

Notiziario Tecnico

Archivio

2/1993

 **TIM**

La propagazione nelle fibre ottiche: un approfondimento necessario

Le geniali intuizioni che hanno portato allo sviluppo dei sistemi di trasmissione in fibra ottica, e lo stesso aggettivo che caratterizza le fibre, si originano, in parte, dall'osservazione di fenomeni di propagazione guidata di radiazioni elettromagnetiche appartenenti alla banda visibile dello spettro. E, sebbene i sistemi trasmissivi concretamente utilizzati abbiano, sin dall'inizio, fatto uso di segnali al di fuori di tale banda, ed in particolare nell'infrarosso "vicino", la spiegazione dei meccanismi di propagazione ha fatto frequentemente riferimento all'ottica geometrica, sufficiente per delineare i meccanismi in questione.

L'ottica geometrica è però sostanzialmente inadeguata a descrivere il comportamento delle fibre monomodali. Con l'affermarsi di tali fibre, ed il prevalere del loro impiego in quasi tutte le applicazioni, si ritiene oggi opportuno tornare sui principi della propagazione guidata ricorrendo alla teoria dell'elettromagnetismo, ed agli strumenti matematici usuali nelle trattazioni concernenti questa teoria.

Per tale motivo, in un ciclo di articoli dedicato ai sistemi in fibra ottica, sono stati inseriti alcuni contributi di richiamo della teoria in argomento, utilizzando in particolare formule ed algoritmi adatti alla trattazione riguardante i mezzi dielettrici e la specifica geometria delle fibre ottiche.

Nel proporre questi contributi, si è consapevoli del fatto che i temi trattati, e le nozioni fornite, sono per lo più già presenti nel bagaglio conoscitivo dei lettori del Notiziario Tecnico SIP; e che, inoltre, avranno utilizzazione diretta piuttosto sporadica da parte di tecnici che operano nell'esercizio delle telecomunicazioni. Si ritiene tuttavia che, proprio nell'espletamento di attività di esercizio, un eventuale troppo vago ricordo dei fondamenti scientifici su cui si basano i sistemi in fibra ottica costituisca una pericolosa carenza, che può condurre, tra l'altro, ad un'interpretazione banale, se non erronea, dei problemi affrontati nella pratica quotidiana.

Sotto tale premessa, un richiamo di principi scientifici, nell'ambito di questo o di altri cicli di articoli, è da considerare in termini di facilitazione e di comodità per i lettori; resta inteso che la frequenza con cui, nell'economia del Notiziario, si proporranno testi di taglio scientifico su argomenti basilari, terrà conto delle indicazioni dei lettori, ai quali pertanto si richiedono reazioni e commenti.

p.r.

Fibre ottiche: motivi e tempi di un successo

C. G. Smeda (*)

Nella rete pubblica di telecomunicazioni, in tutti i Paesi evoluti, le fibre ottiche sono ormai talmente diffuse da doversi considerare come il prodotto di una tecnologia altrettanto consolidata quanto quelle dei mezzi che le hanno precedute nel tempo, che con esse convivono, ma che con esse hanno ben poco in comune, sia sotto il profilo tecnologico, sia sotto quello progettuale. Osservando ciò, anche chi non conosce le vicende storiche della trasmissione su fibra ottica ne ricava la sensazione che il passaggio dai portanti metallici a quelli in silice debba essere stato brusco e traumatico. E poiché in ambito ingegneristico i traumi hanno, sperabilmente, una loro logica, non sembra inutile, soprattutto pensando ai lettori più giovani e a coloro che non ebbero modo di seguire da vicino la fase iniziale della trasmissione ottica, ripercorrere velocemente le strane vicende che hanno caratterizzato l'avvio delle ricerche e i primi sviluppi nel settore delle fibre e dei dispositivi optoelettronici, e hanno scandito i tempi rapidissimi dei loro successi applicativi.

1. Lo scenario alla metà degli anni '60

Gli anni '50 erano stati, se misurati con i metri di allora, un periodo di formidabile espansione delle telecomunicazioni. La domanda di nuovi collegamenti interurbani, trainata, su scala mondiale, dalla crescita economica che proseguiva senza pause fino dall'inizio della ricostruzione postbellica, aveva trovato eccellenti risposte tecniche, in larga misura grazie alle ricadute in campo civile di studi e ricerche precedentemente effettuati a fini militari, primo fra tutti il radar. Ma già negli stessi anni, nei Paesi più lungimiranti, si era compreso che i sistemi trasmissivi a grande distanza tipici dell'epoca - cavi coassiali e ponti radio - non potevano né moltiplicarsi indefinitamente come numero, né continuare ad evolversi nella stessa maniera in cui le comunicazioni elettriche erano progredite dai tempi di Morse fino ad allora, cioè nella direzione di una crescita graduale e continua delle frequenze impiegate. Infatti, nei cavi coassiali l'"effetto pelle", che, come noto, provoca un aumento dell'attenuazione proporzionale alla radice quadrata della frequenza, comporta un calo delle lunghezze delle tratte di ripetizione al crescere del numero dei canali multiplati sullo stesso cavo. Sul

fronte dei ponti radio, le frequenze portanti erano ormai giunte alla soglia di bande nelle quali i gas costituenti l'atmosfera e, soprattutto, i fenomeni meteorologici causano attenuazioni forti e, ancor peggio, variabili aleatoriamente.

Sempre negli anni '50, fu individuata una soluzione teorica di questi problemi, degna di venire approfondita con una massiccia sperimentazione: l'impiego di guide d'onda a sezione circolare, nelle quali esistono modi di trasmissione che, almeno sulla carta, hanno un'attenuazione la quale, in barba all'effetto pelle, decresce sempre al crescere della frequenza. La traduzione di questa idea dalla carta alla pratica era pesantemente condizionata dal fatto che nessuno tra questi modi a bassa attenuazione può propagarsi *da solo* in una guida circolare; nonostante ciò, attraverso una lunga serie di brillanti idee e di tenaci realizzazioni, fu effettivamente possibile costruire guide d'onda ove l'attenuazione era sufficientemente bassa (dell'ordine di 2-3 dB/km) per presentare interesse applicativo nelle trasmissioni a grande distanza, su bande passanti dell'ordine di varie decine di GHz, alle frequenze delle onde millimetriche (cioè sopra i 30 GHz).

Questi successi tecnici furono però possibili solo a fronte di costi molto elevati, sia di produzione (le guide, spesso costituite da fili avvolti ad elica, avevano tolleranze meccaniche sub-millimetriche, su diametri di alcuni cm e su lunghezze delle singole pezzature di alcuni m), sia di installazione (i raggi di curvatura non

(*) Prof. Carlo Giacomo Smeda - Università degli Studi di Padova - Dipartimento di Elettronica ed Informatica.

potevano scendere sotto alcune centinaia di m). Fu dunque chiaro, fin dai primi anni '60, che questi mezzi trasmissivi avrebbero potuto venire utilizzati nella rete solo a condizione che la loro capacità trasmissiva venisse sfruttata appieno. Il loro limite, dunque, era nella scarsissima flessibilità d'impiego.

Nel frattempo, però, all'inizio degli anni '60 era nato il laser; si era così fatto un salto di oltre 3 ordini di grandezza nelle frequenze generabili con un oscillatore ad elevata coerenza temporale. L'attenzione per il nuovo venuto fu immediata anche nel mondo delle telecomunicazioni, come in molti altri settori della scienza e della tecnica. Dopo aver preso atto, nel giro di pochi anni, dei gravi problemi inevitabilmente connessi con lo sfruttamento della propagazione libera di fasci laser in atmosfera (attenuazioni elevate, mancanza di riservatezza, ecc.), si iniziò a studiare a fondo l'ipotesi della trasmissione guidata, non però nelle fibre, inizialmente, bensì in guide di grosso diametro (tipicamente, una decina di cm), costituite da sequenze periodiche di lenti, in vetro o gassose. Erano scelte che oggi possono apparire strane, suggerite in parte dalla precedente cultura "microondistica", ma forse ancor più dalla non disponibilità, a quell'epoca, di sorgenti laser che fossero di piccolo ingombro, di basso consumo, ed in grado di emettere un'onda continua a temperatura ambiente. Infatti, se già nel 1962 più laboratori avevano annunciato, quasi contemporaneamente, la nascita dei primi laser a semiconduttore, fino al 1970, per ragioni di cui si dirà più avanti, essi furono limitati a un funzionamento impulsato a basso *duty cycle* e alla temperatura dell'azoto liquido (pari a -195°C ovvero 77.4 K).

A somiglianza di quelle per onde millimetriche, le guide ottiche a lenti avevano pregi tecnici, ma elevatissimi costi di produzione e d'installazione, e quindi scarsissima flessibilità d'impiego. Ma non si vedevano all'orizzonte altre soluzioni in grado di competere con le due appena descritte; e i top managers delle telecomunicazioni si chiedevano, ansiosi e preoccupati, se, quando, come e dove iniziare a trasferire almeno una delle due dalla ricerca all'esercizio.

2. Le quattro grandi vittorie delle fibre ottiche

A parere di chi scrive, se le fibre ottiche hanno bruciato le tappe nel passare dai laboratori di ricerca alle applicazioni in campo, ciò si deve soprattutto ai risultati eccellenti che sono stati conseguiti su quattro fronti, tutti d'importanza decisiva, e soprattutto alla stupefacente rapidità con cui le battaglie su questi quattro fronti sono state vinte. Benché nella realtà esse si siano intrecciate e sovrapposte, con un po' di semplificazioni si può dare della loro sequenza logica e cronologica una visione schematica, che si cercherà ora di esporre.

2.1 L'attenuazione

La prima proposta di ricorrere a fibre di vetro per la trasmissione di luce laser, avanzata da Kao e Hockham nel 1966, fu considerata da taluni una spiritosa goliardata, da altri un'intollerabile eresia; ben pochi la presero per una cosa seria. Non è dato sapere se le prime ricerche sperimentali siano state avviate solo per non lasciare nulla di intentato, o per vera fiducia nell'idea; sta di fatto che, oltre ad alcuni fra i principali centri di ricerca sulle telecomunicazioni, anche l'americana Corning Glass Works, l'azienda vetraria più famosa del mondo, decise di indagare in proprio, nei suoi laboratori, sull'effettiva realizzabilità delle fantasiose proposte di Kao e Hockham.

La prima causa del diffuso scetticismo nei confronti delle fibre ottiche era, alla fine degli anni '60, l'attenuazione elevatissima delle fibre allora esistenti: i valori tipici si aggiravano sul migliaio di dB/km. Ma Kao e Hockham avevano affermato che queste perdite altissime erano causate da imperfezioni tecnologiche, prima fra tutte l'alta concentrazione di impurezze; essi avevano anche scritto che la trasmissione su fibra avrebbe presentato un interesse concreto se l'attenuazione fosse scesa sotto la soglia di 20 dB/km. Questa cifra appariva lontanissima dalla pratica realizzabilità e fu quindi commentata con sorrisi di sufficienza fino al primo di ottobre del 1970. Quello storico giorno, presso la sede londinese dell'Institution of Electrical Engineers, quasi al termine di una tavola rotonda dedicata al confronto fra guide per onde millimetriche e fibre ottiche, confronto il quale, fino a quel momento, pendeva nettamente a favore delle prime, con un magistrale *coup de théâtre*, Corning annunciò, per bocca del Dr. Maurer, di avere realizzato una fibra con attenuazione inferiore ai 20 dB/km.

Era il segno inequivocabile che la prima, fra le quattro grandi battaglie attraverso cui le fibre ottiche si sarebbero fatte largo, stava prendendo una piega favorevole; ma nessuno allora avrebbe potuto immaginare che, su questo fronte, i passi successivi sarebbero stati così rapidi e così trionfali come mostra la curva storica riportata in fig. 1. Essa indica che già prima del 1980 l'attenuazione era scesa sostanzialmente ai valori minimi teoricamente possibili. Anche grazie alla scelta di passare dalle lunghezze d'onda sfruttate nella fase iniziale (gli 850 nm della cosiddetta "prima finestra") ad altre (i 1300 nm della "seconda finestra" e i 1550 della "terza"), le perdite erano scese a valori 100 volte inferiori a quella soglia dei 20 dB/km che era stata posta da Kao e Hockham. Chi scrive non è a conoscenza di nessun altro esempio di tecnologia, nel campo delle telecomunicazioni, che nel suo primo decennio di vita abbia superato di un fattore 100 l'obiettivo che si era prefisso di conseguire; in questo dato storico si identifica la prima grande molla del

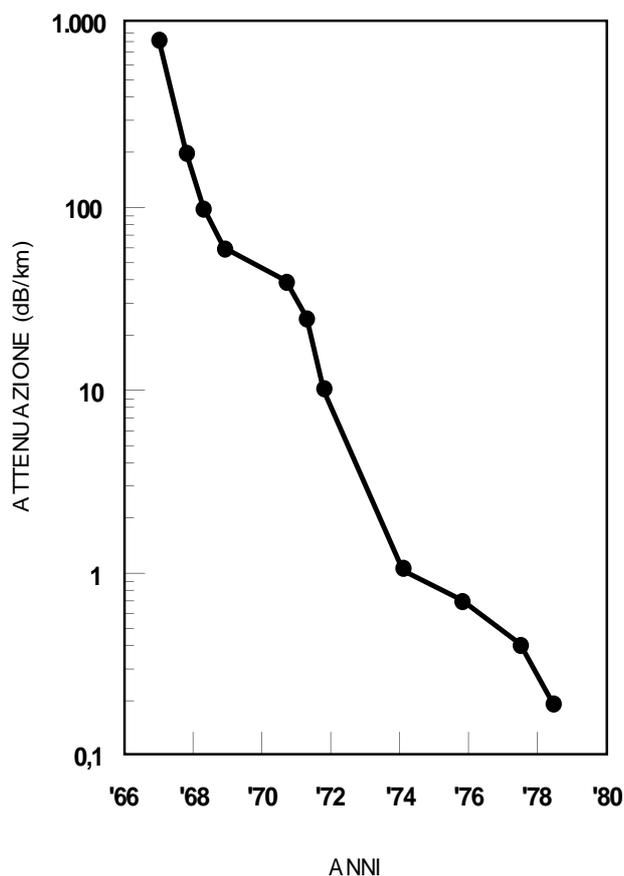


Figura 1

successo commerciale della trasmissione ottica, e soprattutto la sorgente delle grandi energie, e dell'instinguibile ottimismo, che furono necessari per affrontare e risolvere gli altri problemi che restavano sul tappeto.

2.2 La robustezza meccanica

La battaglia dell'attenuazione era stata vinta abbandonando completamente le tecnologie vetrarie tradizionali, ivi comprese quelle del vetro d'ottica, ed adottando invece tecnologie mutuata dai semiconduttori, in grado di garantire i necessari livelli bassissimi di impurezze nelle materie prime (non più di alcune parti per miliardo, in peso). Era giocoforza che ciò comportasse un prezzo da pagare in termini di inesperienza; una delle più gravi, fra le conseguenze inattese, fu l'estrema fragilità delle nuove fibre a bassa attenuazione.

Un aneddoto, del quale non si può garantire la veridicità, afferma che la famosissima fibra annunciata dalla Corning a Londra nel 1970 era fragile al punto che non fu mai possibile toglierla dal tamburo su cui era stata avvolta all'atto della filatura:

essa si spezzava allo svolgimento di ogni spira. Chi scrive ebbe invece la ventura di adoperare personalmente una fibra, anch'essa di produzione Corning, ma della generazione immediatamente successiva, per effettuare, presso i Bell Laboratories, alcuni esperimenti di connettorizzazione: la fragilità era tale da procurare allo sperimentatore, per la prima - e fino ad ora unica - volta, una serie sistematica di incubi notturni. Un terzo fatto analogo ai precedenti si riferisce invece alle prime sperimentazioni effettuate dallo CSELT, nei primissimi anni '70: alcune fibre, deposte la sera, dopo la loro filatura, con estrema delicatezza su vassoi rivestiti di velluto, vennero trovate misteriosamente rotte in più pezzi il mattino seguente.

Eliminare in tempi rapidi questo gravissimo inconveniente era, per la nascente tecnologia, una questione di vita o di morte; fortunatamente, questa seconda battaglia si rivelò molto meno dura e molto più breve del temuto. Per vincerla, bastò ricordarsi che, proprio come ci è stato raccontato a scuola, il vetro non è un corpo solido, ma assomiglia piuttosto ad un liquido ad altissima viscosità; pertanto la sua resistenza meccanica è affidata per gran parte alla tensione superficiale. Una volta che la superficie sia stata intaccata da piccole incisioni, la robustezza di un pezzo di vetro risulta drasticamente ridotta, cosa per altro ben nota da molto tempo, tanto ai vetrai quanto agli scassinatori d'appartamenti. E poiché i primi contatti che possono compromettere lo stato di salute della superficie di una fibra hanno luogo quando, all'atto della sua filatura, essa viene avvolta su un cilindro, la soluzione del problema è consistita nel rivestirla immediatamente "in linea", prima ancora della raccolta sul tamburo, con un sottile straterello di plastica, detto appunto *rivestimento primario*, il quale viene rimosso solo in caso di assoluta necessità, come ad esempio per la realizzazione dei giunti a fusione.

