordoni, svolge studi su codifiche video innovative, essenzialmente con tecniche di segmentazione spazio temporale di regioni. In quest'ambito l'attività, oltre allo studio delle tecniche, è rivolta alla messa a punto di simulazioni di codificatori di tipo avanzato ed a confronti di prestazioni con l'H.263.

Appendice: Formati dell'immagine

I formati previsti dall'ITU per videoconferenza/videotelefonía definiscono una gerarchia di risoluzioni video ottenute per sottocampionamento del formato "di base" della televisione numerica ITU-R 601 4:2:2 interlacciato e non [32]. In questo formato la luminanza (Y) è definita da una matrice di 720 x 576 pixel e le crominanze (U e V) sottocampionate 2:1 lungo le righe per le note proprietà psicovisive, sono definite da due matrici di 360x576 pixel. I campioni delle crominanze coincidono (sono co-locati) con i campioni dispari della luminanza.

I primi due formati, definiti in [8], sono:

- CIF – Common Intermediate Format (n. di campioni 352 x 288 per la luminanza Y e 176 x 144 per ogni crominanza U e V).
- QCIF – Quarter of CIF (n. di campioni 176 x 144 per la luminanza Y e 88 x 72 per ogni crominanza U e V).

Il CIF è ottenuto dal formato ITU-R 601 4:2:2 per sottocampionamento 2:1 bidirezionale della luminanza e sottocampionamento delle crominanze 2:1 lungo le righe e 4:1 lungo le colonne. Il tipo di filtri previsti per il sottocampionamento comporta che ciascuna coppia di campioni delle crominanze sia collocata al centro di quattro campioni di luminanza. Tale formato è denominato 4:2:0 ed è sempre non interlacciato qualsiasi sia il formato 601 da cui è stato generato (interlacciato o non).

Dal formato ottenuto per sottocampionamento, vengono inoltre scartate 8 colonne (4 per lato) di campioni di luminanza con le corrispondenti 4 colonne (2 per lato) di campioni di crominanza. Ciò allo scopo di ottenere un formato d'immagine in cui i campioni di luminanza siano multipli di 16, dimensione del macroblocco usato nella compensazione del movimento. Il formato QCIF è ottenuto dal CIF per sottocampionamento 2:1 bidirezionale della luminanza e delle crominanze.

Il terzo formato sub-QCIF, mirato alla videotelefonía a basso bit rate, è definito in [17]. La risoluzione di sub-QCIF, pari a 128x96 pixel, è ottenibile per sottocampionamento 5:1 del formato comunemente usato nei PC (640x480 pixel) e per sottocampionamento 3:1 delle righe e 11:4 delle colonne del formato CIF (352*288 pixel). La ragione di quanto esposto può essere trovata nella ricerca di compatibilità del terminale videotelefonico con i PC tramite un semplice algoritmo di sovra/sottocampionamento. Viene invece sacrificata la semplicità di conversione da e per il formato CIF assicurando un sottocampionamento semplificato solo per le righe.

Il formato sub-QCIF presenta la stessa struttura a campioni non coincidenti (4:2:0) del CIF e del QCIF ed è normalmente presentato su schermo in forma sovracampionata.

In tab. 1, tratta da [16], vengono riassunti i formati d'immagine previsti per i terminali H.320 ed H.324.

Bibliografia

- [1] *Narrow-band Visual Telephone Systems and Terminal Equipment*. ITU-T Rec. H.320, Marzo 1993.
- [2] ITU-T Rec. I.400 series relevant to ISDN terminals: - *User-Network Interface* (I.410, I.411, I.412); - *User-*

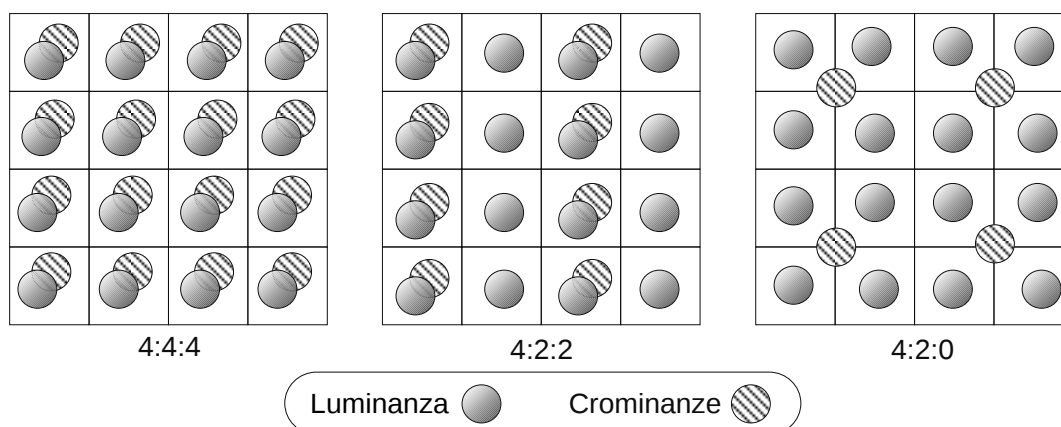


Figura 10 Strutture di campionamento

Picture Format	Luminance pixels	Encoder		Decoder	
		H.261	H.263	H.261	H.263
CIF	352 x 288	Opzionale	Opzionale	Opzionale	Opzionale
QCIF	176 x 144	Obbligatorio ⁽³⁾	Obbligatorio ^(1,2)	Obbligatorio	Obbligatorio ⁽¹⁾
SQCIF	128 x 96 ⁽³⁾	Opzionale ⁽³⁾	Obbligatorio ^(1,2)	Opzionale ⁽³⁾	Obbligatorio ⁽¹⁾

- (1) Opzionale per gli IWA (Interworking Adapters).
 (2) Obbligatoria la codifica di uno dei due formati QCIF o SQCIF; opzionale la codifica di entrambi.
 (3) Il formato SQCIF è definito 128x96 in H.263, in H.261 una qualunque risoluzione inferiore al QCIF, contornata da un bordo nero, è considerata QCIF.

Tabella 1 Formati dell'immagine previsti in H.320 e H.324

- Network Signalling* (I.430, I.431, I.440, I.441, I.450, I.451, I.452).
- [3] *Frame Structure for a 64 to 1920 kbit/s channel in audiovisual teleservices*. ITU-T Rec. H.221, 1990.
- [4] *Framework for Recommendations for Audiovisual services*. ITU-T Rec. H.200, Marzo 1993.
- [5] *Pulse code modulation (PCM) of voice frequencies*. ITU-T Rec. G.711, 1988.
- [6] *Coding of analogue signals by methods other than PCM - 7 kHz audio-coding within 64 kbit/s*. ITU-T Rec. G.722, 1988.
- [7] *Coding of analogue signals by methods other than PCM - Coding of speech at 16 kbit/s using low-delay code excited linear prediction*. ITU-T Rec. G.728, 1992; Annex to 16 kbit/s fixed point specification. G.728, 1994.
- [8] *Video Codec for Audiovisual Services at 64 kbits*. ITU-T Rec. H.261, Marzo 1993.
- [9] *Audio Graphic Conferencing Recommendations*. ITU-T Draft Rec. T.120 series, Febbraio 1995.
- [10] *System for establishing communication between audiovisual terminals using digital channels up to 2 Mbit/s*. ITU-T Rec. H.242, 1993.
- [11] *Frame-synchronous control and indication for audiovisual systems*. ITU-T Rec. H.230, 1990.
- [12] *Procedures for starting sessions of data transmission over the general switched telephone network*. ITU-T Rec. V.8, 1994.
- [13] *A modem operating at data signalling rates of up to 28000 bit/s for use on the general switched telephone network and on leased point-to-point 2-wire telephone type circuits*. ITU-T Rec. V.34, 1994.
- [14] *Error-correcting procedures for DCEs using asynchronous-to-synchronous conversion*. ITU-T Rec. V.42, Marzo 1993.
- [15] *Data compression procedures for Data Circuit Terminating Equipment (DCE) using error-correction procedures*. ITU-T Rec. V.42bis, 1990.
- [16] *Multimedia terminal for low bitrate visual telephone services over the GSTN*. ITU-T Draft Rec. H.324, Maggio 1995.
- [17] *Video coding for narrow telecommunication channels at < 64 kbit/s*. ITU-T Draft Rec. H.263, Maggio 1995.
- [18] *Multiplexing protocol for low bitrate audiovisual communication*. ITU-T Draft Rec. H.223, Maggio 1995.
- [19] *Control of communication between multimedia terminals*. ITU-T Draft Rec. H.246, Maggio 1995.
- [20] *Dual rate speech coder for multimedia telecommunication transmitting at 5,27 & 6,3 kbit/s*. ITU-T Draft Rec. G.723, Maggio 1995.
- [21] Schaphorst, R.: *Report of the Rapporteur of Very Low BitRate Visual Telephony Q.2/15*. Documento ITU-T LBC-95-004, Geneve, 6-17, Febbraio 1995.
- [22] Schaphorst, R.: *Report of the sixth meeting on Question 15/2 in Grimstad, Norway, July 25-27, 1994*. Documento ITU-T LBC-94-132, Agosto 1994.
- [23] *Generic Coding of Moving Picture and Associated Audio Information — Part II: Video*. ISO-IEC IS 13818-2 / ITU-T Rec. H.262, Novembre 1994.
- [24] *Video Codec Test Model TMN 5*. Telenor Research, Documento ITU-T LBC-95-, Gennaio 1995.
- [25] *Digital Compression and Coding of Continuous-tone Still Images — Part I: Requirements and Guidelines*. ISO-IEC IS 10918-1 / ITU-T Rec. T.81, Ottobre 1992.
- [26] *Video Codec Test Model TMN 1*. NCR, Documento ITU-T LBC-94-004.