2.3 L'affidabilità delle sorgenti

Per un tempo non breve, attorno e dopo la metà degli anni '70, al crescente ottimismo dei produttori di fibre e di cavi, che avevano risolto i problemi meccanici, si contrapposero le preoccupazioni dei "device men", o almeno dei più sinceri tra loro, le quali traevano origine dall'insufficiente affidabilità delle sorgenti laser. Anche esemplari selezionati attraverso una scrupolosa procedura di rodaggio ("burn in") raggiungevano a stento durate di vita di poche migliaia di ore, durante le quali si assisteva ad un lento degrado delle prestazioni, seguito da una rottura catastrofica. Le diagnosi dei guasti furono a lungo timide, incerte, contraddittorie; ciò non deve stupire, perché chi

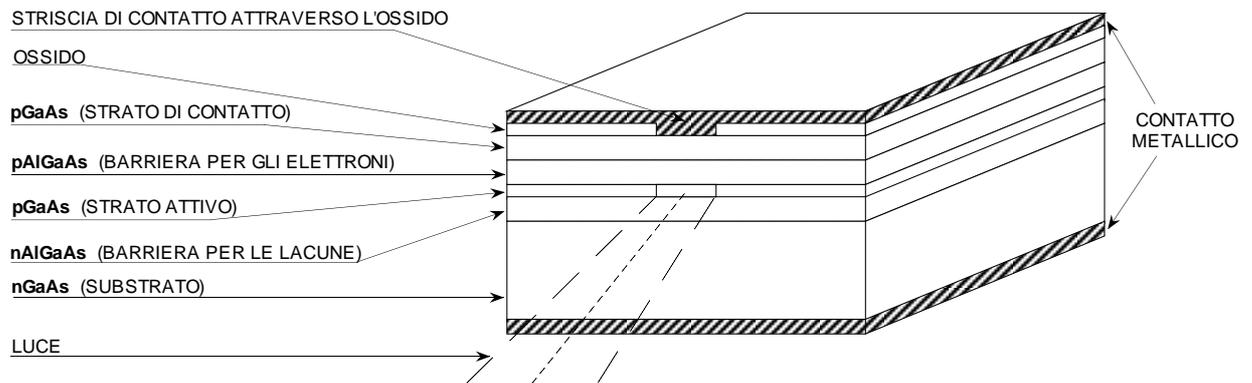


Figura 2

lavorava nel campo delle sorgenti si trovava ad avere a che fare contemporaneamente con materiali (leghe basate sull'arseniuro di gallio) ancora poco sperimentati, specie in termini di produzione di massa, e con tecnologie (i processi epitassiali) molto meno padroneggiate che non quelle classiche del silicio, basate sulla diffusione dei droganti. Una struttura laser come quella schematizzata in fig. 2, che, con gli occhi di oggi, appare più semplice di qualsiasi dispositivo attualmente in commercio, circa vent'anni fa era ben compresa nel suo funzionamento, ma ancora misteriosa nei suoi meccanismi di guasto. Di questi si sapeva, in termini quasi esclusivamente qualitativi, che avevano a che fare con migrazioni delle impurezze verso zone difettose dal punto di vista cristallografico, con conseguente innalzamento a livelli distruttivi delle correnti di soglia necessarie per ottenere il funzionamento laser. I pochi ricercatori che, sulla base di queste conoscenze, affermavano ottimisticamente che nel giro di pochi anni le sorgenti laser avrebbero raggiunto il livello di affidabilità di qualsiasi altro dispositivo a semiconduttore, venivano apertamente sbeffeggiati quali candidati visionari.

Per contro, i diodi elettroluminescenti, o LED, ancorché realizzati in versioni nettamente diverse da quelle per display alfanumerici, e con prestazioni molto più adatte di questi per l'impiego nei sistemi di trasmissione su fibra, potevano già allora vantare durate di vita media di centinaia di migliaia di ore. Non deve quindi stupire se, come si vedrà anche nel seguito, almeno per tutti gli anni '70 non vi fu una forte spinta a favore delle fibre monomodo, in contrapposizione alle multimodo; non si disponeva infatti di una sorgente di luce che fosse accoppiabile con alta efficienza con le fibre monomodo, e al tempo stesso garantisse un tempo medio fra guasti compatibile con le esigenze di manutenzione della rete telefonica.

Ma pure su questo fronte fu dimostrato, nel giro di pochi anni, che gli ottimisti avevano visto giusto.

Anche sotto la spinta delle sue applicazioni nei circuiti a microonde, l'arseniuro di gallio, a lungo definito sarcasticamente "l'eterno materiale del futuro", fu padroneggiato in tutti i suoi aspetti tecnologici; inoltre, e forse sorprendentemente, le lezioni imparate su questo semiconduttore, il cui uso in ottica è ristretto solo alla prima finestra (850 nm circa), consentirono di non ripetere la stessa trafila, in fatto di affidabilità, quando, i potenziali vantaggi della seconda e della terza finestra fecero avviare una sperimentazione intensiva sui semiconduttori basati sul fosforo di indio, oggi dominatori nel campo delle trasmissioni su fibra ottica.

2.4 La capacità di trasmissione

Se misurate in unità di frequenza, le regioni dello spettro elettromagnetico nelle quali l'attenuazione delle moderne fibre ottiche è bassa sono amplissime: come ordine di grandezza, oltre 10^{13} Hz. Ma ciò non basta affatto a garantire di potervi trasmettere segnali a larga banda; per ottenere questa possibilità occorre garantire che la distorsione del segnale sia inferiore a certi limiti, i quali dipendono a loro volta dal tipo di modulazione e rivelazione. Per i sistemi a modulazione d'intensità (*On-Off Keying*, OOK) e rivelazione diretta, sui quali la trasmissione ottica si concentrò in prevalenza fin dall'inizio, e i quali sono tuttora i soli impiegati commercialmente, apparve subito chiaro, fin dai primi anni '70, che la trasmissione su una fibra multimodo a salto d'indice (vedi fig. 3) risulta affetta da una distorsione molto elevata, a causa della diversa velocità di propagazione dei suoi modi; questo fenomeno è detto "dispersione intermodale" e limita le capacità trasmissive teoriche a pochi Mbit/s su distanze di alcuni chilometri.

Due erano, sulla carta, le soluzioni possibili per migliorare sensibilmente questa prestazione, invero assai deludente: le fibre multimodali a gradiente d'indice

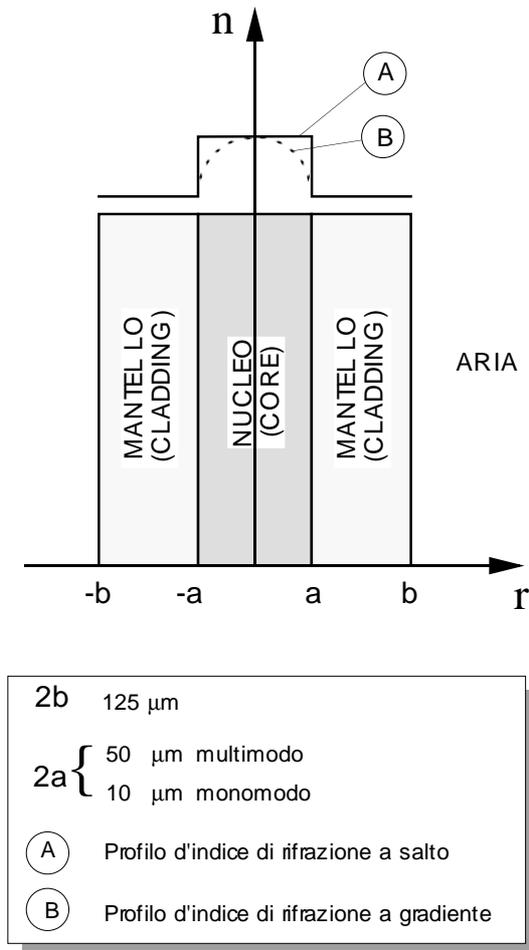


Figura 3

e le fibre monomodali, anch'esse definite in fig. 3. La loro contesa fu lunga e con alterne vicende, motivo per cui è sembrato opportuno tenerla per ultima, fra le grandi battaglie storiche con cui le fibre ottiche si sono conquistate l'attuale posizione di predominio nella trasmissione sulle medie, grandi e grandissime distanze.

Negli anni '70, in assenza, come detto poc' anzi, di una sorgente affidabile ed idonea all'impiego con una fibra monomodo, e forse anche per eccessivo timore delle difficoltà pratiche legate alla realizzazione, al cablaggio e alla giunzione di fibre il cui nucleo centrale doveva avere un diametro di non più di 10 μm, la bilancia sembrò pendere sensibilmente a favore delle multimodo a gradiente d'indice. Per quasi un decennio, un problema "alla moda" fu quello del profilo ottimale, cioè la ricerca dell'andamento dell'indice di rifrazione, in funzione della coordinata radiale, che minimizzava la distorsione intermodale, a parità di tutti gli altri parametri.

Nella risposta sperimentale a questo quesito, che è condensata graficamente nella fig. 4, stanno le ragioni sia delle effimere fortune delle fibre a gradiente d'indice, sia della loro decadenza, fino al quasi completo abbandono, a favore delle monomodali.

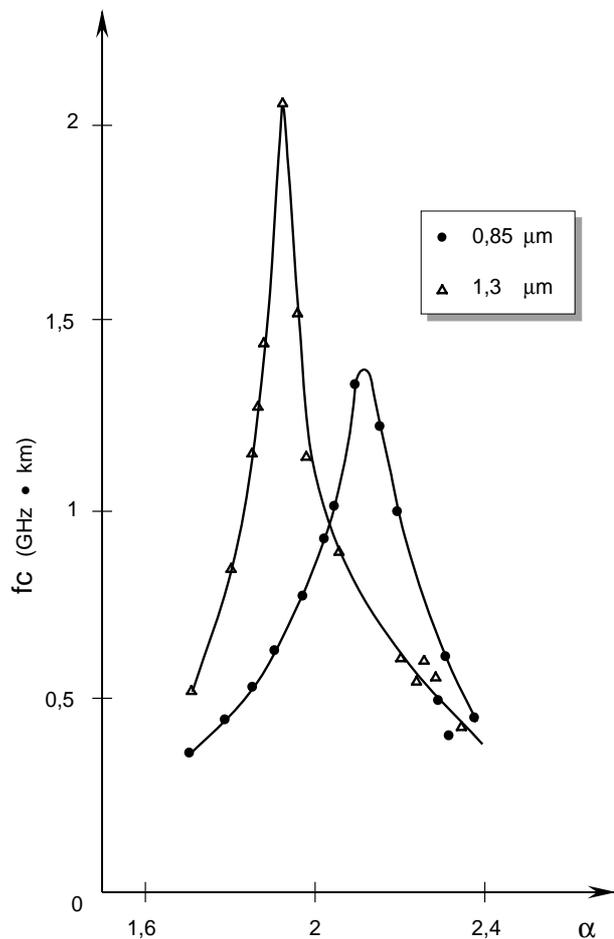


Figura 4

La fig. 4 rappresenta l'andamento del prodotto (larghezza di banda)×(distanza), ottenibile con fibre a gradiente d'indice, al variare del parametro α del profilo. In particolare il profilo d'indice, ovvero l'andamento dell'indice di rifrazione $n(r)$ nella fibra per una data sezione trasversale (supposta circolare), in funzione del raggio r dal centro della sezione, risulta dato da:

$$n^2(r) = \begin{cases} n_0^2 \left[1 - 2\Delta \left(\frac{r}{a} \right)^\alpha \right], & r \leq a \\ n_1^2, & r > a \end{cases}$$

ove a è il raggio del nucleo, n_1 è il valore dell'indice di rifrazione al bordo del nucleo e Δ è l'apertura numerica della fibra, espressa da

$$\Delta = \frac{n_0^2 - n_1^2}{2n_0^2}$$

Le curve tracciate in fig. 4 dimostrano in primo luogo che con l'adozione di un profilo d'indice "ottimizzato" si possono ottenere capacità trasmissive molto superiori a quelle delle fibre a salto d'indice. Ma

esse mostrano anche che il profilo ottimale è estremamente critico: la teoria ci dice che errori dell'ordine di qualche parte per mille nel profilo riducono dette capacità di oltre un ordine di grandezza. Un'altra cosa mostrata dalla stessa figura è che l'ottimizzazione del profilo è molto sensibile alla lunghezza d'onda, e che quindi una fibra, realizzata ed installata in certi tempi, non si presta al "retrofitting", qualora, con l'andare del tempo, si dimostri conveniente il passaggio a sorgenti e a rivelatori funzionanti in un'altra finestra. V'è infine un altro inconveniente, intuibile dalla fig. 4 grazie alla forma simmetrica delle curve attorno al loro massimo: la deviazione dall'ottimo può avvenire sia per "sottocompensazione", cioè perché il profilo è troppo simile a un gradino, sia per "sovracompensazione", cioè per il difetto opposto. La connessione in cascata di due fibre appartenenti alle due categorie può dare luogo ad una reciproca equalizzazione e quindi ad un apparente miglioramento della capacità trasmissiva dell'intero collegamento, ma la natura intrinsecamente aleatoria di questo effetto lo rende difficilmente modellizzabile e quindi pericoloso, anziché gradito, all'atto pratico.

Per questi motivi, le soluzioni multimodali a gradiente d'indice, anche se segnarono una tappa molto importante nel progresso delle comunicazioni ottiche, in quanto per prime dimostrarono la realizzabilità di collegamenti a larga banda su grandi distanze, a lungo andare persero la battaglia della capacità trasmissiva negli anni '80, a favore delle fibre monomodali. La minimizzazione della distorsione che in esse si può ottenere grazie alla reciproca compensazione fra gli effetti del materiale (dispersione cromatica⁽¹⁾) e quelli strettamente propagativi (dispersione di guida⁽²⁾), e che può aversi a qualsiasi lunghezza d'onda d'interesse pratico compresa fra 1300 e 1550 nm, è una realtà ben consolidata. In particolare, nelle cosiddette fibre a dispersione spostata, cioè quelle in cui la compensazione di cui si è ora detto ha luogo attorno a 1550 nm, si è conseguito l'importante obiettivo di far coincidere la lunghezza d'onda di minima attenuazione con quella di minima dispersione totale. Ma questa è realtà di oggi, certamente vissuta in prima persona da tutti i lettori del Notiziario Tecnico.

(1) Fenomeno dovuto alla larghezza spettrale non nulla delle sorgenti laser (le cui emissioni contengono quindi un tutti i valori di lunghezze d'onda compresi in un intervallo centrato attorno ad un valore nominale) e alla variabilità dell'indice di rifrazione della fibra in funzione della lunghezza d'onda stessa.

(2) Fenomeno dovuto all'esistenza di una molteplicità (numerabile) di campi e.m. che sono soluzioni delle equazioni di Maxwell applicate alla propagazione nella fibra (modi), ciascuna delle quali ha una propria velocità di fase e di gruppo.

Prima di passare ad una rapida sbirciata al futuro, il cui scopo principale è quello di evitare che il lettore tragga da quanto ora detto l'erronea conclusione che le battaglie siano definitivamente terminate, è bene dedicare un breve cenno alle vicende di casa nostra, cioè al contributo italiano negli anni eroici di cui si è fino a qui parlato, per evitare un'altra conclusione erronea, quale sarebbe il pensare che in Italia non sia stato compreso tempestivamente il peso della nuova tecnologia, o che si sia colpevolmente tardato nel destinare risorse a queste ricerche.

3. Le fibre ottiche in Italia negli anni '70

Sul finire degli anni '60, era accaduto un fatto assolutamente nuovo, per il nostro Paese, nell'area della ricerca applicata. Erano stati varati, a brevissima distanza di tempo l'uno dall'altro, due progetti di respiro nazionale: il primo era il Progetto Sirio, che, come è noto a molti, riguardava il campo delle trasmissioni via satellite; il secondo riguardava invece le trasmissioni su guide d'onda millimetriche per grandi distanze. Il carattere innovativo consisteva nel fatto che in entrambi, sotto la leadership di un organo scientifico altamente qualificato (il Centro Comunicazioni Spaziali del CNR nel primo caso, un Centro costituito congiuntamente dalla Fondazione U. Bordoni e dalla Fondazione G. Marconi nel secondo), venivano chiamati a collaborare in maniera coordinata gruppi di ricerca universitari, laboratori del CNR e di terzi, ed aziende industriali. Si tratta di un'organizzazione della ricerca che oggi può apparire ovvia, ma che allora era del tutto inesplorata in Italia.

Quando, all'indomani stesso dei primi grandi passi a livello mondiale di cui si è detto prima, sembrò opportuno partire di gran carriera anche da noi nel campo delle fibre ottiche, buona parte delle risorse disponibili presso le Università e le altre strutture pubbliche di ricerca era saturata da questi due progetti, o da altri interessi più isolati. Non era facile prevedere allora che le guide d'onda millimetriche sarebbero uscite battute dallo scontro con le fibre, e non sarebbe stato serio abbandonare a metà un lavoro ben pianificato e per il quale erano state chieste e ottenute risorse finanziarie cospicue, se misurate sulla scala dell'epoca e del Paese.

Fortunatamente, non si creò un vuoto perché fu prontissima e lungimirante la risposta da parte industriale, in primo luogo quella del Gruppo Stet e della Società Cavi Pirelli; se così non fosse stato, probabilmente l'Italia avrebbe perso clamorosamente, fin dall'inizio, l'autobus delle fibre ottiche. Fra i fatti che sono rimasti a documentare, in maniera incontrovertibile, l'immediato intervento della componente industriale vi sono le prime fibre per telecomunicazioni realizzate dallo CSELT già nel 1972; fra quelli che dimostrano invece che detta

risposta non è stata momentanea né effimera, v'è il fatto che 5 anni dopo, nel primo esperimento italiano di collegamento in fibra fra due centrali SIP - il famoso esperimento COS2 effettuato a Torino nel 1977 - il cavo (Pirelli) conteneva fibre CSELT a fianco delle fibre Corning, di qualità assolutamente equivalenti. Si tratta di risultati che ancor oggi, retrospettivamente, mostrano che l'Italia era allora più avanti in questo campo di molti altri tra i Paesi tecnologicamente più sviluppati del mondo.

Purtroppo, fra le conseguenze di questi successi della ricerca industriale ve ne fu una negativa, assolutamente impensabile alla vigilia: i gestori della ricerca pubblica ne trassero la conclusione che ormai il più era stato fatto, e che quindi non era il caso di investire denaro dello Stato nella ricerca in questo settore. Accadde così che i risultati di uno studio di fattibilità, svolto in perfetta armonia da Università e industrie, per un Progetto Finalizzato CNR "Comunicazioni su fibre ottiche", alla metà degli anni '70, venne lasciato cadere, pur fra un coro di elogi e di promesse, nel disinteresse, fino alla completa dimenticanza. E quanto ci fosse ancora da studiare sull'argomento, anche nelle sedi che istituzionalmente sono chiamate ad occuparsi di ricerca di base, lo ha dimostrato l'esperienza vissuta da noi tutti, e lo dimostra tuttora il contenuto del prossimo paragrafo.

4. I fronti ancora caldi

È opinione estremamente diffusa, almeno tra i giovani studenti d'ingegneria, che la tecnologia delle fibre ottiche, dopo i brillanti successi conseguiti nell'ultimo ventennio, sia ormai giunta ad uno stato stazionario e non riservi sorprese particolarmente eccitanti sul versante della ricerca. Sarebbe grave non cercare di sfatare questa credenza, citando almeno alcuni fra i fronti sui quali si sta combattendo oggi, per ottenere ulteriori, significativi miglioramenti nei sistemi di trasmissione, o per aprire alle tecnologie ottiche altri spazi in aree contigue.