- | | |
|---|--|
| <p>[27] Bjontegaard, G.: TMN 1: A block-based coder for coding of moving images <i>at very low bit rates</i>. ISO IEC/JTC 1/SC 29/WG 11, Seminario MPEG-4, Parigi, Marzo 1994.</p> <p>[28] <i>Video Codec Test Model TMN 2</i>. NCR, Documento ITU-T LBC-94-124.</p> <p>[29] <i>New MCPBC vlc table for TMN 2</i>. KPN (PTT Research), Documento ITU-T LBC-94-094.</p> <p>[30] <i>Video Codec Test Model TMN 4</i>. NCR, Documento ITU-T LBC-94-, Ottobre 1994.</p> <p>[31] <i>Comments on Draft Recommendation H.26P/M</i>. Bosch, IENT RWTH, Documento ITU-T LBC-95-004, Leidschendam, Gennaio 1995.</p> <p>[32] <i>Encoding parameters of digital television for studios</i>. ITU-R Rec. 601, 1990.</p> <p>[33] Schaphorst, R.: <i>Report of the Rapporteur meeting for Very Low Bitrate Visual Telephony, August 30, 1993</i>. ISO IEC/JTC 1/SC 29/WG 11/MPEG93-775, Brussels, Settembre 1993.</p> <p>[34] Schaphorst, R.: <i>Report of the Rapporteur meeting for Very Low Bitrate Visual Telephony, Geneva, 6-17 September, 1993</i>. ISO IEC/SC29/WG11/MPEG93-774, Brussels, Settembre 1993.</p> <p>[35] Schaphorst, R.: <i>Report of the fifth meeting on Question 15/2 in Paris, 28-30 March, 1994</i>. Documento ITU-T LBC-94-062, Aprile 1994.</p> <p>[36] Schaphorst, R.: <i>Report of the Rapporteur meeting for Very Low Bitrate Visual Telephony, Geneva, 16-27 May, 1994</i>. ISO IEC/SC 29/WG 11/MPEG94-289, Grimstad, Luglio 1994.</p> <p>[37] Schaphorst, R.: <i>Report of the seventh meeting on Question 15/2 in Piscatway, New Jersey, USA, October 10-13, 1994</i>. Documento ITU-T LBC-94-200.</p> <p>[38] Schaphorst, R.: <i>Report of the eighth meeting on Question 2/15 in Leidschendam, Netherland, January 09-12, 1995</i>. Documento ITU-T LBC-95-082.</p> <p>[39] ITU-T Draft Rec. V.25ter, 1995.</p> | <p>MPEG Motion Picture Expert Group
 ISO International Standardization Organization
 JPEG Joint Photographic Expert Group
 TM Test Model
 ITU-R International Telecommunication Union - Radiocommunication sector</p> |
|---|--|

Acronimi

ITU-T	International Telecommunication Union - Telecommunication sector
ISDN	Integrated Services Digital Network
MCU	Multipoint Control Unit
GSTN	General Switched Telephone Network
IWA	Inter Working Adapter
CIF	Common Intermediate Format
QCIF	Quarter of CIF
DCT	Discrete Cosine Transform

Impianti di alimentazione per sistemi di telecomunicazione: stato dell'arte e prospettive future

G. Calzetti, M. Grossoni (*)

Il servizio telefonico, per l'importanza che riveste come servizio di pubblica utilità, deve essere offerto dagli Enti che lo gestiscono con un livello affidabilistico molto elevato.

Di regola per apparati di commutazione e di trasmissione l'obiettivo affidabilistico al quale si tende, espresso in termini di indisponibilità, è di poche ore in quarant'anni con conseguenti notevoli oneri di impianto e di gestione.

Ne consegue che anche la fornitura dell'energia elettrica a detti apparati deve essere offerta con una affidabilità congruente con quella ora detta; di regola per l'alimentazione si fissa un obiettivo di indisponibilità inferiore a un'ora in quarant'anni.

Ragioni di convenienza economica, di praticità e di sicurezza portano normalmente a scegliere come energia elettrica da utilizzare quella offerta dalla rete pubblica di distribuzione che, come noto, è fornita in corrente alternata, senza alcuna garanzia di continuità della fornitura stessa

In alcuni casi, comunque, sia quale contributo delle telecomunicazioni alla soluzione del problema dei risparmi energetici, sia per alimentare apparati in località isolate e non servite dalla rete pubblica di distribuzione, vengono utilizzate fonti alternative di energia quali l'eolica e la solare a integrazione e/o sostituzione di quella tradizionale.

Anche queste fonti alternative, tuttavia, non possono garantire la continuità di fornitura dell'energia.

Ne consegue che i problemi dell'energia nei sistemi di telecomunicazioni sono soprattutto problemi di adattamento dell'energia elettrica, fornita dalla rete pubblica di distribuzione, alle esigenze funzionali degli impianti da alimentare.

In questa relazione vengono descritte le esigenze di alimentazione dei sistemi per telecomunicazioni, sia tradizionali che di nuovo tipo, e vengono analizzati i mezzi scelti per soddisfare tali requisiti.

Vengono inoltre valutati i cambiamenti in atto nelle esigenze di alimentazione di tali sistemi e le relative soluzioni che permetteranno, nei prossimi anni, di soddisfare i nuovi requisiti.

1. Esigenze dei sistemi da alimentare

La quasi totalità degli apparati telefonici attualmente in uso richiede una fornitura di energia elettrica in corrente continua.

Questa energia deve essere fornita con caratteristiche ben definite come limiti di tensione ammessi in regime statico e in regime dinamico (transitori), e come eventuale residuo di corrente alternata sovrapposto alla continua.

A questo riguardo gli apparati di telecomunicazione possono essere raggruppati in due categorie:

- apparati di tipo tradizionale (tipicamente gli autocommutatori elettromeccanici, ancora in uso,

pur se destinati nel tempo a scomparire);

- apparati di tipo elettronico (tipicamente gli autocommutatori numerici e gli apparati trasmissivi).

Per i primi è richiesto un solo valore nominale di tensione continua (12V, 24V, 48V o 60V a seconda della tecnica costruttiva) con limiti statici e residuo c.a. sovrapposto fissati dalle norme CEI (v. tab. 1); questi limiti di tensione in regime transitorio possono essere anche sensibilmente superati.

Per i secondi sono richiesti più valori nominali di tensione continua (+5V, -5V, +12V, -12V, -48V, ecc.) con i limiti statici e residuo c.a. stabiliti per ciascuno di essi; questa seconda categoria di apparati, inoltre, è molto sensibile alle variazioni transitorie di tensione, soprattutto come velocità di variazione e come integrale tensione-tempo.