Iniziando dagli argomenti che appaiono già più maturi, per andare poi verso quelli più proiettati nel futuro, non si può non citare per primi gli *amplificatori ottici* realizzati mediante fibre drogate con ioni di terre rare, i quali stanno già oggi sconvolgendo la progettazione dei sistemi di trasmissione, iniziando da quelli transoceanici e da quelli terrestri a grandissima capacità, riducendo enormemente il numero dei rigeneratori elettronici. Nel contempo, i *laser tutto-fibra*, parenti strettissimi, in termini tecnologici, degli amplificatori testè citati, avanzano con la dichiarata intenzione di contendere ai laser a semiconduttore uno spazio che, fin dai tempi di Kao e Hockham, nessuno aveva mai pensato potesse venire occupato da altri.

Un secondo fronte incandescente sul quale potrebbero venire sfornate presto grandissime novità a livello applicativo, nelle trasmissioni a grande distanza, è lo sfruttamento degli effetti non lineari nella propagazione lungo le fibre ottiche, o, in termini più specifici, la *propagazione solitonica*. Si tratta di un argomento talmente rivoluzionario rispetto alla cultura tradizionale di un ingegnere di telecomunicazioni, da rendere assolutamente velleitario qualsiasi tentativo di spiegare in poche parole di che cosa si tratti: meglio restare in attesa di un contributo *ad hoc*.

Un fronte più lontano, ma certo non meno affascinante, è lo sfruttamento degli effetti non lineari, o nelle fibre, o in dispositivi ottici di tipo bi- o tri-dimensionale, per realizzare sistemi di *commutazione ottica*, o più in generale di *elaborazione interamente ottica* dei segnali. Anche questi argomenti sono troppo vasti e delicati per poterli liquidare qui, con poche e frettolose parole. Si è ritenuto opportuno citarli solo per poter concludere questa nota suggerendo ai lettori una piccola scommessa: chi fra loro è disposto a credere che, fra vent'anni, su queste pagine, o su quelle che ne avranno preso il posto, vi sarà spazio per una noticina storica affine a questa, dedicata alle vicende iniziali della trasmissione solitonica, oppure a quelle della commutazione ottica?

Fibre ottiche per Telecomunicazioni: propagazione

P. Di Vita, V. Lisi, M. Giaconi, G. Vespasiano (*)

Questo articolo presenta la teoria della propagazione su fibra ottica collocandola nella cornice attuale delle comunicazioni su fibra e delineando gli sviluppi prevedibili per il medio termine.

La fibra è descritta nella sua costituzione; viene giustificato l'andamento dell'attenuazione con la lunghezza d'onda della luce a partire dalle cause fondamentali di assorbimento; vengono presentati i tipi di fibre monomodali che sono oggi di più largo e comune impiego.

La teoria della propagazione è sviluppata per le fibre con profilo dell'indice di rifrazione a gradino; per esse viene presentata la classificazione dei modi di propagazione e la condizione di taglio del modo; infine, si presentano i parametri di dispersione modale e cromatica. Sebbene la trattazione risulti assolutamente generale, si vuole porre l'accento sul caso di fibre monomodali che, come già detto, sono oggi quelle più largamente utilizzate.

La propagazione nelle fibre monomodali verrà ulteriormente ripresa ed approfondita in un lavoro successivo.

1. Introduzione

Per gli innumerevoli vantaggi che offrono, le fibre ottiche monomodali sono il mezzo trasmissivo più largamente usato nelle reti di telecomunicazione. Ad oggi sono state posate nel mondo, in impianti terrestri, più di 10 milioni di chilometri di fibra. Nella assoluta maggioranza dei casi si tratta di fibra ottica monomodale ottimizzata in dispersione per la lunghezza d'onda di 1310 nm (fibre standard). D'altra parte l'attenuazione delle fibre ottiche in silice ha il suo minimo intorno ai 1550 nm; sono state sviluppate quindi fibre a dispersione spostata (*dispersion-shifted*: DS) caratterizzate dall' avere il punto di zero della dispersione cromatica proprio in corrispondenza di tale lunghezza d'onda (terza finestra). Queste fibre, che quindi esprimono al meglio le loro prestazioni in collegamenti a lunga distanza ad elevata velocità di cifra, verranno sempre più utilizzate nelle future installazioni. Considerando l'attuale stadio di sviluppo di nuovi componenti, ed il maturare di tecniche trasmissive quali la moltiplicazione a divisione di lunghezza d'onda (*Wavelength Division Multiplexing*: WDM), si può collocare intorno ai 100 Gbit/s la capacità trasmissiva

cumulativa di un collegamento in fibra ottica con gli strumenti già oggi disponibili.

Questo panorama è arricchito da un ulteriore elemento: l'amplificatore ottico a fibra attiva, capace di amplificare segnali in una banda di circa 40 nm (migliaia di GHz) nella regione della terza finestra. L'introduzione dell'amplificatore ottico risulterà fondamentale in tutti i sistemi di trasmissione su fibra, ma in particolare per i lunghi collegamenti sottomarini; esso consente infatti di evitare i costosi rigeneratori optoelettronici oggi utilizzati e realizza di fatto la trasparenza del collegamento rispetto al formato trasmissivo, nel senso che, in momenti successivi alla prima installazione, si potrà elevare la velocità trasmissiva o passare a tecniche WDM senza sostituzione dei ripetitori in linea. Sono già state dimostrate trasmissioni a 2.5 Gbit/s, 5 Gbit/s, 10 Gbit/s su collegamenti sperimentali, ma in condizioni reali, lunghi fino a 9000 km [1]. Ciò rende ancora più concreta la previsione di un primo collegamento transatlantico amplificato in esercizio a partire dal 1995. D'altra parte collegamenti sottomarini a 2.5 Gbit/s lunghi fino a quasi 300 km potranno essere realizzati con una linea completamente passiva, cioè localizzando gli amplificatori solo nei punti di approdo [2].

Si è dimostrato, per via teorica e sperimentalmente, che la trasmissione su lunghe distanze ad oltre 10 Gbit/s per segnali "convenzionali" (di tipo NRZ) è molto difficoltosa. Il limite principale è dato in questo caso dagli

(*) dott. Pietro Di Vita - CSELT - Torino; dott.ssa Vincenza Lisi - Università di Torino - Torino; ing. Mauro Giaconi, ing. Gaetano Vespasiano - SSGRR - L'Aquila

effetti dispersivi lungo la fibra. Tuttavia esiste una alternativa possibile che è quella della trasmissione di solitoni. Nel caso ideale, un solitone è un impulso ottico molto breve, con involuppo dato da una secante iperbolica⁽¹⁾, che può propagarsi in un mezzo privo di perdite senza cambiare forma. Ciò può ottenersi perché per effetto Kerr (dipendenza non-lineare dell'indice di rifrazione della fibra dalla potenza ottica trasmessa) si riesce a cancellare l'allargamento degli impulsi dovuto a dispersione [3]. Requisito essenziale perché ciò possa verificarsi è che all'impulso ottico sia associata una potenza adeguata, che dipende dalla larghezza dell'impulso, dalla lunghezza d'onda e dai parametri della fibra (dispersione cromatica e diametro di campo modale, vd. § 2). Ad esempio per poter stimolare la produzione di solitoni in una fibra monomodale standard, alla lunghezza d'onda di 1550 nm, è necessaria una potenza dell'ordine dei 48 mW [4], per una fibra *dispersion-shifted*, in condizioni analoghe, la potenza necessaria si riduce di circa un ordine di grandezza. Per ottenere una potenza del genere in trasmissione, ma soprattutto per garantire un adeguato livello di potenza lungo la linea ottica, è necessario ricorrere ad amplificatori ottici distribuiti lungo il collegamento. Con tali tecniche trasmissive sono state realizzate diverse sperimentazioni che hanno dimostrato la possibilità di collegamenti a 10 Gbit/s su 20000 km e di 20 Gbit/s su 13000 km usando fibre a dispersione spostata [5], oppure collegamenti a 4 Gbit/s su 310 km con fibre standard [6].

“La rete di distribuzione attuale è stata dimensionata per fornire il servizio telefonico analogico. I servizi a larga banda del XXI secolo richiederanno una rete completamente ottica ma, quando la domanda di servizi a larga banda sarà nata, sarà troppo tardi per realizzare una nuova rete” [7]. L'introduzione delle fibre ottiche nella rete di distribuzione avverrà, e sarà massiccia, quando sarà sciolto il dilemma rappresentato nelle poche righe precedenti. La rete di distribuzione esistente soddisfa le esigenze attuali dell'utenza; il costo di una nuova rete, a larga banda, può giustificarsi solo con l'offerta di nuovi servizi, che tuttavia per essere richiesti devono essere conosciuti ed apprezzati, e lo diventeranno solo quando saranno disponibili. Per superare quel fantomatico “ultimo miglio” che impedisce alle fibre l'accesso in rete di distribuzione, è necessario sviluppare delle tecnologie adeguate, che riducano il costo dei sistemi ottici nella fornitura dei servizi esistenti raggiungendo la parità con i sistemi su rame. La soluzione che sembra delinearsi come più efficace fa ricorso alla trasmissione punto-multipunto, realizzata con divisori 1xN del fascio ottico, distribuiti lungo la linea che dalla centrale penetra nell'area

(1) La funzione secante iperbolica è definita dalla relazione $sech(x) = 1/\cosh(x) = 2/[\exp(x) + \exp(-x)]$, $x \in (-\infty, \infty)$; la forma del grafico di $sech(x)$ è una campana con il massimo in $x=0$.

di distribuzione. Il terminale di linea (ottico) lato utente potrà essere dedicato o subire un grado variabile di condivisione a seconda delle dimensioni dell'utenza servita e dell'intensità di traffico da essa generato/richiesto.

In un'epoca in cui esplodono le comunicazioni personali di tipo *wireless* c'è peraltro un interesse considerevole nel realizzare una “giudiziosa” integrazione tra fibre ottiche e trasmissione radio a microonde o con onde millimetriche [8, 9]. In generale la funzione per cui le fibre ottiche possono essere vantaggiosamente impiegate è il collegamento tra la centrale e le antenne di un sistema radio o radiomobile, superando tutti i limiti imposti finora dall'uso di guide d'onda metalliche o cavi coassiali in alta frequenza. Tale applicazione, particolarizzata ogni volta alle specifiche funzionalità richieste, assume ulteriore importanza quando il segnale radio debba essere distribuito in ambienti particolari quali aeroporti, stazioni ferroviarie, complessi fieristici, dove la distribuzione del traffico costringe ad operare con celle molto piccole (picocelle o femtocelle) e con sezioni di passaggio per i cavi che possono essere molto anguste.

In questo articolo vengono presentate le caratteristiche generali delle fibre ottiche per telecomunicazioni, soprattutto con riferimento ai parametri trasmissivi e presentando in maniera succinta, ma completa, la teoria della propagazione in fibra. In un lavoro successivo verranno analizzati i parametri caratteristici delle fibre ottiche monomodali che, come già accennato in precedenza, costituiscono la assoluta maggioranza della rete mondiale in fibra ottica.

2. Caratteristiche delle fibre ottiche per Telecomunicazioni

Dopo un periodo iniziale di proliferazione nei tipi e nelle dimensioni delle fibre, l'azione di razionalizzazione in sede normativa internazionale ha consentito di delineare un panorama relativamente chiaro ed assestato. La Fig. 1 mostra la costituzione di una fibra. La regione più interna è detta nucleo (*core*) ed è ricoperta da un mantello (*cladding*) che, come si dirà meglio nel seguito, ha indice di rifrazione leggermente minore. L'insieme di queste due parti vetrose della fibra è poi protetto esternamente da una guaina, detta

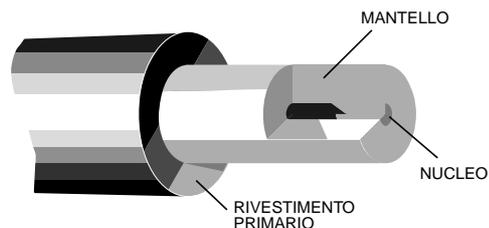


Figura 1 Costituzione di una fibra ottica

Denominazione	Racc.	\varnothing_{co}/MFD (μm)	\varnothing_{cl} (μm)	λ_{cc} (nm)	λ (nm)	α (dB/km)	D (ps/nm km)
multimodale "graded-index"	G.651	50 \pm 3	125 \pm 3	---	850 1310	<4 <2	--- ---
monomodale "standard"	G.652	(9-10) \pm 10% @1310nm	125 \pm 2	<1270	1310 1550	<1 <0.5	3.5 20
monomodale "dispersion- shifted"	G.653	(7-8.3) \pm 10% @1550nm	125 \pm 2	<1270	1310 1550	<1 <0.5	(20) 3.5
monomodale "att.ne minimizzata"	G.654	10.5 \pm 10% @1550nm	125 \pm 2	<1530	1550	<0.22	20

Tabella 1 Caratteristiche delle fibre ottiche secondo le raccomandazioni CCITT (\varnothing_{co} : diametro del *core* (fibre multimodali); MFD: diametro di campo modale (fibre monomodali); \varnothing_{cl} : diametro del *cladding*; λ_{cc} : lunghezza d'onda di taglio misurata sulla fibra assemblata nel cavo; λ : lunghezza/e d'onda consigliata/e per l'esercizio; α : attenuazione; D : dispersione cromatica

rivestimento primario (*primary coating*), che è di solito in acrilato plastico. In una fibra monomodale il rapporto tra il raggio del nucleo e la lunghezza d'onda di esercizio è piccolo; è molto piccola poi la differenza tra l'indice di rifrazione del nucleo e quello del mantello, quindi l'effetto di confinamento della luce nel nucleo è limitato (si parla di debole guida) e, di conseguenza, la distribuzione di potenza ottica in fibra interessa anche il mantello. Per questo motivo non basta il valore del diametro del nucleo per definire la regione dove è confinata la luce, si ricorre allora ad un ulteriore parametro detto diametro di campo modale (*mode field diameter*: MFD) che, in linea di massima, definisce le dimensioni della regione entro cui è contenuta la maggior parte della energia luminosa che attraversa una certa sezione. In tabella I sono riportate le caratteristiche delle quattro classi di fibre ottiche così come sono state individuate dalle raccomandazioni del CCITT (adesso ITU-T, International Telecommunication Union - Telecommunication Standardization Sector).

2.1 Attenuazione della fibra

Parte dell'energia luminosa che si propaga lungo la fibra viene assorbita dal materiale o viene diffusa in esso, costituendo quindi una perdita ai fini del segnale trasmesso. Il rapporto tra la potenza ottica trasmessa e quella ricevuta, dopo una lunghezza di fibra di riferimento, definisce l'attenuazione della fibra stessa, che è una funzione della lunghezza d'onda, del tipo di fibra, e delle eventuali sollecitazioni meccaniche che agiscono sulla fibra. Si distingue in genere tra cause di attenuazione intrinseche ed estrinseche, le prime sono insite nella struttura fondamentale del materiale e non possono essere eliminate, le seconde possono invece ricondursi alla presenza di impurezze ed essere ridotte affinando il processo produttivo. L'andamento dell'attenuazione in funzione della lunghezza d'onda, evidenziando le varie cause che la determinano, è riportato (in termini qualitativi) in Fig. 2.

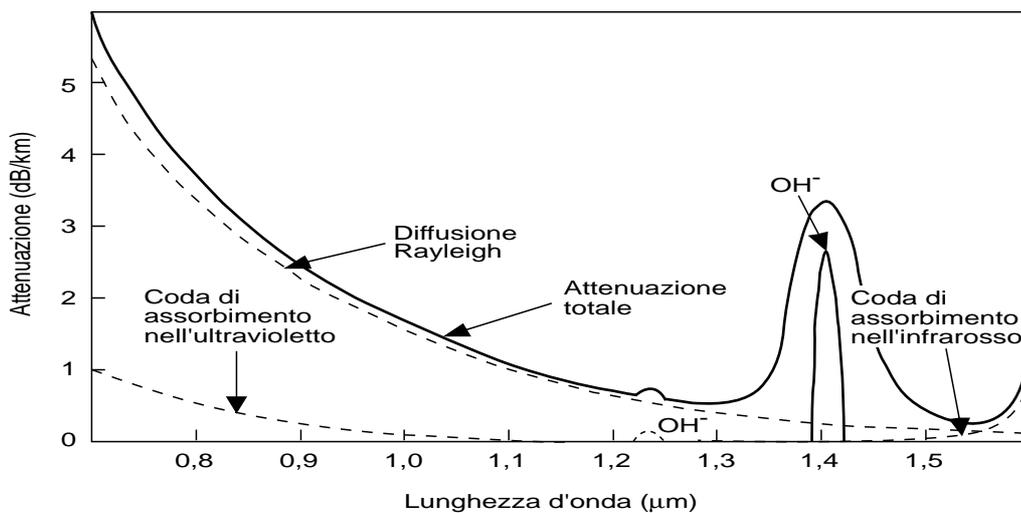


Figura 2 Andamento dell'attenuazione di una fibra ottica in funzione della lunghezza d'onda, con evidenza delle varie componenti che la determinano

Una prima causa di attenuazione intrinseca è l'assorbimento dovuto a transizioni elettroniche. Infatti le transizioni elettroniche, a livello atomico, dalla banda di valenza a quella di conduzione, causano un assorbimento alle lunghezze d'onda più basse, tipicamente nell'ultravioletto, che diminuisce esponenzialmente all'aumentare della lunghezza d'onda. Per la silice pura l'effetto di questo tipo di assorbimento nell'infrarosso è trascurabile. Tuttavia, il drogaggio della silice con modificatori del reticolo quali il GeO_2 , necessario per aumentare l'indice di rifrazione del nucleo, fa spostare la coda della curva di assorbimento verso lunghezze d'onda più elevate. L'entità dell'assorbimento alla generica lunghezza d'onda può essere espresso, in dB/km/ppmw Ge (ppmw Ge = parti per milione in peso di germanio), come

$$\alpha_\lambda = \alpha_0 \exp(E/E_0)$$

dove $\alpha_0 = 1.474 \cdot 10^{-11}$ dB/km/ppmw, $E_0 = 0.268$ eV ed E è l'energia del fotone alla lunghezza d'onda considerata.

Una seconda causa è l'assorbimento dovuto a vibrazioni molecolari. I singoli elementi che costituiscono il vetro, si consideri per esempio la coppia Si-O, sono infatti tenuti insieme da legami chimici più o meno forti. La struttura cui danno luogo tuttavia non è statica, ma assume uno stato vibrazionale continuo e del tutto casuale dovuto ad agitazione termica. D'altra parte ad ogni legame chimico è associato un dipolo elettrico, la cui intensità di campo è modulata dalla vibrazione. Si ha quindi un'interazione tra il campo elettromagnetico associato alla radiazione che si propaga ed il dipolo, con trasferimento di energia dal campo alla struttura. Il tutto si risolve quindi in un assorbimento. Per i vetri ad alto tenore di silice, la vibrazione fondamentale per i legami B-O, P-O, Si-O e Ge-O dà luogo a picchi di assorbimento centrati rispettivamente alle lunghezze d'onda di 7.3 μm , 8.0 μm , 9.0 μm , 11.0 μm . In particolare il B_2O_3 utilizzato nelle prime fibre ottiche ha una curva di assorbimento la cui coda si estende fino a lunghezze d'onda poco superiori a 1.2 μm . Di conseguenza, per le fibre a bassa attenuazione il drogaggio con boro è assolutamente evitato, mentre il fosforo è mantenuto sempre a livelli molto bassi e si cerca di evitarne l'uso nelle fibre monomodali. Per i vetri di silice drogati al germanio il contributo di attenuazione dovuto ad assorbimento nell'infrarosso può essere espresso, in dB/km se la lunghezza d'onda viene espressa in μm , come [10]

$$\alpha_{IR} = 7.81 \cdot 10^{11} \exp(-44.48/\lambda)$$

Si noti che questa espressione, ricavata per le fibre costruite alla fine degli anni '70, dà luogo a valori certamente pessimistici riguardo all'attenuazione in terza finestra delle fibre ottiche attuali, le quali hanno evidentemente un minor tenore in germanio.

Un'altra causa d'attenuazione intrinseca è la diffusione alla Rayleigh nel vetro. Invero, il mezzo dielettrico in cui la luce si propaga presenta delle irregolarità microscopiche. Ognuna di queste, le cui dimensioni sono dell'ordine di un decimo della lunghezza d'onda, costituisce un centro di diffusione della radiazione. Nel vetro le irregolarità sono dovute contemporaneamente a fluttuazioni di densità e di composizione; i vetri sono infatti strutture disordinate, debolmente legate in sequenze altamente casuali. Le fluttuazioni di densità derivano dai moti browniani che interessano la massa liquida prima della solidificazione. La composizione varia spazialmente in quanto il materiale base viene drogato con ossidi di tipo diverso per modificare l'indice di rifrazione. La natura dei fenomeni impedisce una descrizione in termini deterministici, l'effetto complessivo è quello di una perdita di energia per diffusione la cui entità può essere descritta usando la legge di Rayleigh. I vetri di silice pura o solo debolmente drogati al fosforo presentano le minori perdite per diffusione; al contrario il drogaggio con GeO_2 fa aumentare il coefficiente di diffusione della silice. Il contributo di attenuazione per diffusione di Rayleigh, in dB/km se λ è espresso in μm , può essere assunto pari a

$$\alpha_R = A \lambda^{-4}$$

Per vetri a base di silice, nella regione 1.3-1.6 μm , il coefficiente di diffusione A assume valori 0.7-1.0 dB/km· μm^{-4} per fibre monomodali e 1-1.6 dB/km· μm^{-4} per fibre multimodali; la differenza tra i due valori è pienamente giustificabile con la maggior percentuale di germanio presente nelle fibre ottiche multimodali.

Le componenti estrinseche di attenuazione sono riconducibili alla presenza nel vetro di materiali estranei alla composizione desiderata. In particolare i metalli di transizione hanno bande di assorbimento ampie con valori di attenuazione molto elevati, essi devono pertanto essere ridotti al livello di poche parti per miliardo perché il contributo in attenuazione sia trascurabile. In effetti, nei vetri di silice ottenuti per crescita in fase vapore questo tipo di assorbimento è praticamente eliminato. Più rilevante è invece il contributo dovuto all'assorbimento dell'ossidrile OH^- . L'acqua presente nel vetro come OH^- ha la sua vibrazione fondamentale compresa tra 2.7 e 3 μm , a seconda del tipo di composizione. Per la silice pura la prima armonica cade a 1.38 μm dando luogo ad un'attenuazione incrementale di 48 dB/km·ppmw Si-OH, una componente ulteriore è presente a 1.24 μm con un'attenuazione di 2.5 dB/km·ppmw Si-OH. La larghezza di ciascuna di queste regioni di assorbimento è tale da interessare sia la seconda che la terza finestra, è pertanto necessario ridurre l'ossidrile a poche parti per miliardo per evitare un sensibile aumento dell'attenuazione. Peraltro il drogaggio con germanio allarga ulteriormente

la regione di assorbimento verso le lunghezze d'onda più alte (la prima armonica per Ge-OH cade a $1.4 \mu\text{m}$) e il drogaggio con fosforo dà luogo ad una nuova regione di assorbimento collocata esattamente in terza finestra.

L'attenuazione di una fibra non dovrebbe cambiare durante la sua vita utile; tuttavia in presenza di idrogeno questa condizione può venire meno. Nel microambiente costituito dal cavo le sorgenti di idrogeno non mancano: basti pensare ai processi di corrosione in generale ed alla corrosione di natura elettrolitica delle parti metalliche del cavo in particolare. Si ha inoltre produzione di idrogeno da parte di alcuni dei rivestimenti utilizzati per le fibre, tipicamente quelli siliconici. Ancora idrogeno è prodotto per interazione tra i diversi materiali che costituiscono il cavo, né si deve trascurare il fondo di idrogeno presente nell'ambiente (10^{-6} atm). L'aumento di attenuazione indotto dall'idrogeno è dovuto a due cause: l'assorbimento diretto e quello da parte di composti che si formano per la presenza di idrogeno nel vetro. La struttura reticolare aperta dei vetri di silice consente all'idrogeno un'agevole penetrazione; esso quindi si dissolve nel vetro in forma molecolare (H_2) ed occupa gli interstizi legandosi debolmente al biossido di silicio. La vibrazione fondamentale dell'idrogeno dà luogo ad un massimo di assorbimento collocato a $2.42 \mu\text{m}$, con armoniche a 1.245 e $0.851 \mu\text{m}$ e una serie di picchi di assorbimento tra $1.08 \mu\text{m}$ e $1.24 \mu\text{m}$ che fanno aumentare nettamente l'attenuazione in tutta la parte alta dello spettro. È stato rilevato sperimentalmente che alla temperatura di 20°C l'attenuazione aumenta di $0.3 \text{ dB/km/atm H}_2$ a 1300 nm , e di $0.6 \text{ dB/km/atm H}_2$ a 1550 nm . Il fenomeno è reversibile, nel senso che rimuovendo l'idrogeno scompare l'effetto sull'attenuazione. Il secondo effetto dell'assorbimento è invece irreversibile, esso è dovuto alla reazione dell' H_2 con le specie chimiche presenti nel reticolo con formazione di gruppi OH^- , oppure a difetti strutturali indotti dall'idrogeno e legati alla contemporanea presenza di germanio. Per le fibre monomodali al germanio l'assorbimento è centrato a lunghezze d'onda molto basse ed ha solo un debole effetto nella parte di spettro di interesse. Le contromisure, che non sono affatto semplici, consistono nel controbattere una per una tutte le possibili cause di sviluppo di idrogeno e nel realizzare rivestimenti impermeabili all'idrogeno.

L'attenuazione della luce nella propagazione lungo la fibra dipende anche dalla geometria del suo asse, nel senso che, in generale, ogni punto di curvatura causa una perdita. A seconda del rapporto tra raggio di curvatura e dimensioni della fibra si distingue in genere tra macro e micro-curvature, le prime sono le curvature macroscopiche cui la fibra è sempre soggetta in sede di posa, le seconde sono invece caratterizzate da raggi di curvatura molto piccoli e variabili in maniera casuale lungo la fibra. La radiazione dalla fibra curvata non è continua, ma può essere rappresentata con la comparsa di raggi divergenti distribuiti su angoli discreti [11].

L'attenuazione che ne deriva è riconducibile a due fenomeni distinti [12]: la perdita alla transizione tra fibra dritta e fibra curvata, la perdita per radiazione nel tratto con curvatura costante. Nel successivo articolo, specificamente dedicato alle fibre monomodali, si chiarirà che i parametri decisivi per determinare la perdita per curvature sono: il salto d'indice (la differenza relativa tra il valore dell'indice di rifrazione nel nucleo e nel mantello, da cui dipende il grado di confinamento della luce all'interno del nucleo), la lunghezza d'onda, il rapporto tra dimensioni geometriche e diametro di campo modale [13, 14].

2.2 Tipi di fibre

I requisiti essenziali per una fibra ottica per telecomunicazioni sono: basse perdite, caratteristica di dispersione adeguata alle condizioni di impiego, piccole perdite incrementali per micro e macro-curvatura. Assumendo la dispersione come elemento caratterizzante, si possono individuare tre classi di fibre ottiche monomodali (Fig. 3): le fibre a dispersione standard, quelle a dispersione spostata (*dispersion shifted*) e quelle a dispersione appiattita (*dispersion flattened*).

Le fibre più comunemente usate nelle telecomunicazioni sono quelle con dispersione ottimizzata nella regione dei 1310 nm . Con riferimento alla forma del profilo dell'indice di rifrazione esse sono sempre a salto d'indice (*step-index*) e possono essere di tipo a mantello continuo (*matched-cladding*, Fig. 3a) oppure a mantello depresso (*depressed-cladding*, Fig. 3b). Nelle fibre a mantello continuo l'indice di rifrazione del mantello è uniforme e pari a quello tipico della silice, il drogaggio avviene invece nel nucleo per

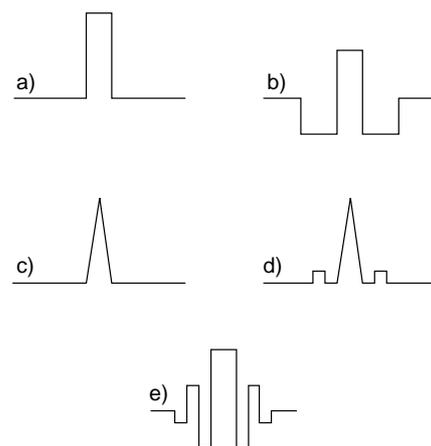


Figura 3 Classificazione delle fibre ottiche monomodali con riferimento al profilo dell'indice di rifrazione e la dispersione cromatica; a), b) fibre a dispersione standard, c), d) fibre a dispersione spostata, e) fibre a dispersione appiattita

umentarne l'indice di rifrazione. Oppure si agisce sul mantello drogandolo uniformemente con fluoro e si mantiene il nucleo al valore della silice (*pure silica core*). Il valore tipico del MFD è di 10 μm , mentre il valore del salto d'indice relativo è dell'ordine dello 0.3%. Per migliorare le prestazioni di queste fibre in terza finestra, soprattutto ai fini delle perdite per curvatura, si sono poi realizzate le fibre dette *reduced* caratterizzate da un MFD di 9.5 μm e da un salto d'indice di 0.37%. Le fibre a mantello depresso sono invece caratterizzate da un mantello che nella regione ad immediato contatto con il nucleo ha indice di rifrazione minore rispetto alla sua parte più esterna. Il MFD tipico è 9 μm ed il salto d'indice rispetto al mantello interno è dello 0.25%. I valori tipici di attenuazione sono 0.33-0.4 dB/km a 1310 nm e 0.18-0.2 dB/km a 1550 nm. Sia le fibre a mantello continuo che depresso fanno riferimento alla raccomandazione G.652 del CCITT.

La lunghezza d'onda a dispersione nulla della fibra può essere spostata in corrispondenza del minimo di attenuazione collocato a 1550 nm; questo può ottenersi riducendo il diametro del nucleo e, contemporaneamente, aumentando il salto d'indice. I profili d'indice più comunemente utilizzati per le fibre a dispersione spostata sono quelli riportati in Fig. 3c-3d. Il rialzo anulare nel profilo d'indice migliora il confinamento del modo e riduce di conseguenza le perdite per curvatura. Il MFD a 1550 nm è di 7-8.3 μm e l'attenuazione minima è dell'ordine di 0.22 dB/km, molto prossima quindi al valore tipico delle fibre standard.

Con un profilo d'indice particolarmente elaborato (Fig. 3e) si può ottenere una fibra la cui dispersione cromatica si mantiene limitata a valori molto bassi, minori di 2 ps/nm-km, in tutta la regione che va dalla seconda alla terza finestra (fibre a dispersione appiattita). Il vantaggio potenziale di una tale soluzione è evidente se solo si pensa alla possibilità di una trasmissione con tecnica WDM a larghissima banda ed alla capacità trasmissiva che ne deriva. A fronte di ciò, qualcosa resta ancora da fare in termini di sensibilità alle curvature e, soprattutto, di costo del prodotto finito; di conseguenza le fibre a dispersione appiattita hanno ancora qualche difficoltà a penetrare nel mercato.

3. Teoria della propagazione ottica in fibra

3.1 Modi di propagazione

Una fibra ottica ideale può essere descritta come una guida d'onda dielettrica cilindrica, caratterizzata da una distribuzione radiale della costante dielettrica $\epsilon(r)$, tale da essere sostanzialmente maggiore nella parte interna (nucleo) rispetto alla parte esterna (mantello), dove si può considerare uniforme.

3.1.1 Equazioni di campo

Per ricavare le espressioni del campo elettromagnetico in tale struttura si richiede la soluzione delle equazioni di Maxwell, tenendo conto delle condizioni al contorno che, nel caso della fibra, sono costituite dalla regolarità del campo, dal suo rapido decadimento nel mantello, e dalla continuità delle componenti del campo tra nucleo e mantello.

Le equazioni di Maxwell, in un mezzo dielettrico (privo di cariche e correnti) sono:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (1)$$

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t} \quad (2)$$

$$\vec{\nabla} \cdot \vec{D} = 0 \quad (3)$$

$$\vec{\nabla} \cdot \vec{B} = 0. \quad (4)$$

Inoltre si ha:

$$\vec{D} = \epsilon \vec{E}; \quad \vec{B} = \mu \vec{H}, \quad (5)$$

dove μ è la permeabilità magnetica, che per frequenze ottiche può considerarsi pari a quella del vuoto, ed ϵ è la costante dielettrica, che dipende dal punto. Si noti che se la costante dielettrica può assumersi costante, allora

per la (3) risulta $\vec{\nabla} \cdot \vec{E} = -\vec{E} \cdot \frac{\vec{\nabla} \epsilon}{\epsilon} \approx 0$, e questa relazione

verrà utilizzata per ricavare l'equazione d'onda che risolve il problema fin qui impostato. Qualunque onda che si propaga nella fibra può ottenersi dalla sovrapposizione di onde monocromatiche (essendo ω la pulsazione e t il tempo), del tipo:

$$\begin{cases} \vec{E}(r, \psi, z, t) = \vec{E}(r, \psi, z) e^{i\omega t} \\ \vec{H}(r, \psi, z, t) = \vec{H}(r, \psi, z) e^{i\omega t} \end{cases} \quad (6)$$

avendo scelto un sistema di coordinate polari cilindriche, coassiali alla fibra, considerata circolarmente simmetrica (Fig. 4). La coordinata radiale r è la distanza del punto generico dall'asse della fibra; la coordinata azimutale ψ rappresenta l'angolo formato, su una sezione ortogonale della fibra, dalla proiezione del raggio vettore del punto su tale sezione con un semiasse di riferimento; la coordinata longitudinale z rappresenta la distanza del punto da una sezione di riferimento della fibra, per esempio quella d'ingresso.

Sostituendo le (6) nelle (1) e (2) si ottiene, tenendo conto della (5)

$$\vec{\nabla} \times \vec{E} = -i\omega \mu \vec{H} \quad (7)$$

$$\vec{\nabla} \times \vec{H} = i\omega \epsilon \vec{E} \quad (8)$$

Si può verificare che le equazioni di Maxwell restanti (3) e (4) sono implicate dalle (7) e (8). Essendo la struttura cilindrica invariante per traslazioni lungo l'asse z e per rotazioni attorno allo stesso asse, ne consegue che la dipendenza del campo elettromagnetico da z e da ψ può esprimersi come:

$$\begin{cases} \vec{E}(r, \psi, z) = \vec{E}(r)e^{-i(v\psi + \beta z)} \\ \vec{H}(r, \psi, z) = \vec{H}(r)e^{-i(v\psi + \beta z)} \end{cases} \quad (9)$$

essendo v un numero intero (affinché il campo sia monodromo rispetto a ψ), detto numero quantico azimutale, tale che v/r rappresenta la costante di propagazione azimutale dell'onda elettromagnetica, e β è la costante di propagazione longitudinale che, come si vedrà, può assumere solo particolari valori (autovalori).

Introducendo il numero d'onda del mezzo, che si può interpretare come il modulo della costante di propagazione:

$$k = \omega\sqrt{\epsilon\mu} \quad (10)$$

e sostituendo le (9) nelle (7)(8) e svolgendo i relativi calcoli, è possibile esprimere le componenti trasverse del campo in funzione di quelle longitudinali:

$$E_r = \frac{-i}{k^2 - \beta^2} \left(\beta \frac{dE_z}{dr} - i\omega\mu \frac{v}{r} H_z \right) \quad (11)$$

$$E_\psi = \frac{-i}{k^2 - \beta^2} \left(-i\beta \frac{v}{r} E_z - \omega\mu \frac{dH_z}{dr} \right) \quad (12)$$

$$H_r = \frac{-i}{k^2 - \beta^2} \left(\beta \frac{dH_z}{dr} + i\omega\epsilon \frac{v}{r} E_z \right) \quad (11a)$$

$$H_\psi = \frac{-i}{k^2 - \beta^2} \left(-i\beta \frac{v}{r} H_z + \omega\epsilon \frac{dE_z}{dr} \right) \quad (12a)$$

Gli sviluppi fin qui presentati hanno consentito quindi di esprimere tutte le componenti di campo trasverse in funzione di solo quelle longitudinali. Applicando

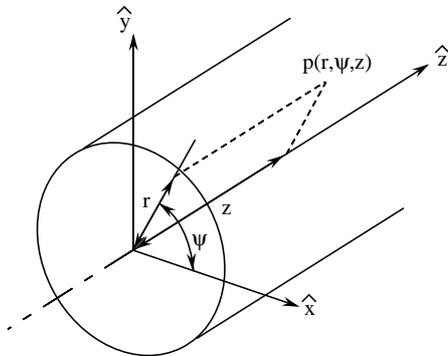


Figura 4 Sistema di riferimento in coordinate cilindriche

l'operatore rotore ad entrambi i membri dell'equazione (7), e sfruttando la nota relazione valida per qualunque vettore \vec{V}

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{V}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{V}) - \nabla^2 \vec{V}$$

grazie anche alle (8) e (10) si ottiene:

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} = k^2 \vec{E} \quad (13)$$

La condizione $\vec{\nabla} \cdot \vec{E} \approx 0$ consente di passare da una formulazione vettoriale del problema ad una scalare valida per ciascuna delle componenti; in questo caso l'interesse è rivolto alla componente longitudinale $E_z = E_z(r)e^{-i(v\psi + \beta z)}$, quindi la (13) può riscriversi come segue

$$\nabla^2 E_z + k^2 E_z = 0 \quad (14)$$

Poiché in coordinate cilindriche si ha

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \psi^2} + \frac{\partial^2}{\partial z^2}$$

e ricordando la (9), segue che la (14) fornisce la seguente equazione differenziale per $E_z(r)$:

$$\frac{d^2 E_z}{dr^2} + \frac{1}{r} \frac{dE_z}{dr} + \left(k^2 - \beta^2 - \frac{v^2}{r^2} \right) E_z = 0 \quad (15)$$

Ed un'identica equazione è valida anche per $H_z(r)$.