(*) Ing. Giuseppe Calzetti, ing. Maurizio Grossoni -Telecom Italia DG- Roma

Tensione normale	Valori limite ammessi	
	Massimo	Minimo
12V	14V	11V
24V	26V	23V
48V	52V	46V
60V	64V	58V

Tensione psfometrica generata dagli impianti di energia:

- gli impianti di energia destinati all'alimentazione microfonica degli apparecchi e alla trasmissione di comandi sui fili di conversazione non devono presentare sui conduttori principali di alimentazione dell'auto-commutatore, in corrispondenza del quadro principale di distribuzione, una **tensione psfometrica superiore a 3mV** in condizioni normali di esercizio;
- è ammessa una tensione psfometrica fino a 6mV in condizioni eccezionali (ad esempio durante il funzionamento di generatori di riserva o in fase di ricarica delle batterie dopo una interruzione dell'erogazione di energia dalla rete).

Tabella 1 Valori normali e valori limite ammessi per la tensione continua fornita dalla stazione d'energia

In taluni casi per gli apparati di tipo elettronico (ad es. per i sistemi di gestione TLC) è, altresì, richiesta un'alimentazione in corrente alternata con un valore nominale di tensione di 230V - 400V a 50 Hz.

Appare evidente, da quanto precede, come sia necessario interporre fra la fornitura di energia (da rete c.a. o, eventualmente, da altre fonti) e gli apparati da alimentare un insieme di macchine, opportunamente interconnesse, il quale sia in grado di adattare l'energia di rete c.a., o di altre fonti, alle esigenze funzionali degli apparati, assicurando nel contempo una continuità di fornitura a questi ultimi tale da raggiungere l'obiettivo affidabilistico prefissato.

Questo insieme di macchine, chiamato impianto di alimentazione o anche stazione di energia, è normalmente composto da:

- una parte destinata a convertire la corrente alternata proveniente dalla rete pubblica di distribuzione, in corrente continua e/o in corrente alternata di caratteristiche idonee all'alimentazione degli apparati (ad adattarla a queste caratteristiche se si tratta di fonti alternative);
- una parte destinata ad assicurare la continuità di alimentazione durante i periodi in cui la fornitura di energia da rete c.a. (o da altre fonti) si interrompe.

2. Caratteristiche dell'impianto di alimentazione

Analizzando attentamente le esigenze di alimentazione degli apparati di TLC si evidenzia che, mentre per quelli di tipo elettromeccanico occorre fornire energia con un unico valore nominale di tensione, e quindi con un unico impianto di alimentazione, per gli apparati di tipo elettronico, stante i diversi valori di tensione da essi richiesti, esistono due possibili soluzioni:

- realizzare tanti impianti di alimentazione indipendenti quante sono le tensioni da fornire;
- realizzare un solo impianto di alimentazione, adattando l'unica tensione da questo fornita alle esigenze delle varie parti degli apparati alimentati tramite convertitori continua-continua decentrati a livello apparati.

Tenuto conto delle complicazioni impiantistiche e degli elevati oneri economici della prima soluzione nonché della disponibilità di convertitori c.c./c.c., di potenza anche molto piccola, caratterizzati da rendimenti di conversione e affidabilità elevati, la soluzione normalmente adottata è quella che prevede un unico impianto di alimentazione predisposto, generalmente, per una tensione nominale in c.c. di 48V.

La soluzione utilizzata per convertire la corrente alternata proveniente dalla rete di distribuzione pubblica in corrente continua con caratteristiche idonee all'alimentazione degli apparati di TLC, è quella che prevede l'uso di uno o più convertitori statici c.a./c.c. (raddrizzatori) funzionanti fra loro in parallelo.

Qualora il carico alimentato richieda anche un'alimentazione in corrente alternata, sono normalmente previsti uno o più convertitori c.c./c.a. (invertitori) funzionanti fra loro in parallelo e idonei a essere alimentati dai predetti raddrizzatori; in taluni casi sono utilizzati appositi sistemi c.a./c.c. e c.c./c.a. con tensione nominale in c.c. di valore diverso da quello utilizzato dagli apparati di TLC (UPS).

La soluzione adottata per assicurare la continuità di alimentazione è quella che prevede l'uso di una o più batterie di accumulatori elettrochimici in quanto è l'unica in grado di consentire, fra l'altro in modo relativamente economico, la realizzazione di impianti di alimentazione ad affidabilità molto elevata.

Questi accumulatori infatti sono macchine che (salvo casi eccezionali) non sono soggette a guasti tali da interrompere istantaneamente l'erogazione di energia.

Di regola esse manifestano gradualmente nel tempo un eventuale peggioramento delle loro caratteristiche funzionali per cui è possibile, con una corretta gestione dell'impianto di alimentazione, correre ai ripari in tempo utile. Qualsiasi altra soluzione (gruppi di continuità a volano, ecc.) per quanto ben realizzata, non consente di ottenere livelli di affidabilità altrettanto elevati.

Gli accumulatori normalmente usati negli impianti in questione sono quelli al piombo-acido, sia per ragioni

economiche, sia, soprattutto, per il minor scarto di tensione (nei confronti di altri tipi) che si manifesta ai loro morsetti nelle varie fasi di funzionamento.

La loro tensione, pari a circa 2V nominali a circuito aperto, varia infatti come segue:

- da 2V a $1,8V \div 1,75V$ durante il funzionamento in scarica ai regimi cui sono normalmente sottoposti negli impianti in questione;
- da 2V a $2,30V \div 2,40V$ durante il funzionamento in ricarica, così come viene eseguita normalmente per un rapido ripristino delle capacità, senza spreco di energia e senza pregiudicare l'integrità;
- costante a $2,22V \pm 1\%$ o a $2,25V \div 2,27V \pm 1\%$ (a seconda del tipo impiegato), durante il funzionamento in carica di conservazione.

Per l'impianto di alimentazione gli accumulatori, oltre a rappresentare la parte indispensabile a ottenere la continuità, costituiscono anche una riserva di energia.

La loro idoneità ad accumulare anche grandi quantità di energia in volumi relativamente modesti, infatti, è tale che in molti impianti essi sono l'unica sorgente prevista per continuare l'alimentazione degli apparati durante le interruzioni dell'energia primaria oppure durante i guasti dei convertitori c.a./c.c..

In questi casi il loro dimensionamento deve essere tale da assicurare il livello affidabilistico previsto per gli impianti stessi. Tuttavia, frequentemente, in relazione anche all'importanza che riveste per la continuità del servizio telefonico il complesso di apparati alimentati e, soprattutto negli impianti di tipo elettronico per i quali è indispensabile il funzionamento dell'impianto di climatizzazione, la riserva di energia costituita dagli accumulatori viene integrata mediante l'uso di gruppi elettrogeni (2a riserva di energia), i quali consentono di prolungare l'alimentazione degli apparati anche per le interruzioni dell'energia primaria di durata notevole.

L'uso degli accumulatori, ovviamente, impone all'impianto di alimentazione, e più precisamente alla parte destinata a convertire la c.a. in c.c., di provvedere a ripristinare la capacità degli accumulatori stessi dopo un loro funzionamento in scarica, e a mantenere successivamente questi pienamente carichi.

Di fronte alle due esigenze, fra loro contrastanti, di ripristinare il più rapidamente possibile la piena capacità degli accumulatori dopo una loro scarica e di non danneggiare la loro integrità, fra i vari metodi possibili di ricarica, come è noto, è normalmente adottato quello I-U, il quale:

- in una prima fase effettua la carica a corrente autolimitata e tensione crescente;
- in una seconda fase prosegue la carica a tensione costante (generalmente compresa fra 2,3V e 2,4V) e corrente decrescente;
- successivamente mantiene la piena capacità con una tensione autoregolata di $2,22V \div 2,27V \pm 1\%$ (in funzione del tipo di accumulatore utilizzato).

3. Caratteristiche degli apparati costituenti gli impianti di alimentazione

3.1 Accumulatori

I tipi di accumulatori disponibili sul mercato si contraddistinguono in funzione della coppia elettrochimica impiegata (piombo-acido, nichel-cadmio, ecc.) nonché in funzione del campo di applicazione (stazionario, trazione, avviamento, ecc.).

Nel caso degli impianti per telecomunicazioni il tipo di accumulatore più utilizzato è quello al piombo-acido di tipo stazionario.

Gli accumulatori stazionari al piombo si caratterizzano essenzialmente per l'esecuzione dell'elettrodo positivo costituito da un gruppo di piastre collegate in parallelo che sono fondamentalmente dei seguenti tipi:

- a) piastra a grande superficie -planté- realizzata in lastre di piombo puro a struttura lamellare; la materia attiva è ricavata per formazione degli strati superficiali delle piastre stesse;
- b) piastra corazzata -a tubetti- nella quale la materia attiva è contenuta in un insieme di tubetti, isolanti e permeabili all'elettrolito all'interno di ciascuno dei quali è disposto coassialmente un conduttore in lega di piombo;
- c) piastra a griglia impastata -a piastra piana- nella quale la materia attiva è pressata su di una griglia in lega di piombo.

L'elettrodo negativo è costituito in ogni caso da piastre a griglia impastata.

Le principali caratteristiche e i rapporti economici di tali tipi di accumulatori sono riassunti nella tab. 2.

È opportuno ricordare l'importanza della progettazione e della costruzione della piastra positiva da cui dipende in pratica la durata dell'accumulatore; uno dei punti critici è infatti la formulazione delle leghe per le griglie.

Il piombo, notoriamente metallo tenero, viene legato con altri componenti, in particolare antimonio, per conferire la necessaria resistenza meccanica e per consentire la movimentazione delle griglie durante le operazioni di fabbricazione delle piastre e il loro montaggio in batteria, nonché per impartire al metallo fuso una conveniente fluidità per le operazioni di stampaggio delle griglie in conchiglia, per gravità o sotto pressione.

La maggior parte delle griglie è costituita da leghe antimoniali ($1 \div 11\%$) anche se nella panoramica mondiale se ne trovano un'ampia gamma (ad es.: piombo-calcio o leghe senza antimonio, tipo Astatin).

Purtroppo l'antimonio ha anche degli effetti negativi.

Infatti il metallo durante la ricarica della batteria viene rilasciato dalla griglia positiva; per effetto dell'attacco corrosivo anodico migra, attraverso il separatore, alla piastra negativa dove viene depositato.

Questo cosiddetto avvelenamento da antimonio della

Tipo di accumul.	Capacità elemento [Ah]	Densità elettrolito [Kg/dmc]	Tensione di carica [V]	Tensione di mantenim. [V]	Corrente di mantenim. [mA/Ah]	Resist. interna (valore medio) [Mohm/Ah]	Auto - scarica [%]	Rendim. in Ah %	Rendim. in Wh %	Energia specif. [Wh/dmc]	Durata [anni]	Prezzo
Al piombo	(Caratteristiche riferite alla scarica in 10 ore e alla temperatura di 25 °C)											
Platté (GROE)	25 ÷ 3000	1.21	2.20 ÷ 2.40	2.20 ÷ 2.25	0.15 ÷ 0.25	100	3 ÷ 8	84	70	15 ÷ 45	25	200
Tubolari (DIN)	25 ÷ 12000	1.21 ÷ 1.25	2.20 ÷ 2.40	2.20 ÷ 2.25	0.05 ÷ 6	300	0.5 ÷ 5	84	70	19 ÷ 65	15	100
A piastre impastate	12.5 ÷ 6000	1.21 ÷ 1.25	2.20 ÷ 2.40	2.20 ÷ 2.25	0.15 ÷ 6	120	2 ÷ 5	84	70	30 ÷ 76	7 ÷ 12	80

Tabella 2 Confronto tra le tipologie di accumulatori ad elettrolito libero

piastra negativa comporta due distinti effetti: un incremento del consumo di acqua durante le reazioni di ricarica dovuto all'abbassamento delle sovratensioni di scarica dell'idrogeno alla piastra negativa; la creazione di pile locali dove l'ossidazione del piombo è facilitata dall'effetto depolarizzante dell'antimonio sulla scarica dell'idrogeno.

L'assenza di antimonio nelle griglie, viceversa, provoca nella piastra positiva tre distinti effetti:

- minore capacità ovvero minore utilizzazione della materia attiva positiva;
- minore vita ciclica in carica e scarica complessiva, quindi minor numero di Wh estraibili dalla batteria durante la sua vita utile;
- possibilità di decadimenti prematuri della piastra positiva.

L'esperienza maturata nell'ultimo decennio evidenzia che un approccio "elettrochimico" per le batterie stazionarie senza manutenzione può essere il seguente:

- piastre positive tubolari con griglie intorno al 2% di antimonio;
- piastre negative con griglie in leghe prive di antimonio.

Come si è visto negli accumulatori convenzionali al piombo ad acido libero, le reazioni collaterali nelle fasi di carica e/o di ricarica portano a:

- liberazione di gas, con possibilità di avere nell'ambiente miscele esplosive;
- necessità di rabboccare gli elementi al fine di ripristinare l'acqua perduta.

Queste due situazioni che comportano oneri di manutenzione e d'investimento, soprattutto per la necessità di avere locali separati per le batterie, hanno favorito il nascere e il perfezionarsi degli accumulatori al piombo regolati con valvole (comunemente chiamati ermetici).

La tecnologia attualmente offre due soluzioni:

- accumulatori a elettrolito gelatinoso;
- accumulatori a elettrolito assorbito.

Nel caso di accumulatori gelificati, l'elettrolito viene immobilizzato mediante l'aggiunta di silice fino a ottenere del gel tixotropico, caratterizzato almeno

teoricamente da una configurazione spugnosa che permette la diffusione dell'ossigeno.

È da sottolineare, in quanto costituisce un aspetto critico, che questo gel può liquefare in determinate condizioni estreme (ad es.: pressione, temperatura, concentrazione acido, ecc.) e ridiventare gel quando cessano tali condizioni.

Nel caso del tipo assorbito, l'elettrolito è trattenuto da un separatore ad alto potere assorbente in genere di microfibre di vetro; in questo caso sono garantiti una porosità uniforme per la maggiore diffusione dell'ossigeno, un intimo contatto con le piastre per la migliore efficienza del processo elettrochimico nonché per la sua stabilità nel tempo; inoltre tale separatore è caratterizzato da una bassa resistenza interna che conferisce agli accumulatori costruiti con questa tecnologia un'ottima attitudine alle scariche rapide.

Un confronto fra i vari tipi è riportato nella tab. 3.

Gli accumulatori ermetici sono a tutti gli effetti accumulatori al piombo e pertanto, a parte la loro possibilità specifica della ricombinazione dei gas, presentano caratteristiche elettriche e comportamenti funzionali uguali a quelli noti di tutti gli accumulatori al piombo acido.

È utile riassumere di seguito alcune caratteristiche degli accumulatori ermetici, in particolare di quelli a elettrolito assorbito, che li differenziano dagli accumulatori tradizionali ad acido libero.

Queste caratteristiche si possono così sintetizzare:

- assenza di esalazioni acide;
- bassissime emissioni di idrogeno;
- nessuna necessità di rabbocchi periodici con acqua;
- curva di scarica più alta, soprattutto alle scariche rapide;
- minor ingombro complessivo, sia per i minori spazi di rispetto necessari alla manutenzione, sia per la maggior energia specifica, soprattutto alle scariche rapide;
- possibilità di una loro sistemazione in qualunque posizione;
- fornitura allo stato carico pronto;

GEL	AGM
Elettrolito: Gel di acido solforico e silice ($d > 1,280 \text{ kg/l}$) - nessuna stratificazione Separatore: in resina fenolica - resistenza leggermente superiore - handling critico Grande volume di elettrolito - buona resistenza alla sovraccarica - grande capacità termica Possibile impiego di piastre tubolari Vita attesa > 15 anni Tenuta poli: for life Esperienza nella tecnologia > 40 anni	Elettrolito: acido solforico ($d > 1,280 \text{ kg/l}$) - possibile stratificazione Separatore: in fibra di vetro - bassa resistenza - buon comportamento con alte correnti - handling critico Piccolo volume di elettrolito - moderata resistenza alla sovraccarica - bassa capacità termica con possibilità di valanga termica Piastre tubolari non adatte Vita attesa dichiarata 10 anni Tenuta poli: critica Esperienza nella tecnologia: circa 15 anni su capacità inferiori

Tabella 3 Differenze tra le caratteristiche dei sistemi GEL ad alta capacità e Assorbito (AGM) corrispondenti

- carica di mantenimento a tensione più alta (2,27V contro 2,22V);
- maggiore vulnerabilità a cariche non controllate;
- impossibilità di verificare il loro stato di carica tramite la lettura della densità dell'elettrolito.

Si può intuire quale impatto può avere sui sistemi di alimentazione, che utilizzano gli accumulatori come continuità di alimentazione e come riserva di energia, l'uso di accumulatori che presentano le sopracitate caratteristiche.