3.1.2 Equazione caratteristica

Alle equazioni proposte per $E_z(r)$ ed $H_z(r)$ devono essere imposte le condizioni al contorno. Sia $F_0(r)$ la soluzione della (15) nel nucleo della fibra ($r < a$; con a raggio del nucleo), regolare (non divergente) nell'origine ($r=0$), e sia $F_1(r)$ la soluzione nel mantello della fibra ($r > a$), regolare all'infinito, cioè che per $r \gg a$ mostri un sostanziale decadimento esponenziale. Allora $E_z(r)$ e $H_z(r)$ possono così esprimersi:

$$E_z = \begin{cases} E_0 F_0(r) & \text{per } r < a \\ E_1 F_1(r) & \text{per } r > a \end{cases} \quad (16)$$

$$H_z = \begin{cases} H_0 F_0(r) & \text{per } r < a \\ H_1 F_1(r) & \text{per } r > a \end{cases} \quad (17)$$

Queste espressioni sono fisicamente accettabili e le costanti E_0, E_1, H_0, H_1 , sono determinate imponendo la continuità delle componenti tangenziali dei campi (E_z, H_z, E_ψ, H_ψ) all'interfaccia tra nucleo e mantello ($r=a$). In realtà, solo tre di queste costanti possono essere determinate in funzione della quarta che si può eventualmente trovare partendo dalle condizioni di eccitazione elettromagnetica della fibra stessa da parte di

una sorgente ottica⁽²⁾. In definitiva si può scrivere un sistema omogeneo di quattro equazioni in quattro incognite, ed affinché la soluzione sia non banale ($E_0=H_0=0$) è necessario che il determinante del sistema si annulli. Svolgendo i calcoli, si ottiene (è sottinteso $r=a$):

$$\frac{v^2}{a^2} \beta^2 F_0^2 \left(\frac{1}{k_0^2 - \beta^2} - \frac{1}{k_1^2 - \beta^2} \right)^2 = k_0^2 C_1 C_2 \quad (18)$$

ove

$$\begin{cases} C_1 = \frac{\frac{dF_0}{dr} - \frac{F_0}{k_0^2 - \beta^2}}{\frac{dF_1}{dr} - \frac{F_1}{k_1^2 - \beta^2}} \\ C_2 = \frac{\frac{dF_0}{dr} - \frac{F_0}{k_0^2 - \beta^2}}{\frac{dF_1}{dr} - \frac{F_1}{k_1^2 - \beta^2}} \frac{\epsilon_1}{\epsilon_0} \end{cases}$$

La (18), detta **equazione caratteristica**, determina i valori della costante di propagazione longitudinale β per i quali si hanno nella fibra configurazioni di campo elettromagnetico in grado di propagarsi effettivamente: queste configurazioni sono dette **modi**.

Per ogni valore di v l'equazione caratteristica ha per soluzioni due serie di modi denominati EH_{vd} e HE_{vd} , dove il pedice d numera ognuna delle soluzioni fisicamente valide di ciascuna serie.

Il caso $v=0$ va trattato separatamente, in quanto le due serie di modi hanno rispettivamente, $E_z=0$ o $H_z=0$; si tratta dunque di modi trasversi elettrici (TE_{0d}) per i quali si annullano anche le componenti E_r e H_ψ , o trasversi magnetici (TM_{0d}) che hanno nulle anche le componenti E_ψ e H_r . Si rammenta che i modi TE e TM sono i soli che si propagano nelle guide d'onda metalliche. Volendo mantenere la notazione EH e HE, si assume convenzionalmente che: $TE_{0d}=HE_{0d}$ e $TM_{0d}=EH_{0d}$.

Nel caso di $v \neq 0$, invece, è necessario che entrambe le componenti di campo siano non nulle. Si possono introdurre notevoli semplificazioni assumendo che la fibra sia debolmente guidante, cioè si abbia solo una piccola variazione della costante dielettrica del nucleo rispetto a quella del mantello. Giacché β dev'essere sempre compresa tra i valori massimi e minimi del numero d'onda $k(r) = \omega \sqrt{\epsilon(r)\mu}$, ne deriva che $\beta \approx k$, o meglio:

$$|\beta - k| \ll k. \quad (19)$$

Inserendo questa approssimazione nell'equazione caratteristica (18) (con $\epsilon_1/\epsilon_0 \approx 1$) e ponendo $D_0 = k_0^2 - \beta^2$ e $D_1 = k_1^2 - \beta^2$, si ottiene:

$$\frac{v}{a} F_0 \left(\frac{1}{D_0} - \frac{1}{D_1} \right) = \mp \left(\frac{dF_0}{dr} - \frac{F_0}{D_0} \frac{dF_1}{dr} \right) \quad (20)$$

Questa **forma semplificata dell'equazione caratteristica** evidenzia le due serie di soluzioni corrispondenti ai modi EH e HE con la presenza del doppio segno (\pm). In questa equazione e nelle seguenti, il segno superiore contraddistinguerà i modi EH e quello inferiore i modi HE.

Con opportune sostituzioni, si può dimostrare che per $v \neq 0$

$$\frac{E_0}{H_0} = \mp i \sqrt{\frac{\mu}{\epsilon}} \quad (21)$$

Questa equazione consente di stabilire la seguente relazione tra E_z e H_z :

$$H_z = \mp i \sqrt{\frac{\epsilon}{\mu}} E_z \quad (22)$$

Quest'ultima spiega perché, come già detto, per $v \neq 0$ le componenti E_z e H_z devono essere entrambe non nulle (altrimenti tutto il campo risulterebbe nullo); in questo senso i modi EH e HE sono detti ibridi (per v diverso da zero) in quanto non possono essere né TE, né TM.

Grazie all'ipotesi di debole guidanza ($\beta \approx k$) si possono ricavare le seguenti interessanti relazioni incrociate:

$$H_r = -\sqrt{\frac{\epsilon}{\mu}} E_\psi \quad (23)$$

$$H_\psi = \sqrt{\frac{\epsilon}{\mu}} E_r \quad (24)$$

Le componenti longitudinali dei campi sono molto piccole rispetto a quelle trasverse, quindi le (23) e (24) permettono di scrivere:

$$\vec{H} \approx \sqrt{\frac{\epsilon}{\mu}} \hat{z} \times \vec{E} \quad (25)$$

E' importante osservare che una analoga relazione è valida per un'onda piana che si propaghi nella fibra nella direzione z : cioè i vettori campo elettrico e campo magnetico sono sostanzialmente ortogonali tra di loro e ortogonali alla direzione di propagazione.

L'equazione caratteristica comporta ulteriori interessanti relazioni tra le componenti dei campi, ad esempio

$$E_r = \mp i E_\psi \quad (26)$$

$$H_r = \mp i H_\psi \quad (27)$$

(2) Non si dimentichi che il sistema di equazioni di partenza, ovvero le equazioni di Maxwell (1)-(4) per la fibra, è *omogeneo*, poiché applicato ad una regione di spazio priva di cariche e correnti.

Le equazioni (23), (24), (26) e (27) mostrano che le componenti trasverse del campo, nella ipotesi di debole guida, soddisfano la medesima equazione differenziale. Tale equazione differenziale può ricavarsi dalle (11), (15) e (22) e si ottiene:

$$\frac{d^2 E_r}{dr^2} + \frac{1}{r} \frac{dE_r}{dr} + \left(k^2 - \beta^2 - \frac{(\nu \pm 1)^2}{r^2} \right) E_r = 0 \quad (28)$$

L'equazione caratteristica legata a questa equazione differenziale risulta più semplice della (20) valida per le componenti longitudinali. Infatti, sia $G_0(r)$ la soluzione della (28) nel nucleo della fibra, regolare nell'origine, e sia $G_1(r)$ la soluzione nel mantello, regolare all'infinito. Allora può scriversi (A_0 e A_1 sono delle costanti):

$$E_r = \begin{cases} A_0 G_0 & \text{per } r < a \\ A_1 G_1 & \text{per } r > a \end{cases} \quad (29)$$

L'equazione caratteristica si ottiene imponendo la continuità di E_r e di E_z (in virtù delle relazioni precedenti ciò è sufficiente per garantire la continuità delle altre componenti); cioè (sottintendendo $r = a$):

$$A_0 G_0 = A_1 G_1 \quad (30)$$

$$A_0 \left(\frac{dG_0}{dr} \pm \frac{\nu \pm 1}{a} G_0 \right) = A_1 \left(\frac{dG_1}{dr} \pm \frac{\nu \pm 1}{a} G_1 \right) \quad (31)$$

Anche questo è un sistema omogeneo e quindi per avere soluzioni A_0 e A_1 significative è necessario che il determinante sia nullo, ciò porta alla seguente equazione caratteristica, di notevole semplicità:

$$\frac{1}{G_0} \frac{dG_0}{dr} = \frac{1}{G_1} \frac{dG_1}{dr} \quad (32)$$

Ciò significa che ciascuna componente trasversale del campo deve essere continua all'interfaccia nucleo-mantello, insieme alla sua derivata, e l'equazione caratteristica si ottiene imponendo la continuità delle derivate logaritmiche delle soluzioni regolari nel nucleo e nel mantello dell'equazione (28).

3.1.3 Linee di campo dei modi

Per identificare i modi della fibra è necessario studiare l'andamento delle linee di flusso dei vettori campo elettrico e campo magnetico. A tal fine si considera l'andamento delle componenti trasverse del campo elettrico in funzione delle coordinate trasverse. Dalle (9) e (26) si ottiene:

$$E_r(r, \psi) = E_r(r) e^{-i\nu\psi} = E_r(r) (\cos \nu\psi - i \sin \nu\psi) \quad (33)$$

$$E_\psi(r, \psi) = \pm i E_r(r) e^{-i\nu\psi} = E_r(r) (\pm \sin \nu\psi \pm i \cos \nu\psi) \quad (34)$$

La parte reale e la parte immaginaria di queste componenti rappresentano due regimi elettromagnetici indipendenti, del tutto simili tra di loro, e che differiscono solo per il diverso orientamento delle linee di flusso dei campi. Se ne può quindi considerare la sola parte immaginaria, ottenendo:

$$\frac{E_r}{E_\psi} = \mp \text{tg } \nu\psi \quad (35)$$

Per definizione, le linee di campo sono tali che il vettore del campo elettrico trasverso è tangente ad esse in ogni punto (Fig. 5); ne consegue che

$$\frac{1}{r} \frac{dr}{d\psi} = \frac{E_r}{E_\psi},$$

avendo indicato con $r(\psi)$ la funzione che descrive la generica linea di flusso. Quindi $r(\psi)$ soddisfa la seguente equazione differenziale (v. (35)):

$$\frac{1}{r} \frac{dr}{d\psi} = \mp \text{tg } \nu\psi, \quad (36)$$

integrando la quale si ha:

$$r^{\mp \nu} \cos \nu\psi = \text{cost} \quad (37)$$

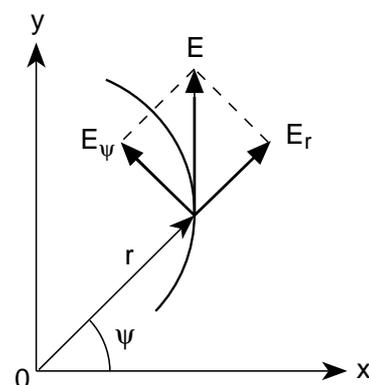


Figura 5 Relazione tra le componenti trasversali del campo

Nella Fig. 6 sono riportati gli andamenti delle linee di flusso del campo elettrico per i modi con v più basso. Si noti che le linee hanno una periodicità azimutale di π/v ; inoltre è evidente la differenza tra i modi HE ed EH: nei primi le linee di flusso si chiudono solo all'infinito, mentre nei secondi le linee sono sempre chiuse su se stesse. E' anche evidente che solo i modi HE_{1d} sono linearmente polarizzati, come può anche dedursi dalla (37) (che per questi modi diventa: $r \cos(\psi) = cost$) che descrive rette parallele all'asse cartesiano verticale.

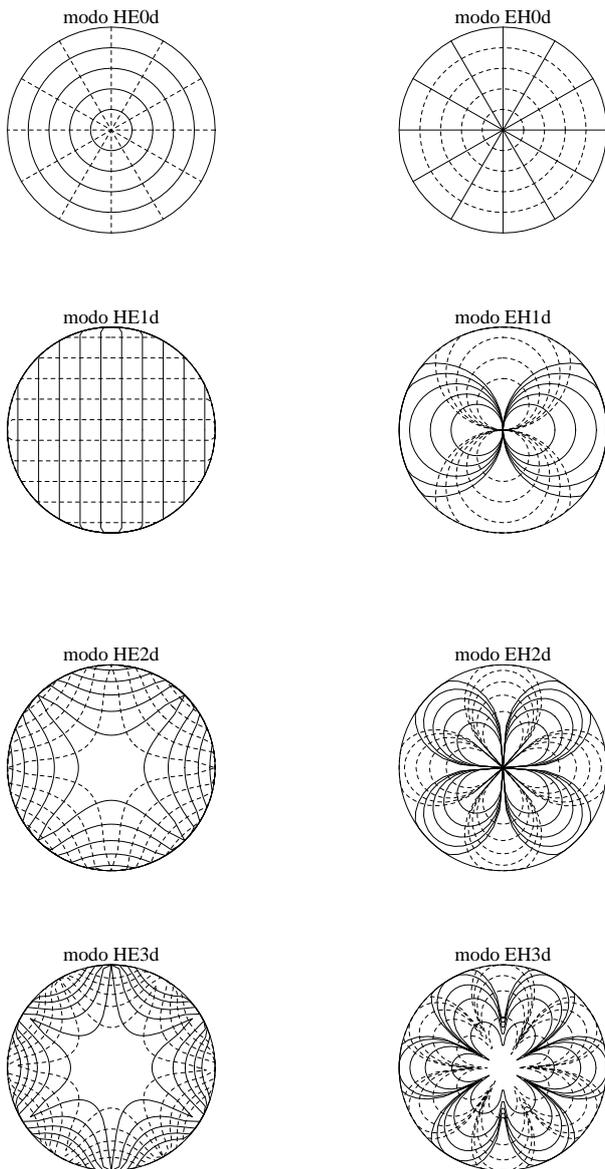


Figura 6 Linee di campo elettrico (linee continue) e di campo magnetico (linee tratteggiate) per i modi HE_{vd} e EH_{vd} , per alcuni valori di v

3.1.4 Modi Linearmente Polarizzati (LP)

Dalla (28) si riconosce facilmente che i modi $HE_{v+1,d}$ ed $EH_{v-1,d}$ hanno gli stessi autovalori β e le stesse caratteristiche di dispersione. Questa degenerazione è valida nell'ipotesi di debole guida e implica che i due modi sono sempre associati: ha senso quindi considerare un modo composto come somma di tali modi.

Detto $E_r^{(v)}(r)$ il campo $E_r(r)$ soluzione della (28) per i modi $HE_{v+1,d}$ o $EH_{v-1,d}$, la (28) medesima può essere riscritta come:

$$\frac{d^2 E_r^{(v)}}{dr^2} + \frac{1}{r} \frac{dE_r^{(v)}}{dr} + \left(k^2 - \beta^2 - \frac{v^2}{r^2} \right) E_r^{(v)} = 0 \quad (38)$$

e componendo il contributo delle parti reali delle componenti E_r (eq.(33)) dei due modi $HE_{v+1,d}$ o $EH_{v-1,d}$ e operando similmente sulle componenti E_ψ , H_r e H_ψ (eq.(34), (23) e (24)) si ha:

$$E_r(r, \psi) = E_r^{(v)} \cos v\psi \cos \psi \quad (39)$$

$$E_\psi(r, \psi) = -E_r^{(v)} \cos v\psi \sin \psi \quad (40)$$

$$H_r(r, \psi) = \sqrt{\frac{\epsilon}{\mu}} E_r^{(v)} \cos v\psi \sin \psi \quad (41)$$

$$H_\psi(r, \psi) = \sqrt{\frac{\epsilon}{\mu}} E_r^{(v)} \cos v\psi \cos \psi \quad (42)$$

Un risultato simile si ottiene componendo le parti immaginarie delle componenti.

Studiando, come fatto precedentemente, il rapporto E_r/E_ψ si osserva che tutti questi modi composti sono linearmente polarizzati: essi vengono dunque denominati modi $LP_{v,d}$ (equivalenti alla combinazione $HE_{v+1,d} + EH_{v-1,d}$). Tutti i modi composti $LP_{v,d}$ sono sintesi di due modi semplici eccetto i modi LP_{0d} che coincidono con i soli modi HE_{1d} e sono dunque semplici, e i modi LP_{1d} che sono invece l'insieme di tre modi semplici: EH_{0d} (TM_{0d}), HE_{0d} (TE_{0d}) e HE_{2d} e che comunque si possono comporre a due a due ($HE_{2d} - TE_{0d}$ o $HE_{2d} - TM_{0d}$).

3.1.5 Fibre a gradino

L'equazione (28) è un risultato molto importante della teoria svolta finora: essa con l'equazione caratteristica (32) permette di determinare le condizioni di propagazione di ciascun modo. Bisogna però precisare che in genere la (28) non è di facile soluzione, in particolare quando la distribuzione di $\epsilon(r)$ non è uniforme nel nucleo. E' comunque interessante esporre la

soluzione per fibre con nucleo uniforme (fibre a gradino di indice di rifrazione), per illustrare le caratteristiche più salienti dei vari modi.