Senza addentrarsi nei particolari, si sottolinea come le implicazioni che il sistemista dovrà aver presenti sono numerose, e riguardano aspetti funzionali, gestionali, impiantistici, di sicurezza nonché legali.

3.2 Raddrizzatori (convertitori c.a./c.c.)

I raddrizzatori per impianti di alimentazione per TLC hanno subito nel corso degli ultimi anni una notevole evoluzione che ha consentito di migliorare le loro prestazioni in genere, e più precisamente, il rendimento di conversione, il peso, gli ingombri e la vita utile.

Tale evoluzione è stata determinata dalla disponibilità di nuovi componenti che hanno permesso di sviluppare differenti tecniche di conversione.

Infatti dopo i raddrizzatori con diodi al selenio e al silicio, si sono diffusi nel corso degli anni sessanta i raddrizzatori con diodi controllati; il loro principio di funzionamento è assai semplice e consiste nell'accensione dei diodi controllati con un ritardo, rispetto all'inizio dell'onda sinusoidale, variabile e controllato in funzione della tensione d'uscita (regolazione di tensione) e della corrente d'uscita (limitazione di corrente).

Lo spegnimento dei diodi controllati è naturale in quanto la corrente si riduce spontaneamente a zero nell'ambito di ogni semiperiodo.

All'inizio degli anni ottanta si sono cominciati a diffondere i raddrizzatori a conversione diretta da rete: essi sono l'ultima generazione in ordine di tempo e utilizzano conversioni intermedie ad alta frequenza (PWM).

Poiché le dimensioni dei trasformatori e dei filtri sono inversamente proporzionali alla frequenza, questa tecnica consente una sostanziale diminuzione di dimensioni (da 4 a 1) e di pesi.

Altri vantaggi sono inoltre presenti: assenza di rumore acustico, eccellente comportamento dinamico, basso rumore in uscita (anche psfometrico), facilità d'installazione e sostituzione. Essi subito negli ultimi dieci anni continui miglioramenti dovuti ai progressi nella teoria dei circuiti di conversione, nelle tecniche di pilotaggio e controllo, nelle tecnologie dei componenti.

Le frequenze di conversione sono salite dai 20-40kHz dei transistori bipolari, ai 100-200kHz dei MOSFET, per arrivare oltre il MHz con tecniche di risonanza.

Le potenze sono salite dai 100W a oltre 10kW, ma sono tuttora limitate al campo medio-basso.

La tendenza è di un sempre maggior impiego nel campo delle piccole e medie potenze con tecniche a scambio di corrente sinusoidale con la rete, a fattore di potenza unitario, e impiego di microprocessore per le regolazioni e i controlli.

3.3 Invertitori (convertitori c.c./c.a.)

La tecnica della conversione da tensioni continue a tensioni alterante risale a molti anni prima della seconda guerra mondiale.

Fino all'avvento dei diodi controllati negli anni '60, solo pochi invertitori commerciali furono sviluppati, date le limitazioni imposte dalle valvole e dai mutatori, e la tecnologia degli invertitori rimase quasi dormiente fino ad allora.

A partire dalla seconda metà degli anni '60 gli invertitori soppiantarono i gruppi di continuità rotanti negli impianti di alimentazione in c.a. per telecomunicazioni.

A seconda delle tecniche impiegate per la regolazione della tensione di uscita e della tecnologia, si sono avute successive generazioni di invertitori.

L'ultima generazione, resa possibile con l'adozione di transistori di commutazione veloci, di alta tensione e corrente, è quella che utilizza la regolazione PWM a frequenza elevata con ricostruzione dell'onda sinusoidale (5° generazione).

Tale soluzione prevede un numero elevato di commutazioni nel semiperiodo, con frequenze di commutazione nell'ordine dei kHz per grosse potenze fino a frequenza ultrasonica di 20kHz per potenze medio-piccole.

La legge di modulazione della commutazione del ponte di conversione è sinusoidale; con ciò il contenuto armonico all'uscita è a frequenza di conversione e quindi facilmente filtrabile con semplici filtri L-C.

Sono poi normalmente applicate varie tecniche di controllo quale ad esempio: la regolazione della sinusoide di uscita punto a punto, con ripresa delle distorsioni introdotte dalla conversione e da carichi distorcenti; ciò consente anche un ottimo comportamento dinamico.

I vantaggi di invertitori realizzati con tale tecnica sono principalmente:

- ottimo comportamento dinamico;
- bassa distorsione all'uscita anche su carichi distorcenti;
- minori dimensioni e pesi.

La tecnica di conversione con regolazione PWM, sopra descritta, appare sufficientemente ottimizzata e destinata a mantenersi anche nei prossimi anni; miglioramenti si avranno nell'aumento delle frequenze di conversione e nei tagli di potenza.

Una sostanziale novità, non nell'architettura della conversione di potenza, ma nei circuiti di comando, controllo e segnalazione si avrà con l'impiego di microprocessori inseriti con funzioni operative.

I vantaggi attesi sono in una diagnostica (locale e a distanza) più spinta, con incremento della disponibilità dei sistemi e una miglior collaudabilità e gestione in opera.

Circa il commutatore statico si ricorda che esso è l'elemento che connette i carichi a una rete di soccorso nel caso in cui la tensione fornita dagli invertitori esca dai limiti assegnati per effetto di guasti o di sovraccarichi agli invertitori stessi.

La parte di potenza è semplice, essendo costituita da una coppia di diodi controllati (per fase) in antiparallelo.

Complessi risultano invece i criteri di gestione e i relativi circuiti che li realizzano.

Essendo il commutatore statico un elemento di riserva deve possedere due requisiti in contrasto tra loro:

- elevata velocità d'intervento;
- immunità a interventi intempestivi.

Inoltre deve risultare molto affidabile (anche nei riguardi di eventuali perturbazioni provenienti dalla rete) per non creare "strozzature" affidabilistiche al sistema, che ne vanifichino l'utilità.

Deve inoltre possedere i necessari requisiti, richiesti espressamente dalle norme, per potersi interconnettere alla rete pubblica; uno essenziale è quello (di tipo antinfortunistico) contro i ritorni di tensione verso rete per guasto al commutatore statico stesso (cortocircuito o comando indebito agli SCR).

3.4 Gruppi elettrogeni

I gruppi elettrogeni disponibili sul mercato si contraddistinguono in funzione del motore primo (endotermico a ciclo otto, endotermico a ciclo diesel e a turbina a gas) e del generatore (dinamo, asincroni e sincroni).

Nelle applicazioni per impianti di telecomunicazioni sono normalmente utilizzati gruppi elettrogeni mossi da motori diesel, la cui potenza nominale può variare da qualche kW a qualche MW e la cui velocità di rotazione è compresa fra 3000g/1' e 1000g/1'.

Circa il generatore, il tipo più diffuso è il generatore sincrono autoeccitante e autoregolante; l'alternatore propriamente detto è costituito da un indotto fisso e un induttore rotante, mentre l'eccitazione in c.c. della ruota polare è fornita, attraverso un ponte di diodi, dall'indotto rotante di un alternatore trifase d'eccitazione avente l'induttore fisso.

Ciascun alternatore è equipaggiato con un dispositivo d'eccitazione statico che, oltre a fornire l'energia necessaria per l'eccitazione dell'eccitatrice, provvede a correlare la corrente di eccitazione con il carico in modo da mantenere costante la tensione dell'alternatore stesso.

Tali alternatori si caratterizzano per:

- elevata rapidità di risposta in quanto è il carico stesso che comanda il valore della corrente di eccitazione;
- elevata capacità di sovraccarico in quanto essendo la corrente di eccitazione legata alla corrente di carico dell'alternatore, quest'ultimo può sopportare notevoli sovraccarichi pur mantenendo la tensione a un valore prossimo a quello nominale;
- mancanza di contatti striscianti, in quanto privi sia di spazzole che di anelli, e quindi, conseguentemente, una ridotta manutenzione.

Il gruppo elettrogeno, per adempiere alla propria funzione di generatore autonomo di energia elettrica, deve poter essere connesso ai carichi che deve alimentare garantendo la massima sicurezza sia di esercizio che delle persone; è

perciò indispensabile la presenza di un quadro elettrico destinato a sovrintendere all'impianto nel suo complesso.

Dai criteri di inserzione dei G.E. negli impianti deriva la minore o maggiore complessità delle funzioni di comando e controllo affidate al quadro elettrico.