La distribuzione radiale (profilo) di indice di rifrazione $n(r)$ lega la distribuzione della costante dielettrica nelle fibre $\epsilon(r)$ alla costante dielettrica nel vuoto ϵ_v come segue:

$$\epsilon(r) = n^2(r)\epsilon_v \quad (43)$$

D'altra parte dalla (10) si vede che

$$k(r) = n(r)k_v \quad (44)$$

dove k_v è il numero d'onda nel vuoto, esprimibile in funzione della lunghezza d'onda e della velocità c della luce nel vuoto come:

$$k_v = \frac{2\pi}{\lambda} = \frac{\omega}{c} = \omega\sqrt{\epsilon_v\mu} \quad (45)$$

Una fibra a gradino di indice di rifrazione è caratterizzata dall'aver il nucleo con indice di rifrazione uniforme n_0 , ed un mantello con indice anch'esso uniforme $n_1 < n_0$.

La soluzione nel nucleo $G_0(r)$ della (38) per i modi $LP_{v,d}$, regolare nell'origine, è in questo caso data da una funzione di Bessel di prima specie:

$$G_0(r) = J_v(\beta_{r0}r) \quad (46)$$

dove si è introdotta la costante di propagazione trasversa nel nucleo β_{r0}

$$\beta_{r0} = \sqrt{k_v^2 n_0^2 - \beta^2} \quad (47)$$

La soluzione nel mantello $G_1(r)$ dell'equazione (28), regolare all'infinito, è invece data da una funzione di Bessel modificata di seconda specie:

$$G_1(r) = K_v(\beta_{r1}r) \quad (48)$$

dove la costante di propagazione trasversa nel mantello β_{r1} è definita (in modulo) come:

$$\beta_{r1} = \sqrt{\beta^2 - k_v^2 n_1^2} \quad (49)$$

Tale soluzione comporta un rapido decadimento di tipo esponenziale se la costante di propagazione trasversa nel mantello β_{r1} è puramente reale.

Si osservi che le costanti β_{r0} e β_{r1} non sono indipendenti, infatti si introduca la **frequenza normalizzata V**

$$V = k_v a \sqrt{n_0^2 - n_1^2} \quad (50)$$

pari al prodotto del raggio del nucleo della fibra per l'apertura numerica massima normalizzato tramite il numero d'onda nel vuoto k_v . La frequenza normalizzata V è una quantità fissata conoscendo la fibra in esame e la lunghezza d'onda della radiazione che si vuole propagare in essa. Dalle (47), (49) e (50) si trova che:

$$a^2 \beta_{r0}^2 + a^2 \beta_{r1}^2 = V^2 \quad (51)$$

Introducendo le quantità adimensionali: $u = a\beta_{r0}$, costante di propagazione trasversa normalizzata nel nucleo, e $w = a\beta_{r1}$, modulo della costante di propagazione trasversa normalizzata nel mantello, la (51) può risciversi:

$$u^2 + w^2 = V^2 \quad (52)$$

Invece le (46) e (48) diventano:

$$G_0(r) = J_v\left(u \frac{r}{a}\right); \quad G_1(r) = K_v\left(w \frac{r}{a}\right) \quad (53)$$

L'equazione caratteristica (32) ottenuta imponendo la continuità delle derivate logaritmiche di queste due soluzioni per $r = a$, diviene in questo caso:

$$\begin{cases} J_{v+1}(u)K_{v-1}(w) + J_{v-1}(u)K_{v+1}(w) = 0 \\ \text{per i modi } LP_{v,d} \text{ con } v \neq 0 \\ uJ_1(u)K_0(w) - wJ_0(u)K_1(w) = 0 \\ \text{per i modi } LP_{0,d} \end{cases} \quad (54)$$

E' questa un'equazione trascendente che per ogni V definisce implicitamente una serie di funzioni $u(w)$ che, mediante la (52), permettono di ottenere i valori possibili di u e w e quindi di β (autovalori) per avere modi guidati. Risolvendo per via numerica il sistema costituito dalle (52) e (54) al variare della frequenza normalizzata V si ottengono diverse funzioni $\beta(V)$, ciascuna per ogni modo $LP_{v,d}$. La Fig. 7 riporta, per alcuni dei modi della fibra e per valori della frequenza normalizzata fino a $V=10$, l'andamento della costante di propagazione normalizzata, b , definita come:

$$b = \frac{\beta^2 - k_v^2 n_1^2}{k_v^2 (n_0^2 - n_1^2)}$$

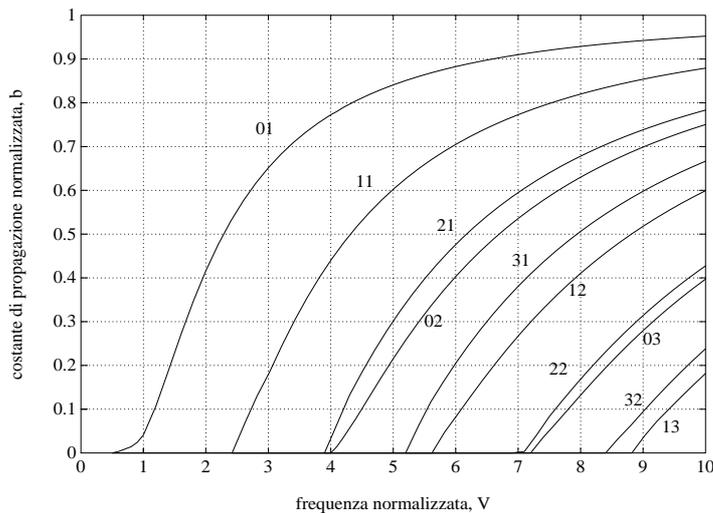


Figura 7 Andamento della costante di propagazione normalizzata in funzione della frequenza normalizzata per alcuni dei modi (di ordine più basso) di una fibra a gradino; per ogni curva è indicata la coppia vd che individua il relativo modo LP_{vd}

3.1.6 Taglio dei modi

La fig. 7 illustra un fenomeno di validità generale. Si nota infatti che per ogni modo LP_{vd} esiste, in genere, un valore particolare di V : V_{vd}^T al di sotto del quale non è possibile avere soluzioni fisicamente accettabili corrispondenti a modi guidati. Questa soglia corrisponde a $\beta = k_v n_1$ per ciascun modo, e cioè a $\beta_{r1} = 0$ e $w = 0$. In realtà, per $V < V_{vd}^T$ è ancora possibile avere soluzioni dell'equazione caratteristica, ma in questo caso β_{r1} ha una parte immaginaria non nulla, che si traduce (v.(49)) in una corrispondente parte immaginaria in β . Questo indica che in tali condizioni il modo non è più guidato, ma si irradia fuori della fibra, e quindi non è di interesse. Per $V < V_{vd}^T$ si dice che il modo LP_{vd} è tagliato dalla guida e V_{vd}^T è detta frequenza normalizzata di taglio; ad essa corrisponde una lunghezza d'onda di taglio

$$\lambda_{vd}^T = \frac{2\pi a}{V_{vd}^T} \sqrt{n_0^2 - n_1^2} \quad (55)$$

Questo valore di lunghezza d'onda è tale che il modo LP_{vd} si può propagare in una data fibra solo per $\lambda < \lambda_{vd}^T$. La condizione di taglio è: $w = 0$, e dalla (52) si vede che ciò implica $u = V (= V_{vd}^T)$. Per ottenere i valori di V_{vd}^T in fibre a gradino, occorre dunque valutare il limite per $w \rightarrow 0$ nella (54). Ricordando che $K_v(w) \approx w^{-v}$ per $v \neq 0$ e $w \rightarrow 0$ e $K_0(w) \approx \ln(w)$ per $w \rightarrow 0$, si vede facilmente che, indicando con j_{vd} il d -esimo zero della funzione $J_v(u)$, i valori V_{vd}^T sono soluzioni delle seguenti equazioni

$$\begin{cases} v \neq 0 & J_{v-1}(V) = 0 \Rightarrow V_{vd}^T = j_{v-1,d} \text{ per } d \geq 1 \\ v = 0 & VJ_1(V) = 0 \Rightarrow V_{01}^T = 0; V_{0d}^T = j_{1,d-1} \text{ per } d \geq 2 \end{cases}$$

Esiste quindi un solo modo che può propagarsi in ogni condizione: il modo LP_{01} o HE_{11} che ha una V di taglio nulla, e che è detto modo fondamentale. Per $V < V_{01}^T = 2.405$ tutti gli altri modi sono irradati, per cui la fibra è detta **fibra monomodo**. Altrimenti, se più modi possono essere guidati, la fibra è detta **fibra multimodo**.

Il campo elettromagnetico trasverso del modo LP_{vd} nel mantello è descritto, per ogni profilo, dalla funzione $K_v(wr/a)$; questa funzione, per $r \gg a$, va come un esponenziale decrescente: $\exp(-wr/a)$, il che assicura che il campo sia ben confinato nella fibra, e spiega anche il meccanismo del taglio. Infatti, quando $w \rightarrow 0$ il campo tende ad aprirsi su tutto lo spazio; viceversa, per $w \gg 1$ il campo si annulla immediatamente nel mantello e risulta quindi confinato nel solo nucleo della fibra ed il modo si dice in condizioni lontane dal taglio. In tali condizioni u tende ad un valore finito, in fibre a gradino, ottenibile dalla (54) per $w \rightarrow \infty$.

$$w \rightarrow \infty: J_v(u) = 0 \Rightarrow u = j_{vd} \quad (56)$$

per cui (v.(52)) in questo caso è: $V \approx w$ e, come si vede dalla Fig. 7, $\beta \rightarrow k_v n_0$.

Nel nucleo il campo del modo LP_{vd} è descritto dalla funzione $J_v(ur/a)$. Tale funzione di Bessel è simile ad una funzione armonica modulata con un fattore $r^{-1/2}$ di metrica, che per $r \rightarrow 0$ va come r^v . Per cui soltanto la J_0 vale 1 nell'origine, mentre le altre sono infinitesime. Quindi il campo all'interno del nucleo ha un andamento pseudoarmonico, cambiando di segno $d-1$ volte (come è suggerito dalla (56)): ciò indica che il suo modulo quadro ha d massimi relativi lungo l'asse r . Inoltre, soltanto i modi LP_{0d} hanno campo non nullo sull'asse.

Quanto detto finora, valido per le fibre con profilo a gradino, resta essenzialmente valido anche per le fibre con qualunque profilo di indice di rifrazione. Per queste ultime, infatti, cambia la forma della soluzione $G_0(r)$ nel nucleo, ma rimane valida la soluzione $G_1(r)$ di eq. (48) nel mantello, in quanto questo si suppone uniforme in ogni caso. Rimangono anche valide le definizioni (47), (49) e (50) di β_{n0} , β_{r1} e V , assumendo però per n_0 l'indice di rifrazione massimo nel nucleo; così pure restano valide le (51), e (52). Ovviamente cambia la forma dell'equazione caratteristica (54), ma rimane valido il concetto di taglio di ciascun modo, anche se i valori V_{vd}^T dipendono dal profilo.

Il fatto che il campo ha nel nucleo un andamento pseudoarmonico, con d massimi nella sua intensità, e nel mantello un sostanziale decadimento esponenziale, è di validità generale. Si può infine mostrare che il numero totale di modi guidati da una fibra è proporzionale a V^2 (per $V \gg 1$), ed il coefficiente di proporzionalità dipende dal tipo di profilo. Per fibre a gradino il numero dei modi guidati è circa $V^2/2$.

3.2 Distorsione dell'impulso ottico

Poiché la dipendenza da ω delle caratteristiche del modo è lentamente variabile rispetto agli altri parametri in gioco, il campo totale della fibra nel punto z può scriversi:

$$\vec{E}(r, \psi, z, t) = \sum_{vd} b_{vd}(z, t) \vec{E}_{vd}(r, \psi) \quad (57)$$

dove

$$b_{vd}(z, t) = \int_0^\infty a_{vd}(\omega) e^{i(\omega t - \beta_{vd}(\omega)z)} d\omega \quad (58)$$

in cui $\beta_{vd}(\omega)$ è la costante di propagazione del modo LP_{vd} , mentre gli $a_{vd}(\omega)$ sono coefficienti di Fourier determinabili dal campo all'ingresso della fibra $E_0(r, \psi, t)$ attraverso la seguente equazione:

$$a_{vd}(\omega) = \frac{1}{4\pi W_{vd}} \int_{-\infty}^\infty dt e^{-i\omega t} \int_0^\infty r dr \int_0^{2\pi} d\psi W_{vd}(r, \psi, t)$$

essendo:

$$\begin{cases} W_{vd}(r, \psi, t) = \vec{E}_0(r, \psi, t) \times \vec{H}_{vd}(r, \psi, t) \cdot \hat{z} \\ W_{vd} = \int_0^\infty r dr \int_0^{2\pi} d\psi \vec{E}_{vd} \times \vec{H}_{vd} \cdot \hat{z} \end{cases}$$

Le (57) e (58) permettono di seguire l'evoluzione spazio-temporale del campo. A partire dalle (57) e (58) e sfruttando il fatto che la larghezza spettrale del campo-

sorgente E_0 è solitamente molto minore della pulsazione centrale ω_0 , si può sviluppare $\beta_{vd}(\omega)$ come segue:

$$\beta_{vd}(\omega) = \beta_{vd}(\omega_0) + \Delta\omega \tau_{vd}(\omega_0) + \frac{\Delta\omega^2}{2} A_{vd}(\omega_0) \quad (59)$$

ove si è posto $\Delta\omega = \omega - \omega_0$; nella (59) sono stati inoltre introdotti il **ritardo unitario di gruppo** τ_{vd} , cioè il tempo che l'impulso impiega a coprire l'unità di lunghezza della fibra su quel dato modo, e la quantità A_{vd} , così definiti:

$$\tau_{vd}(\omega) = \frac{d\beta}{d\omega}$$

$$A_{vd}(\omega) = \frac{d^2\beta}{d\omega^2}$$

3.2.1 Dispersione modale

Sostituendo lo sviluppo (59) fermato al primo ordine, nella (58), si ottiene:

$$b_{vd}(z, t) = e^{i(\omega_0 \tau_{vd}(\omega_0) - \beta_{vd}(\omega_0))z} b_{vd}(0, t - z\tau_{vd}) \quad (60)$$

che riportata nella (57) permette di ottenere:

$$E(r, \psi, z, t) = \sum_{vd} e^{i\phi_{vd}(\omega_0)z} b_{vd}(0, t - z\tau_{vd}) E_{vd}(r, \psi)$$

ove $\phi_{vd}(\omega) = \omega\tau_{vd} - \beta_{vd}(\omega)$.

Questa equazione mostra che il campo di ciascun modo si propaga con diverse velocità di gruppo. Ciò significa che se un impulso ottico viene lanciato all'ingresso della fibra, esso si ripartisce tra i vari modi guidati i quali, viaggiando lungo la fibra con ritardi diversi, causano una dispersione di tale impulso. La larghezza (intesa come scarto quadratico medio, s.q.m.) dell'impulso in z , $\sigma_t(z)$, può facilmente correlarsi alla dispersione (s.q.m.) dei ritardi di gruppo σ_{vd} tra i vari modi, secondo la relazione:

$$\sigma_t(z) = z\sigma_{vd} \quad (61)$$

Questo genere di dispersione è detto intermodale, o semplicemente dispersione modale, cioè tra i vari modi, e l'allargamento che essa produce sarebbe funzione lineare della lunghezza se non ci fosse scambio di energia tra i vari modi. Invece ogni disomogeneità lungo la fibra (giunti, connettori, variazioni del diametro del nucleo) determina un rimescolamento dei modi, e si può assumere che la dispersione cresca con la radice quadrata della lunghezza.

3.2.2 Dispersione cromatica

Nell'equazione (60) si è trascurata la derivata seconda di β , cioè il termine A_{vd} , e per questo motivo, in tale equazione, i singoli impulsi portati da ciascun modo appaiono come una copia non distorta, ma semplicemente ritardata, dell'impulso di partenza. In realtà è presente una certa dispersione intramodale, cioè nell'ambito del medesimo modo, ma essa risulta generalmente trascurabile rispetto a quella intermodale.

Nelle fibre monomodo è presente solo la dispersione intramodale, perché ovviamente manca quella intermodale, detta anche dispersione cromatica in quanto il ritardo di gruppo del singolo modo dipende dalla pulsazione ω , o dalla lunghezza d'onda λ ; infatti il mezzo di cui la fibra è composta è dispersivo. Inoltre le proprietà stesse di propagazione del modo, come si è visto, dipendono da ω , e ciò causa un contributo di guida. La larghezza (s.q.m.) dell'impulso ottico portato dal singolo modo LP_{vd} in z ($\sigma_{rvd}(z)$) è tale che:

$$\sigma_{rvd}^2(z) = \sigma_{rvd}^2(0) + z^2 \sigma_{\omega}^2 A_{vd}^2(\omega_0) \quad (62)$$

dove σ_{ω} è la larghezza spettrale (s.q.m.) del campo-sorgente in ω . Introducendo la larghezza spettrale (s.q.m.) del campo-sorgente in λ (σ_{λ}) e la dispersione cromatica D_{vd} , così definite:

$$\sigma_{\lambda} = -\frac{\lambda^2}{2\pi c} \sigma_{\omega}$$

$$D_{vd} = \frac{d\tau_{vd}}{d\lambda} = -\frac{2\pi c}{\lambda^2} \frac{d^2\beta}{d\omega^2} = -\frac{2\pi c}{\lambda^2} A_{vd}$$

ne segue che la (62) può riscriversi:

$$\sigma_{rvd}^2(z) = \sigma_{rvd}^2(0) + z^2 \sigma_{\lambda}^2 D_{vd}^2(\omega_0) \quad (63)$$

La dispersione intramodale, particolarmente se $\sigma_{rvd}(0)$ è trascurabile rispetto agli altri termini della (63), è funzione lineare della lunghezza della fibra ed è anche funzione lineare della larghezza spettrale del campo-sorgente, e della dispersione cromatica. Se in una fibra multimodo la dispersione intermodale è comparabile con quella intramodale, allora la dispersione totale si otterrà sommando quadraticamente le larghezze $\sigma_r(z)$ e $\sigma_{rvd}(z)$, di eq. (61) e (63), scegliendo per quest'ultima un valor medio tra i vari modi. Questo può avvenire solo se la dispersione intermodale è estremamente bassa, cioè in fibre in cui i ritardi di gruppo τ_{vd} tra i vari modi sono molto simili.