Da un punto di vista generale e analizzando solo il caso, che è anche il più comune, di G.E. in emergenza alla rete ENEL (funzionamento in isola), si può dire che il quadro elettrico ha lo scopo di comandare l'avviamento e l'arresto del gruppo elettrogeno, in relazione alle condizioni della rete esterna di distribuzione, per garantire l'alimentazione di riserva alle utenze; ha inoltre lo scopo di controllare i parametri del G.E. con funzione di protezione e arresto macchina al verificarsi di eventuali anomalie durante il suo funzionamento.

La logica di funzionamento può essere realizzata:

- elettromeccanicamente, con relè di tipo classico;
- elettronicamente, con schede a circuiti integrati;
- con microprocessore, dotati di unità programmabili.

Gli impianti dotati di logica di controllo a microprocessore si vanno sempre più diffondendo, soprattutto presso i maggiori Enti, in quanto integrabili in reti di supervisione generale, che aumentano considerevolmente l'affidabilità complessiva dell'impianto installato.

Gli schemi applicati e il tipo di componenti utilizzati variano moltissimo in funzione del tipo di impianto e delle caratteristiche e/o dell'abilità dei vari costruttori.

Il gruppo elettrogeno, infine, è equipaggiato, in funzione delle specifiche esigenze di impianto o delle necessità proprie del motore, con tutta una serie di dispositivi di controllo e di apparati ausiliari.

Fra questi vanno ricordati i dispositivi di sicurezza

omologati quali:

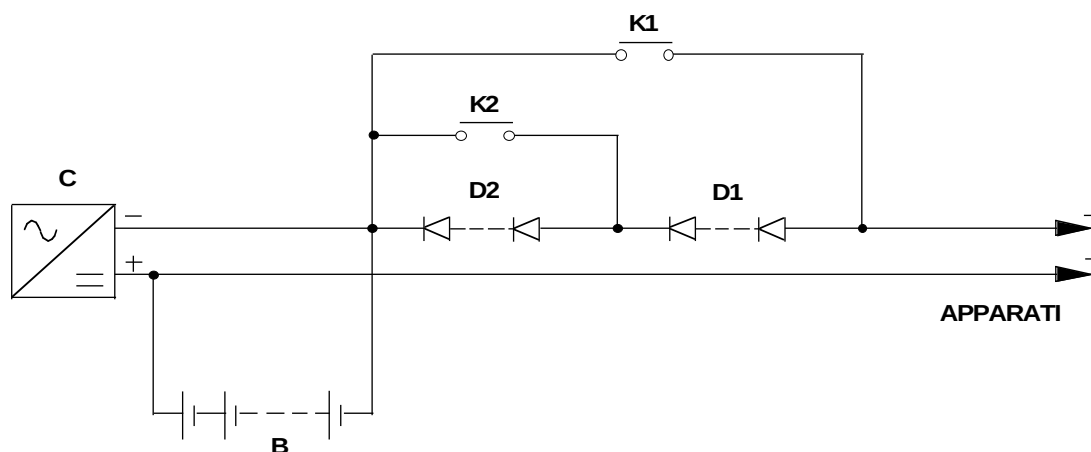
- dispositivo automatico di arresto motore per alta temperatura dell'acqua di raffreddamento (nei motori raffreddati ad aria può essere sostituita dalla temperatura delle teste cilindri o dalla temperatura dell'olio di lubrificazione);
 - dispositivo automatico di arresto motore per bassa pressione dell'olio di lubrificazione;
 - dispositivo automatico di intercettazione (elettrovalvole) del flusso del combustibile in caso di arresto del motore diesel o in mancanza della tensione di controllo degli automatismi;
- che per disposizione ministeriale (circolare Ministero dell'interno n° 31) devono equipaggiare obbligatoriamente il motore primo.

4. Stato dell'arte e prospettive future per gli impianti di alimentazione in ambito Telecom Italia

4.1 Stato dell'arte

Fino all'avvento degli apparati di tipo elettronico, gli impianti di alimentazione in c.c. normalmente usati in Telecom Italia sono stati quelli strutturati come indicato nelle figg. 1, 2, 3.

Nel corso degli anni questi impianti, pur mantenendo inalterata la loro struttura di principio, via via che il progresso tecnologico lo consentiva, hanno subito modifiche tese a migliorarne le prestazioni funzionali e a ridurre gli oneri d'impianto e di gestione.



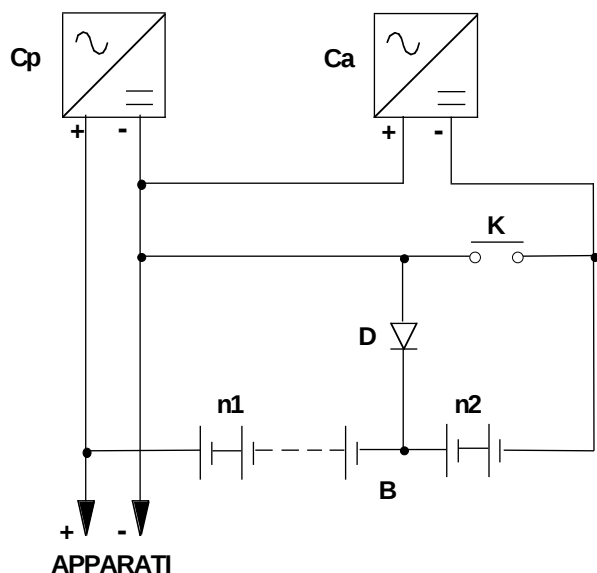
C: convertitore c.a./c.c.; effettua il mantenimento e la ricarica della batteria di accumulatori e, attraverso le celle di caduta D1 o D1+D2, alimenta gli apparati.

B: batteria di accumulatori con un numero di elementi pari a $V_n/2$.

D1-D2: serie di celle di caduta.

K1-K2: contattori; la loro azione, opportunamente comandata, consente di mantenere la tensione verso gli apparati entro i limiti consentiti nelle varie fasi di funzionamento.

Figura 1 Impianto di alimentazione in c.c. per apparati elettromeccanici



Cp: convertitore c.a./c.c.; alimenta gli apparati e, con l'ausilio di Ca, effettua il mantenimento e la carica rapida della batteria.

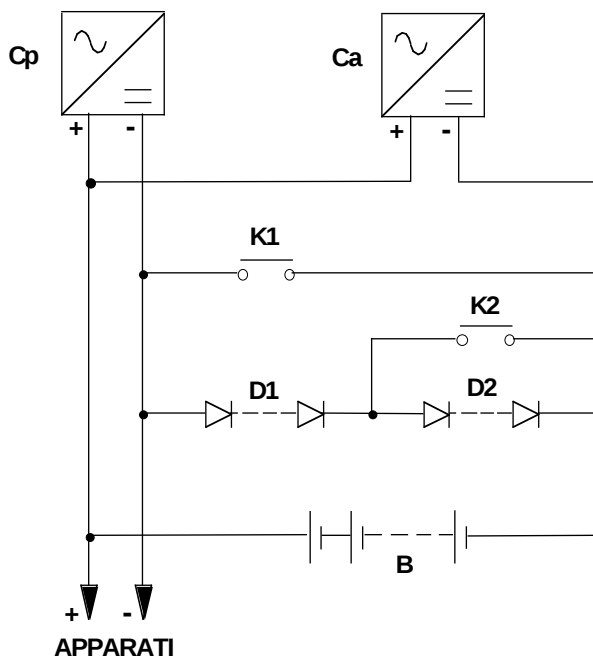
Ca: convertitore c.a./c.c.; sommando la sua tensione (opportunamente regolata) a quella costante Cp, consente di effettuare il mantenimento e la carica rapida della batteria.

B: batteria di accumulatori divisa in due parti, con un numero totale di elementi $n1+n2$ pari a $Vn/2$.

D: diodo; assicura la continuità al mancare della rete c.a..

K: contattore; si chiude quando manca la rete c.a..

Figura 2 Impianto di alimentazione in c.c. per apparati elettromeccanici



Cp: convertitore c.a./c.c.; alimenta gli apparati.

Ca: convertitore c.a./c.c.; effettua il mantenimento e la carica rapida della batteria.

B: batteria con un numero di elementi pari a $Vn/2$.

D1-D2: serie di celle di caduta; assicurano la continuità al mancare della rete c.a..

K1-K2: contattori; la loro azione consente di mantenere la tensione verso gli apparati entro limiti accettabili al mancare della rete c.a..

Figura 3 Impianto di alimentazione in c.c. per apparati elettromeccanici

Con l'avvento degli apparati elettronici, considerate:

- la più elevata sensibilità di questi alle variazioni transitorie di tensione (a fronte ripido soprattutto);
 - la prospettiva di dover alimentare per molti anni ancora i due tipi di apparati elettromeccanico ed elettronico (il primo destinato a scomparire sostituito gradualmente dal secondo);
 - la necessità, durante questo periodo transitorio, di fornire contemporaneamente, in molti casi, l'energia ai due tipi di apparati coesistenti in un'unica sede;
- Telecom Italia ha messo a punto, a fine anni '60, un impianto di alimentazione la cui struttura è rappresentata in fig. 4.

La scelta fatta, dove ciascuna uscita è in grado di fornire la piena potenza dell'insieme convertitori c.a./c.c.-batterie, consente la più completa elasticità di utilizzazione dell'impianto.

È possibile cioè fornire contemporaneamente energia dalle due uscite nelle proporzioni richieste dai due tipi di apparati presenti o, al limite, fornire tutta la potenza disponibile da una soltanto di esse, se è presente un solo tipo di apparati.

Inoltre elimina in ambedue le uscite i transistori a fronte ripido di tensione che nelle strutture precedenti si manifestavano a ogni interruzione di rete c.a. (per l'uscita completa di convertitori c.c./c.c. centralizzati il transitorio a fronte ripido si manifesta soltanto qualora si escludano i convertitori c.c./c.c. per guasto).

Infine consente il completo utilizzo della capacità delle batterie.

Un'altra soluzione adottata è quella di fig. 5 la quale rispecchia la soluzione di fig. 3 con le seguenti varianti:

- sono previste due uscite per i due tipi di apparati;
- i convertitori chiamati Ca in fig. 5 provvedono anche all'alimentazione degli apparati elettronici;
- al mancare della rete c.a. le due serie di celle di caduta vengono cortocircuitate immediatamente da un diodo controllato e successivamente dai contattori K1 e K2; in tal modo si riduce notevolmente la durata e l'entità dell'abbassamento di tensione dovuto alla c.d.t. delle celle di caduta.

Con questa struttura non vengono eliminati i transistori a fronte ripido per l'uscita verso apparati elettromeccanici, pur riducendoli tramite il diodo controllato suddetto, e

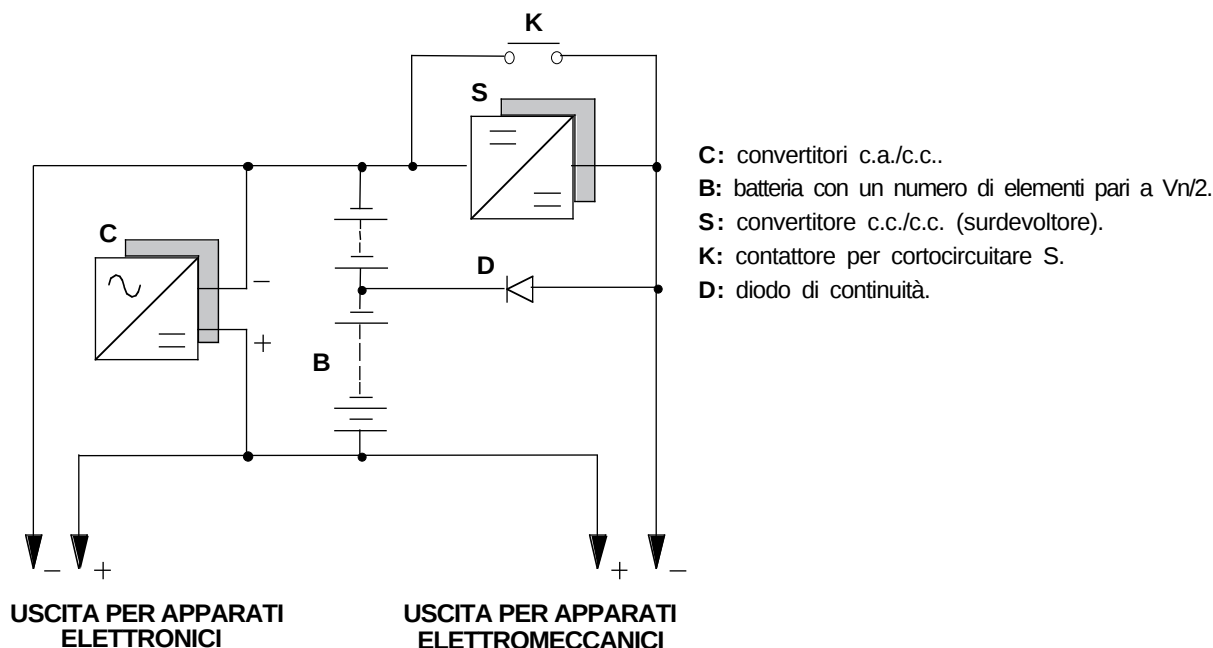
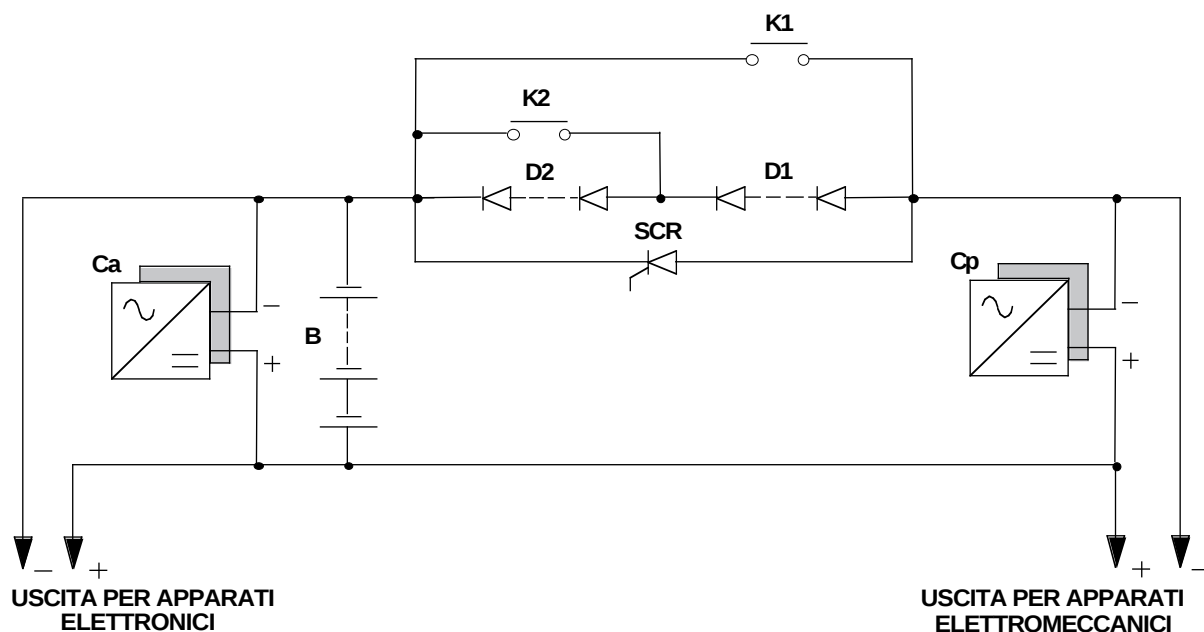


Figura 4 Impianto di alimentazione in c.c. per apparati elettronici ed elettromeccanici



Cp: convertitori c.a./c.c.; alimentano gli apparati elettromeccanici.

Ca: convertitori c.a./c.c.; alimentano gli apparati elettronici ed effettuano il mantenimento e la carica rapida della batteria.

B: batteria con un numero di elementi pari a $V_n/2$.

D1-D2: serie di celle di caduta; assicurano la continuità in uscita verso gli apparati elettromeccanici al mancare della rete c.a..

K1-K2: contattori; la loro azione consente di mantenere la tensione verso gli apparati elettromeccanici entro limiti accettabili.

SCR: diodo controllato; l'intervento di esso riduce la durata dell'abbassamento di tensione verso gli apparati elettromeccanici al mancare della rete c.a..

Figura 5 Impianto di alimentazione in c.c. per apparati elettronici ed elettromeccanici

non è possibile, sempre per questa uscita, utilizzare completamente la capacità delle batterie.