Per rimuovere gli effetti della dispersione modale, una maniera drastica è quella di utilizzare fibre monomodo, nelle quali l'unica dispersione presente è quella intramodale che, come già visto, presenta un contributo di guida e uno di materiale. Per materiali a base di silice (quelli comunemente usati), il contributo di materiale è preponderante rispetto a quello di guida per lunghezze d'onda $\lambda < 1100$ nm: per $\lambda \approx 1300$ nm è possibile compensare i due contributi annullando anche la dispersione intramodale. In tal caso, sono possibili enormi capacità di trasmissione di informazione, limitate solo dalla dispersione del terzo ordine che è veramente modesta.

4. Conclusioni

In questo articolo sono state presentate le fibre ottiche per TLC esprimendo, in termini di prospettive di impiego, l'importanza sempre maggiore che esse rivestiranno nel sistema di comunicazione del futuro. Si è pure ritenuto utile riassumere l'attuale articolazione tipologica delle fibre, con riferimento ai parametri più significativi ed alle prestazioni offerte.

La teoria della propagazione della luce in fibra è stata presentata cercando di snellire la trattazione e puntando piuttosto ad esprimere le conclusioni in maniera intuitiva. I concetti relativi ai modi di propagazione, alla distribuzione di campo, alle condizioni di taglio, alla dispersione dell'impulso, sono stati espressi in forma generale, ma è evidente la finalizzazione alle fibre monomodali (normali o a dispersione spostata) che saranno poi oggetto specifico della seconda parte di questo lavoro.

In questa parte l'obiettivo è stato quello di richiamare i parametri fondamentali e le relazioni di base che governano la propagazione. Nella parte successiva verranno affrontati invece aspetti specifici delle fibre monomodali quali l'effetto delle curvature, l'importanza del profilo dell'indice di rifrazione, il problema della giunzione tra fibre.

Bibliografia

- [1] Li, T.: *Next-Generation Lightwave Transmission Systems*. MoC0.2, ECOC '93, Montreux, vol. I, pp. 7-10.
- [2] Gabla, P.M. et al.: *Progress and Perspectives of Repeaterless Submarine Systems*. WeC10.1, ECOC'93, Montreux, vol. I, pp. 107-112.
- [3] Hasegawa, A. et al.: *Transmission of Stationary Nonlinear Optical Pulse in Dispersive Dielectric Fibers*. «Applied Phys. Lett.», vol. 23, agosto 1973, pp.142-144.
- [4] Mollenauer, L.F. et al.: *Soliton Propagation in Long Fibers with Periodically Compensated Loss*. «IEEE Journal of Quantum Electron.», vol. QE-22, gennaio 1988, pp. 157-173.
- [5] Mollenauer, L.F. et al.: *Demonstration, using Sliding-Frequency Guiding Filters, of Error-Free Soliton Transmission over more than 20 Mm at 10 Gb/s Single Channel, and over more than 13 Mm at 20 Gb/s in a Two Channel WDM*. «Electron. Letters», vol. 29, maggio 1993, pp. 910 - 911.
- [6] Christensen, B. et al.: *Soliton Communication on Standard Non Dispersion-Shifted Fiber*. MoC2.1, ECOC'93, Montreux, vol. I, pp. 27 - 31.
- [7] Okada, K. et al.: *Fiber-Optic Subscriber Systems using Point-to-Multipoint Transmission*. ThC11.3, ECOC'93, Montreux, vol. I, pp. 128 - 131.
- [8] O'Reilly, J.J. et al.: *Wireless Connection using mm-waves and Fiber*. ILSS'93, Vancouver, pp. 222 - 226.
- [9] Wale, M.J.: *Current Developments in RF and Microwave Fiber Optics*. WeP9.1, ECOC'93, Montreux, vol. I, pp. 99 - 104.
- [10] Miya, T. et al.: *Ultra Low-Loss Single Mode Fibers at 1.55µm*. «Electron. Letters», vol. 15, 1979, pp. 106 - 108.
- [11] Badar, A.H. et al.: *Transition and Pure Bending Losses in Multimode and Single-Mode Bent Optical Fibers*. «IEE Proceedings-part J», vol. 138, n. 4, agosto 1991, pp. 261 - 268.
- [12] Gambling, W.A. et al.: *Radiation from Curved Single-Mode Fibers*. «Electron. Letters», vol. 12, 1976, pp. 567 - 569.
- [13] Sameda, C.G.: *Radiation of Discrete Beams from Curved Single-Mode Fibers*. «Electron. Letters», vol. 13, 1977, pp. 712 - 713.
- [14] Artiglia, M. et al.: *Bending Loss Characterization in Single-Mode Fibers*. ECOC'87, Helsinki, vol. I, pp. 437 - 443.

Appendice

Si riportano di seguito le espressioni che definiscono il prodotto scalare e vettoriale, nonché la divergenza ed il rotore, allo scopo di facilitare la lettura del testo.

Siano \vec{A} e \vec{B} due vettori dello spazio tridimensionale, nel quale si suppone fissato un riferimento cartesiano ortogonale con versori \hat{x} , \hat{y} e \hat{z} . Le componenti dei

vettori nella direzione di un versore \hat{u} sono individuate mediante il pedice u , $u = x, y, z$. Allora si ha per definizione

- prodotto scalare di \vec{A} e \vec{B} :

$$\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z;$$

- prodotto vettoriale di \vec{A} e \vec{B} :

$$\vec{A} \times \vec{B} = (A_y B_z - A_z B_y) \hat{x} + (A_z B_x - A_x B_z) \hat{y} + (A_x B_y - A_y B_x) \hat{z}.$$

Si osservi che il prodotto vettoriale si può ottenere anche sviluppando il determinante della matrice

$$\begin{bmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{bmatrix}$$

secondo gli elementi della prima riga. Infine, il prodotto misto di tre vettori \vec{A} , \vec{B} e \vec{C} , ovvero $\vec{A} \times \vec{B} \cdot \vec{C}$, è dato dal determinante di una matrice le cui righe coincidono ordinatamente con le componenti cartesiane dei tre vettori moltiplicati.

Si introduce inoltre un vettore operatoriale detto "nabla", definito come

$$\nabla = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}.$$

L'applicazione del vettore operatoriale nabla ad una funzione scalare f dà luogo ad un vettore, denominato *gradiente* di f . L'interpretazione geometrica del vettore gradiente mostra che il corrispondente versore è diretto secondo la normale alla superficie definita nello spazio euclideo tridimensionale dalla equazione $f(x, y, z) = 0$.

Le operazioni di divergenza e rotore di un vettore \vec{F} , le cui componenti sono funzioni delle variabili x , y e z , sono formalmente definite come prodotto scalare e vettoriale rispettivamente del vettore nabla e del vettore \vec{F} . In modo esplicito si ha

$$\nabla \cdot \vec{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$

per la divergenza e

$$\nabla \times \vec{F} = \det \begin{bmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}$$

per il rotore. Un ulteriore operatore adoperato nel testo e' il laplaciano, formalmente esprimibile come la divergenza del vettore nabla, ovvero:

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

La sincronizzazione: l'obiettivo di sistema ed i problemi specifici

E' nel concetto stesso di telecomunicazione la necessità di cooperazione tra apparati localizzati in punti fisicamente distinti e spesso geograficamente lontani. Questa cooperazione implica che apparati interconnessi operino secondo regole reciprocamente note -utilizzando codici e protocolli opportuni-, ed in tempi correlati in misura più o meno stringente, in relazione alla natura ed alle caratteristiche dell'informazione da trasferire.

La conoscenza reciproca dei protocolli, e la correlazione dei tempi sono, di fatto, finalizzate a limitare allo stretto indispensabile (ma comunque, per sicurezza, con adeguate ridondanze) lo scambio dei messaggi di servizio tra apparati interconnessi, concentrando così la funzione trasmissiva sulle informazioni che sono veramente tali.

Quando un elemento di protocollo o un elemento di messaggio (ad esempio un blocco di simboli) è interpretabile indipendentemente dalla sua collocazione in una sequenza di elementi di messaggio, o dalla sua posizione temporale, la cooperazione tra apparati è di tipo asincrono; altrimenti, nel caso in cui sia necessaria la conoscenza della citata collocazione o la temporizzazione, la cooperazione è di tipo sincrono. In entrambi i casi, tuttavia, si pone un problema di sincronizzazione.

Con questa classificazione semplificata non si pretende di istituire una tassonomia esauriente ed ineccepibile; si vuole solamente richiamare l'attenzione sull'ampiezza del concetto di sincronizzazione che, come si è accennato, non è confinato in una dimensione temporale. Di fatto, il coordinamento logico e quello temporale si completano e si integrano, al fine di assicurare una corretta cooperazione tra apparati e tra nodi della rete di telecomunicazione.

Si intuisce che, in questa accezione di significato ampia e comprensiva, l'esigenza di sincronizzazione si è presentata già nei primi sistemi di telecomunicazione e, successivamente, in molti passaggi del processo di sviluppo sino ai sistemi attuali. Durante tale processo si è peraltro specificata in vari problemi, risolti, ovviamente, utilizzando gli strumenti tecnologici che si rendevano via via disponibili.

Per quanto esposto, nell'avviare un ciclo di articoli sul tema della sincronizzazione, si ritiene opportuno avvertire che è improprio considerare la sincronizzazione di una rete numerica, di dimensione nazionale, come un singolo problema, anche se di grande complessità e di rango sistemistico; ciò non toglie che, in relazione alla progressiva sostituzione,

nella trasmissione e nella commutazione, delle tecniche analogiche con le numeriche, aumenti notevolmente l'importanza della sincronizzazione complessiva della rete. Ed è perciò indispensabile che venga erogato uno sforzo progettuale espressamente rivolto alla configurazione di una strategia generale per assicurare tale sincronizzazione.

Nel perseguire questo obiettivo, non può essere tuttavia ignorato che i dati da immettere nel modello della rete sono, a loro volta, di complessa caratterizzazione; consistono infatti nelle prestazioni effettive di soluzioni tecniche, per lo più già operanti, escogitate per risolvere gli specifici problemi di coordinamento temporale e logico tra i processi, ed all'interno dei processi, che concretizzano le funzioni di una rete numerica.

Richiamare perciò con ragionevole dettaglio tali specifici problemi, risponde non tanto ad un desiderio di conoscenza storica dei fatti, quanto alla reale esigenza di disporre degli elementi da cui muovere per disegnare la citata strategia.

Gli studiosi che si sono cimentati con questi specifici problemi sono talvolta pervenuti, per vari motivi, a soluzioni diverse, di efficacia confrontabile, delle quali solamente una successiva sperimentazione in esercizio ha potuto far emergere gli effettivi vantaggi ed inconvenienti.

Con riferimento ad alcune scelte, si sono anche sviluppate vere correnti di pensiero, con il formarsi di schieramenti ed il consolidarsi di preferenze da parte dei tecnici interessati all'argomento. A tal proposito, è opportuno avvertire il lettore che, nei contributi che saranno oggetto di pubblicazione, potranno emergere opinioni ed orientamenti personali, che possono non coincidere con quelli assunti dalla SIP nei confronti della sincronizzazione.

E' peraltro doveroso sottolineare come alcune divergenze nascano da un problema di fondo, che si presenta con notevole frequenza nell'impostazione sistemistica delle grandi reti di telecomunicazione. Nelle soluzioni filosoficamente più valide è di norma maggiore l'interdipendenza tra le parti del sistema, ma ciò ne comporta anche una maggiore vulnerabilità; è soprattutto più probabile che guasti e disfunzioni si propaghino dal punto di insorgenza a notevoli aree della rete. D'altra parte, le tecniche di controllo e di supervisione, studiate per fronteggiare tali evenienze, possono divenire esse stesse un ulteriore fattore di vulnerabilità. Ben si comprende, perciò, l'attaccamento di alcuni tecnici di provata esperienza nei confronti di soluzioni più semplici e meno integrate, i cui limiti sono comunque ben noti.

p.r.

La sincronizzazione nelle reti e nei sistemi di Telecomunicazione

P. Rumboldt (*)

Nelle reti di telecomunicazione si presenta con grande frequenza, e con riferimento a varie funzioni essenziali delle reti stesse, il problema della sincronizzazione tra apparati che operano sia in punti (della rete) tra loro lontani, sia nell'ambito di un singolo nodo o di una stessa centrale. La sincronizzazione tra apparati è quella che riguarda più direttamente il progettista ed il gestore di una rete complessa, ed è pure la più difficile da assicurare; e ciò per varie ragioni tra cui è fondamentale la localizzazione fisica dei dispositivi da sincronizzare reciprocamente, data la velocità di propagazione finita dei segnali di informazione e dei segnali di tempo (cronosegnali), e l'influenza che i mezzi trasmissivi hanno sui tempi di propagazione di tali segnali. Ma ha anche notevole rilevanza il processo di sviluppo della rete; un processo continuo, e che si avvale di "tecniche" di più fornitori, con le inevitabili eterogeneità che permangono malgrado i grandi sforzi di standardizzazione compiuti in varie sedi; gli apparati, tra l'altro, appartengono spesso a diverse generazioni tecnologiche. E' infine da rilevare che, oltre a quelli citati, vi sono numerosi problemi di sincronizzazione all'interno degli apparati, la cui soluzione compete tipicamente ai progettisti che operano nel comparto industriale.

Con questo testo introduttivo, con cui si avvia un ciclo di articoli sull'argomento, si intende richiamare preliminarmente l'attenzione dei lettori su come il tema generale della sincronizzazione, pur essendo caratterizzato da una certa unitarietà concettuale, si sia concretizzato, nel tempo, a fronte di specifici problemi che si sono presentati nelle diverse funzioni delle reti: una specificazione cui corrisponde una diversificazione della natura stessa dei singoli problemi da risolvere, e che pone spesso in ombra le reciproche interrelazioni tra tali problemi. L'importanza di queste interrelazioni non è tanto da riferire alla citata unitarietà concettuale, quanto alle implicazioni che esse hanno sul funzionamento di una moderna rete complessa. In sostanza, la sincronizzazione di una moderna rete costituisce un obiettivo progettuale di natura sistemistica -e come tale unitario- che deve essere però perseguito tenendo conto delle effettive prestazioni delle soluzioni studiate ed adottate per gli specifici problemi.

La terminologia utilizzata è discorsiva e di uso corrente; per la precisa definizione di alcune fondamentali grandezze si rinvia agli altri articoli del ciclo.

1. Premessa

In una grande rete di telecomunicazioni a funzionamento automatico, sono numerosi i processi tra loro interrelati, il cui corretto svolgimento presuppone un coordinamento temporale tra fasi di processi diversi, e tra specifiche funzioni all'interno dei processi. Questa constatazione solleva qualche dubbio sulla opportunità di adottare il termine ellittico "sincronizzazione" per far riferimento ad un insieme di problemi eterogenei in cui è essenziale il citato coordinamento temporale.

Assumendo comunque per la sincronizzazione un'accezione di significato molto ampia e comprensiva, questa include anche, in un apparente paradosso, il funzionamento di sistemi asincroni: si intuisce come ciò derivi da un uso *assoluto* dell'aggettivo "asincrono", che richiederebbe invece la specificazione dei processi o delle funzioni *non sincrone* cui l'aggettivo allude.

Prescindendo comunque da ogni considerazione lessicale, è da notare che una trasmissione telex, fondata sull'uso di telestampanti asincrone, implica che nel terminale di ricezione l'apparato riceva un comando di *start*: che si sincronizzi per riconoscere correttamente il valore semantico di ogni bit della *parola* che si appresta a ricevere. Questa modalità di sincronizzazione sembra

(*) ing. Paolo Rumboldt - SIP DG - Roma

essere molto lontana da quelle utilizzate nelle moderne reti numeriche; essa ha invece, nelle grandi linee, notevoli analogie con il criterio con cui si assicura l'interoperatività tra nodi di commutazione plesiocroni, cioè temporizzati da orologi indipendenti.

Il confronto tra la sincronizzazione dei terminali telex e quella tra nodi plesiocroni, di per sé banale ed improprio (se analizzate in modo approfondito), fornisce l'occasione per sottolineare come nei problemi attinenti alla sincronizzazione siano rilevanti gli aspetti *quantitativi* e quelli *statistici*. Nel telex, data la semplicità e l'*imprecisione* degli orologi presenti nei due terminali, il riallineamento è attuato *sistematicamente* per ogni parola (di 5 bit); nei nodi di commutazione, invece, tale esigenza si presenta ad intervalli comparativamente molto lunghi e *non con sistematicità aprioristicamente determinata*.

Si deve anche osservare che, dall'analisi delle due situazioni richiamate, emerge un'altra essenziale differenza, che è appunto legata alla utilizzazione dei segnali di sincronismo inviati dal terminale di trasmissione: nei sistemi asincroni i segnali di sincronismo tipicamente *comandano* il processo di sincronizzazione del terminale di ricezione, mentre in quelli sincroni o "quasi sincroni", tali segnali vengono *normalmente* impiegati per *monitorare* la situazione di reciproco sincronismo, e per avviare eventualmente la risincronizzazione quando si sia perduta. Questa constatazione suggerisce l'opportunità di designare con termini distinti i segnali di sincronismo aventi le due diverse funzioni, come si chiarirà successivamente.

A partire dagli anni '80, in varie sedi internazionali si è cercato di dare un inquadramento all'intera materia attinente alla sincronizzazione, che si presenta come molto articolata anche per motivi storici; secondo una classificazione piuttosto generale, i vari problemi possono essere raggruppati in tre grandi aree:

- sincronizzazione di portante e di bit;
- sincronizzazione di trama;
- sincronizzazione di rete.

Un esame non superficiale di tali aree mette in luce tuttavia, insieme ad una certa difficoltà logica nel ripartire razionalmente la materia, la presenza di numerosi punti di intersezione tra i problemi delle diverse aree. Ed è per converso facile scoprire che ogni problema tecnico ed ingegneristico, che si deve risolvere con riferimento agli specifici processi da sincronizzare, risente di fattori presenti nelle tre aree, e rilevanti in termini di cause e/o di effetti.

Va anche osservato che oggi, quando si parla di sincronizzazione delle reti di Tlc, ci si riferisce per lo più ai problemi sopra richiamati, che però non esauriscono *tutti i problemi di temporizzazione coordinata* presenti nei processi riguardanti le fondamentali macrofunzioni su cui si fondano le grandi reti: trasduzione, trasmissione, commutazione e segnalazione. Per tale motivo un'analisi più approfondita, che non ha peraltro presunzione di

completezza, richiede che l'argomento sia tralasciato dal punto di vista degli specifici processi in cui si pongono problemi di sincronizzazione.