La situazione attuale in Telecom Italia è pertanto la seguente:

- le strutture di impianto (in c.c.) di figg. 1, 2, 3 non sono più in produzione, tuttavia sono ancora presenti in servizio;
- le strutture di impianto (in c.c.) di figg. 4 e 5 verranno utilizzate fintantoché rimarranno in servizio apparati elettromeccanici;
- gli impianti di alimentazione in c.c. di nuova fornitura, completati con uno o più gruppi elettrogeni e con uno o più invertitori, in funzione della natura e dell'importanza del carico alimentato nonché in presenza di apparati che richiedano un'alimentazione in c.a. senza soluzione di continuità, sono visibili in fig. 6; in taluni casi (quando non sia possibile installare invertitori o la potenza richiesta sia elevata) sono previsti appositi gruppi di continuità statici in c.a. (UPS).

4.2 Prospettive future

Nel campo degli impianti di alimentazione per telecomunicazioni si possono prevedere le seguenti prospettive:

- la progressiva riduzione del numero di apparati di tipo tradizionale in servizio fa prevedere strutture dell'impianto di alimentazione in corrente continua soltanto orientate all'alimentazione di apparati elettronici.

Le strutture di figg. 4 e 5, consentono di seguire questo orientamento; infatti in esse la parte di impianto destinata a realizzare l'uscita verso apparati tradizionali è resa opzionale.

In prospettiva l'unica struttura utilizzata sarà quella che prevede la sola uscita verso apparati elettronici.

- la disponibilità di adatti sensori e dei microprocessori lascia prevedere un orientamento verso il controllo in tempo reale (locale e a distanza) delle varie condizioni di funzionamento dell'intero impianto di alimentazione e dei singoli apparati costituenti (valori di corrente e/o tensione, temperature, intervento protezioni, ecc.).

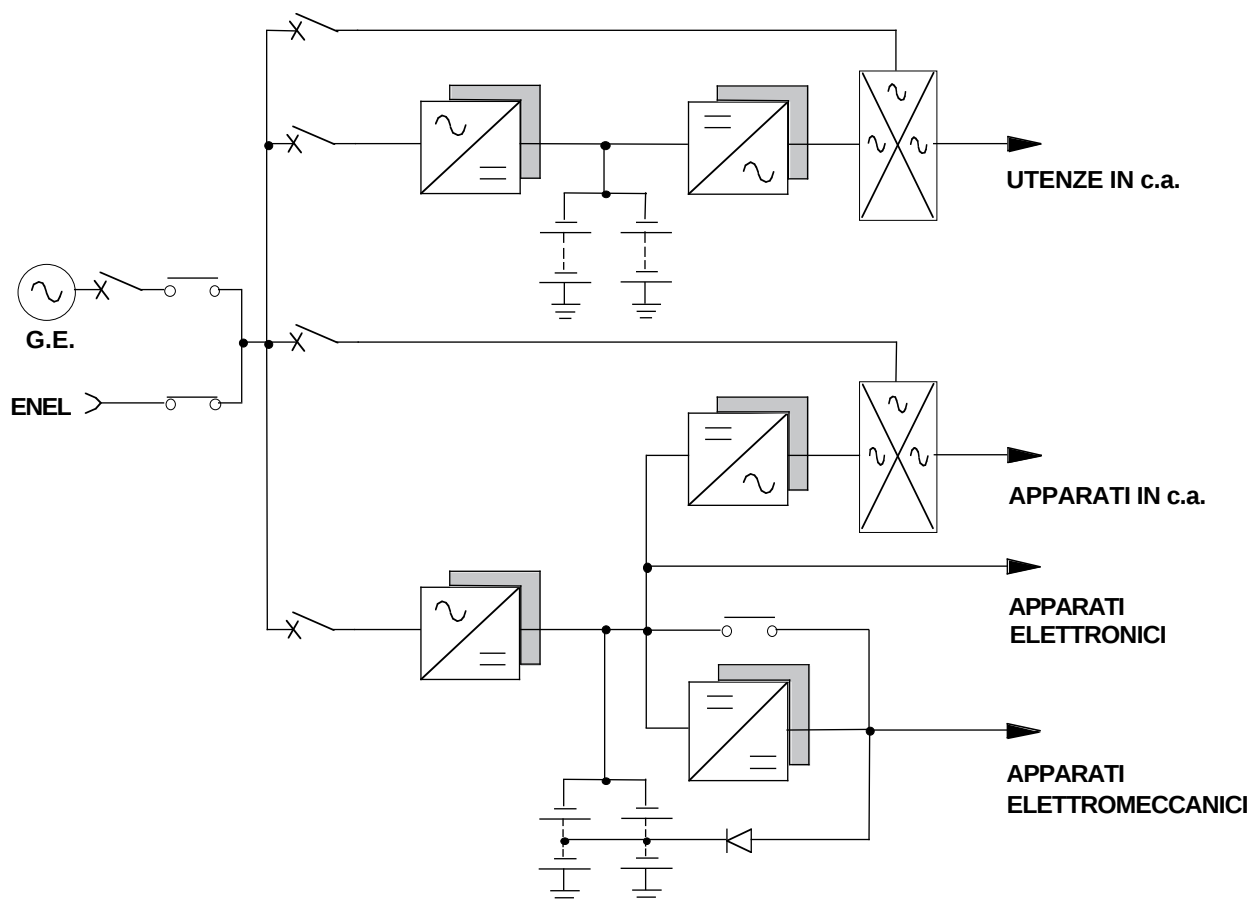


Figura 6 Impianto di alimentazione in c.c. e in c.a. per apparati di TLC e centri di elaborazione dati

È possibile infatti portare su adatto quadro di segnalazione locale e, tramite un'adeguata rete di microprocessori a più livelli, a uno o più centri di sorveglianza presidiati, le varie segnalazioni generate dai sensori predisposti nelle diverse parti dell'impianto di alimentazione.

Un'altra possibilità che si prospetta è quella di utilizzare la stessa rete di microprocessori per azioni di comando tese a effettuare automaticamente le riconfigurazioni dell'impianto di alimentazione che si rendessero necessarie al manifestarsi in esso di qualche anomalia o al miglior sfruttamento dell'impianto stesso (risparmio energetico).

Questa possibilità, ovviamente, dovrà essere subordinata a un'attenta valutazione degli aspetti affidabilistici che coinvolge.

- la disponibilità ormai affermata degli accumulatori al piombo-acido del tipo a ricombinazione, fra l'altro per capacità unitarie via via crescenti, lascia prevedere la possibilità di una realizzazione impiantistica degli impianti di alimentazione anche sensibilmente diversa da quella attuale.

Infatti, considerato che questi tipi di accumulatori non producono esalazioni acide e, anche in condizioni di carica rapida, producono quantità di idrogeno di gran lunga inferiori ai limiti di pericolosità, si può affermare che non è più necessario adibire per essi un locale apposito, locale che per gli accumulatori ad acido libero, come noto, deve essere opportunamente attrezzato (protezioni antiacido, impianti di ventilazione ed elettrico rispondenti a severe norme di sicurezza).

Gli accumulatori quindi possono essere installati nello stesso locale dei convertitori c.a./c.c. o di altre apparecchiature elettriche e/o di TLC, solo che il locale stesso non sia ermeticamente chiuso, cioè che in esso sia attivo un sia pur minimo ricambio d'aria. Si comprende allora che ubicando convertitori c.a./c.c. e batterie in uno stesso locale è possibile:

- ridurre la lunghezza dei collegamenti fra convertitori e batterie con conseguenze positive di costo e, funzionalmente, di comportamento ai transitori nel passaggio dalle condizioni di carica a quelle di scarica per le batterie stesse;
- evitare che le batterie si trovino a funzionare con temperatura ambiente anomala, con conseguente miglior sfruttamento della loro capacità nominale.

E' possibile, inoltre, trasformare l'unico tradizionale impianto di alimentazione centralizzato in più impianti ubicati, ciascuno, nelle vicinanze o nello stesso locale dei singoli apparati da alimentare.

Così facendo vengono eliminati i lunghi e costosi collegamenti in corrente continua fra l'insieme convertitori-batterie e gli apparati, sostituiti dai più economici collegamenti in c.a. fra l'insieme rete c.a.-gruppo elettrogeno e i singoli impianti di alimentazione.