2. La sincronizzazione delle reti: qual è il problema?

2.1 La sincronizzazione dei processi

Nel termine sincronizzazione è già implicitamente contenuto il concetto dell'aggancio tra due o più fenomeni fisici, diversamente localizzati, che devono realizzarsi in reciproca correlazione temporale: simultaneamente o dopo un tempo predeterminato. Questo aggancio non è però di tipo causale: i fenomeni considerati si attuano nell'ambito di processi che si evolvono secondo ritmi tendenzialmente indipendenti, nei quali è proprio la sincronizzazione ad introdurre un ulteriore vincolo temporale.

Il richiamo ai processi è tutt'altro che casuale, dato che la sincronizzazione può riguardare, invece che singoli fenomeni, *fasi di processi*, il cui aggancio deve avvenire in base a criteri logici, escludendo comunque un rapporto diretto tra cause ed effetti.

Nel primo dei due casi, l'aggancio tra fenomeni fisici avviene mediante segnali che trasferiscono un'informazione di tipo sostanzialmente temporale e che, convenzionalmente, definiamo *cronosegnali*.

Nel secondo caso, l'informazione di sincronizzazione deve *avvertire* il nodo, o il dispositivo in cui si svolge un processo, dello stadio raggiunto da un altro processo, *non appena* lo stadio è stato raggiunto, ma indipendentemente da *quando* è stato raggiunto rispetto ad un riferimento temporale assoluto. L'informazione di sincronizzazione è ora di tipo logico; i segnali che la convogliano sono convenzionalmente definiti *sincrosegnali*.

I cronosegnali controllano la *velocità* di svolgimento dei processi, il *ritmo* secondo cui si evolvono, mentre i sincrosegnali servono ad assicurare che processi diversi, che si devono svolgere in modo coordinato, si sviluppino rimanendo *in passo*. O, ancora, che possano ritornare in passo quando abbiano perduto il coordinamento.

La distinzione è piuttosto importante, soprattutto nell'ambito delle tecniche numeriche, perché processi coordinati (addirittura, in molti casi, reciproci), possono anche differire lievemente nei ritmi di svolgimento (che dipendono dai cronosegnali), ma necessitano di una rigorosa correlazione per quanto attiene ai caratteri logici e numerici, che è assicurata dai sincrosegnali. Per chiarire tale concetto basta ricordare che, mentre una differenza di ritmo influisce generalmente solo sulla qualità dell'informazione ricevuta, una perdita di sincronismo "logico" si traduce in una totale equivocazione del significato e della funzione che ciascun bit espleta nell'ambito di un flusso numerico.

E' ancora da osservare che la frequenza (cioè il numero di volte, nell'unità di tempo) con cui è necessario operare una *risincronizzazione* è legata alla differenza tra i cronosegnali dei processi tra loro coordinati (tale operazione è però attuata grazie ai sincrosegnali). Inoltre, essendo segnali di ritmo, i cronosegnali sono tipicamente periodici e come tali vengono utilizzati: essi sono anche tipicamente segnali di comando.

I sincrosegnali, per converso, possono essere anche aperiodici; se periodici, nel funzionamento normale degli apparati, cioè *non nella fase di recupero della sincronizzazione*, sono utilizzati per monitoraggio.

L'attuale tendenza riguardante la *sincronizzazione generale* delle reti è di assicurare l'identità del ritmo di riferimento in tutti gli apparati ed in tutti i punti della rete; e ciò, distribuendo i cronosegnali generati in un unico punto, e relegando pertanto i sincrosegnali in una funzione di monitoraggio. E' peraltro ovvio che, come si è già detto, i sincrosegnali assumono invece una funzione *attiva* per il recupero del sincronismo o per la sincronizzazione iniziale.

Per tutti gli apparati operanti in un determinato nodo i ritmi di funzionamento possono essere per lo più ricondotti ad un unico cronosegnale; mentre, per quanto concerne i sincrosegnali, essi riguardano le varie esigenze di interoperatività tra gli apparati presenti nel nodo o in più nodi della rete; sono pertanto molteplici, anche in relazione ai vari livelli gerarchici di moltiplicazione e di affasciamento dei segnali nei nodi e negli apparati.

Un chiarimento su quanto esposto emerge da alcune considerazioni sulla natura dei processi da sincronizzare.

2.2 Sincronizzazione per il riconoscimento di simbolo

Il termine sincronizzazione allude chiaramente alla esigenza di un coordinamento temporale tra due o più processi, cioè ad una temporizzazione *relativa*. I processi in questione hanno spesso carattere tipicamente *inverso o reciproco* (in senso matematico). Tale reciprocità è in alcuni casi molto evidente, come nella moltiplicazione/demoltiplicazione, mentre in altri emerge da un'analisi piuttosto raffinata dei processi stessi. E' questo, ad esempio, il caso della commutazione con tecniche a divisione di tempo.

L'obiettivo complessivo che viene perseguito in una rete -il trasferimento indirizzato dell'informazione- comporta appunto, nel caso di informazioni che devono essere restituite nella loro caratteristica temporale (come ad esempio l'informazione vocale), che venga assicurata la sincronizzazione con riferimento ai processi che riguardano i vari livelli della struttura OSI. Ma non è difficile individuare, nelle modalità per

perseguire questo obiettivo, la necessità, per lo più ricorrente, di risolvere due problemi basilari: a) il riconoscimento, nel segnale elementare (elettrico, elettromagnetico, ottico) del simbolo che esso vuole rappresentare; b) il riconoscimento del valore semantico, cioè del significato e della funzione, di ciascun simbolo nell'ambito della struttura numerica cui appartiene (parola, trama, supertrama, canale logico ecc.).

2.2.1 Sincronizzazione di simbolo elementare, <di "bit">; forme d'onda e simboli

Come si è accennato, si pongono problemi di sincronizzazione *anche* all'interno dei singoli processi in cui si concretizzano le *macrofunzioni* di una rete numerica. Uno di tali problemi è così ricorrente e fondamentale, nei confronti della funzione *trasmissione*, da aver condizionato sostanzialmente le scelte operate nella impostazione dei sistemi trasmissivi numerici e, prima ancora, di quelli analogici. Si può anzi affermare che un intero filone di attività di studio e di progettazione si è sviluppato per risolvere tale specifico problema.

In alcuni dei trattamenti riguardanti la funzione trasmissione, e nelle relative modalità di attuazione, è possibile individuare una comune finalità di fondo consistente nella necessità di riconoscere "al meglio", nei segnali ricevuti, il vero contenuto informativo, separandolo dai disturbi di varia natura (rumore termico, prodotti di intermodulazione, diafonia, interferenza intersimbolica ecc.); si tratta, in pratica, di ridurre il campo di incertezza *ermeneutica*, cioè interpretativa, del processo di rivelazione (riferita ai bassi livelli semantici: simbolo o lettera di un alfabeto ecc.), assicurando così una buona corrispondenza tra i segnali inviati e quelli rivelati. L'ottimizzazione del processo di rivelazione implica che la ricerca del *più probabile valore simbolico* elementare di un segnale impulsivo venga effettuata in particolari intervalli di tempo -detti intervalli di decisione- la cui posizione temporale sia la più opportuna in relazione al rapporto tra i valori istantanei del segnale da riconoscere e quelli dei segnali di disturbo che lo inquinano. A questo fine, dati gli sbandamenti di fase del segnale stesso rispetto alla sua posizione temporale "nominale" (causati dai fenomeni di propagazione, dai mezzi trasmissivi ecc.) è necessario che gli intervalli di decisione seguano i citati sbandamenti; in altre parole che ne sia assicurato *dinamicamente* il posizionamento ottimale.

I problemi di questa classe sono, in certa misura, riconducibili a quelli della *rivelazione coerente*, nella quale non è rilevante la fase assoluta della portante di demodulazione, quanto quella relativa al segnale da rivelare.

Alla sincronizzazione per il riconoscimento di simbolo è affidato il compito di contenere il tasso d'errore nei processi di rivelazione e di rigenerazione o, in altre parole, di limitare *l'equivocazione*; questa sincronizzazione, che riguarda principalmente la trasmissione con tecniche numeriche, risponde, come si è detto, alle esigenze della *rivelazione coerente*, e corrisponde alla *sincronizzazione di portante* nell'ambito delle tecniche analogiche: le analogie attengono sia agli effetti sulla correttezza e sulla intellegibilità delle informazioni restituite dal processo, sia ai metodi con cui estrarre la portante dal flusso informativo. E' ancora da osservare che, nelle tecniche di trasmissione numerica a banda traslata, il problema della rivelazione coerente si articola appunto nei due aspetti *di portante* (estrazione e ricostruzione della portante) e *di simbolo* (ricerca del simbolo in un intervallo di decisione), tra loro interconnessi, cosicché i dispositivi di rivelazione più efficaci li trattano simultaneamente.

Su questo argomento si tornerà nei testi del ciclo; è invece necessario chiarire subito che, a differenza di quanto verrà detto con riferimento alla sincronizzazione tra processi logici, nella sincronizzazione di portante e di simbolo non è possibile prescindere dall'analisi fine delle forme d'onda: sia di quelle che materializzano i cronosegnali utilizzati nei processi specifici, che di quelle relative ai segnali da trattare.

Tale analisi è peraltro indispensabile in tutti quei casi in cui, in relazione alle frequenze di cifra molto elevate, ed alla presenza di disturbi rilevanti, i segnali da trattare ed i cronosegnali non possono essere nemmeno approssimativamente stilizzati nelle funzioni matematiche normalmente usate nello studio dei segnali impulsivi (funzione "impulsiva unitaria", funzione "gate" ecc.).

La sincronizzazione di simbolo, pur essendo concettualmente collegata, per diversi aspetti, a quella di parola e di trama, se ne diversifica piuttosto sostanzialmente anche sotto un profilo semplicemente storico; essa è usualmente trattata nei testi tecnici con riferimento alla *frequenza di simbolo*, cioè alla grandezza direttamente legata, nei sistemi numerici, agli orologi ed ai cronosegnali. Si parla spesso, a questo proposito, di sincronizzazione "di bit", anche quando riferita a simboli non binari.

L'argomento in questione, avendo costituito uno dei principali problemi da risolvere nei primi sistemi PCM per cavi metallici (a coppie simmetriche e coassiali) negli anni '60 e '70, si è oggi decisamente ridimensionato nel caso dei sistemi in fibra ottica, mentre continua ad essere rilevante nella trasmissione numerica su ponte radio e per i sistemi radiomobile numerici. In un prossimo futuro, tuttavia, è possibile che il tema si riproponga, sia pure con diverse modalità, nella trasmissione su fibra ottica fondata sull'impiego dei solitoni, e nelle "tecniche coerenti".

2.2.2 Sincronizzazione a livello logico; i processi reciproci

Con l'espressione livello logico si intende far riferimento alla sincronizzazione delle reti numeriche *in quanto tali*: sistemi complessi che trasferiscono, instradano e trattano *simboli*. E' questa una precisazione necessaria per i riflessi sui problemi di sincronizzazione.

Nelle tecniche di trasmissione, multiplazione e commutazione numerica, il valore semantico di ciascun simbolo dipende dalla sua posizione temporale nell'ambito di *parole*, e di *trame* di vario livello gerarchico; per tale motivo al problema, di cui si è già detto, di separare il segnale dal disturbo per riconoscere correttamente il simbolo, si aggiunge quello di riconoscere la posizione temporale di ciascun simbolo, relativamente alla parola e alla trama cui appartiene, e di ciascuna trama nell'ambito di trame di livello superiore.

In particolare, nelle tecniche numeriche binarie, in cui i segnali impulsivi sono uguali tra loro, ed inseriti tipicamente in flussi continui, il citato riconoscimento deve attuarsi *inserendo ogni singolo simbolo binario in un intervallo di tempo appartenente ad un flusso di intervalli di tempo, organizzato, in parole e trame di vario ordine, come il flusso da interpretare, e sincronizzato con esso*. In un nodo ricevente, in cui si attua il riconoscimento, deve pertanto operare un dispositivo capace di predisporre il flusso di intervalli di tempo che si sincronizzerà con il flusso di simboli grazie ai sincrosegnali. Questa operazione è richiamata nella letteratura tecnica come *sincronizzazione di trama* (frame synchronization). Essa può riguardare trame aventi bit-rate lievemente diversi, e consiste nell'assicurare una corrispondenza biunivoca tra gli intervalli di tempo del flusso numerico che perviene ad un nodo e quelli del flusso ricostruito localmente.

I sincrosegnali sono *marcaggi riconoscibili*, inseriti nei flussi di simboli binari che convogliano le informazioni di messaggio. La necessità di utilizzare, anche per i sincrosegnali, gli stessi simboli binari impiegati per i messaggi, costringe a caratterizzarli grazie alla loro ricorrenza statistica, diversa da quella dei simboli che derivano dalla codifica dei messaggi.

Per quanto concerne i sincrosegnali, è necessario un chiarimento: il marcaggio temporale che essi forniscono è *relativo* ai flussi di simboli di cui fanno parte, e nell'ambito dei quali permettono di identificare il valore semantico di ciascun simbolo. Ciò significa che i sincrosegnali non sono propriamente segnali di tempo, ma particolari sequenze di simboli (raggruppati in parole o distribuiti) inseriti nei flussi continui di bit per scandirne l'organizzazione in trame, parole ecc.

In relazione alla organizzazione gerarchica dei flussi di simboli, in una rete numerica si presentano esigenze di sincronizzazione ai vari livelli della gerarchia; d'altra parte, le soluzioni adottate per la sincronizzazione di

ciascun livello devono essere strettamente coordinate con quelle relative agli altri livelli. Ciò, sia per evitare situazioni di blocco della rete, a causa di interazioni tra i dispositivi di monitoraggio e di recupero del sincronismo, sia per ridurre i tempi di risincronizzazione, e gli effetti della perdita di sincronismo sui servizi di Tlc; sia, infine, per evitare l'avvio di procedure di risincronizzazione non indispensabili⁽¹⁾.

La configurazione di una strategia di sincronizzazione è peraltro molto complessa, non solo in relazione alle citate interazioni tra i vari livelli di sincronizzazione, ma anche a causa della aleatorietà dei punti del sistema in cui insorgono le cause di perdita del sincronismo.

3. I cronosegnali: generazione e trasferimento

Per effettuare concretamente la sincronizzazione, data la tipica reciprocità tra i processi da sincronizzare, è necessario disporre, nei punti della rete in cui i processi si attuano, di *cronosegnali* identici o molto prossimi, e di *sincrosegnali*.

Come si è già detto, sarebbe auspicabile che i ritmi di svolgimento dei processi reciproci, che si attuano nei vari nodi sincronizzati, coincidessero esattamente, anche se ciò in molti casi non è indispensabile; poiché tali ritmi sono determinati dai cronosegnali, si intuisce la possibilità di operare, con riferimento ai cronosegnali, secondo due diverse strategie: generarli localmente con generatori plesiocroni (vedi par.3.1), ovvero trasmettere in tutti i nodi della rete quelli generati in un solo punto (o in pochi punti).

Si deve però osservare che, nella traduzione dei due criteri citati nei programmi di realizzazione delle reti, sono intervenuti molti fattori connessi sia alle possibilità tecnologiche che si sono rese disponibili via via nel tempo, sia all'innestarsi dei programmi su reti già esistenti, e su soluzioni già adottate in precedenza.

Sotto questa premessa, i problemi attinenti alla generazione ed alla trasmissione dei segnali di tempo vanno studiati non solo in una visione ingegneristica, ma con specifica attenzione ad una realtà nazionale. E' da aggiungere che l'alternativa tra il decentramento della generazione e la distribuzione (e quindi il trasferimento) riguarda tipicamente i cronosegnali, dato che per i sincrosegnali la loro stessa natura ne implica la trasmissione tra i punti in cui si attuano i processi reciproci da sincronizzare.

Di fatto, le due possibili soluzioni atte ad assicurare la presenza dei cronosegnali in tutti i punti (nodi ed apparecchiature) in cui è necessaria una sincronizzazione di ritmo, si sono confrontate nel tempo ed è prevalsa l'una

o l'altra, in ragione del peso che assumevano i diversi fattori tecnico economici in gioco; un confronto dal quale, in pratica, è emersa la tendenza ad integrarle.

I principali fattori in questione sono:

- la precisione, la stabilità e l'affidabilità dei generatori primari "indipendenti";
- le tecniche di trasmissione dei cronosegnali ed il loro peso relativo nell'economia dei sistemi di trasmissione, con particolare riferimento all'ingombro dei cronosegnali a scapito dei segnali di informazione;
- le previdenze necessarie per assicurare il sincronismo logico tra apparati operanti in ambito plesiocrono;
- le tecniche necessarie per compensare, in ricezione, le variazioni di fase dei cronosegnali causate dal mezzo trasmissivo (tipicamente in ambito *mesocrono*, vedi par.3.3);
- le tecniche con cui assicurare la continuità di funzionamento dei nodi subordinati (che utilizzano cronosegnali generati in altri nodi) in caso di mancata ricezione dei cronosegnali;
- l'affidabilità della rete di trasferimento della sincronizzazione.

Nel valutare i fattori citati intervengono sempre, direttamente o indirettamente, considerazioni di costo.

I fattori elencati, d'altra parte, possono essere oggi considerati e valutati simultaneamente per le scelte di natura progettuale o programmatica da attuare in un quadro di tecniche e di tecnologie disponibili ben noto, anche per quanto attiene ai costi; ma il loro peso è mutato notevolmente nell'ultimo cinquantennio, durante il processo di evoluzione dei sistemi di Tlc. Ed è pure da richiamare il sostanziale mutamento di esigenze e di prospettive, che si è determinato quando la sincronizzazione, prima confinata nell'ambito delle tecniche di trasmissione, ha interessato anche la commutazione e, soprattutto, l'interoperatività delle due tecniche.

3.1 I cronosegnali nei sistemi di trasmissione: generazione plesiocrona o trasmissione

Il problema di come trasferire i cronosegnali si è posto concretamente negli anni '50 nella multiplazione FDM (frequency division multiplexing); si è infatti manifestata l'esigenza di fornire al demultiplicatore una *portante di conversione* isofrequenziale con quella utilizzata nella multiplazione⁽²⁾.

(1) Si veda nei successivi articoli quanto attiene agli "slip controllati".

(2) Un'analisi più di dettaglio del problema pone in luce inoltre un'ulteriore esigenza riguardante la fase della portante. Per evitare gli effetti degli scostamenti di fase della portante di conversione dal valore nominale, nei demodulatori di tipo più diffuso, si immette un segnale di grande ampiezza che ai fini del processo di demodulazione può essere assimilato ad una funzione "gate".

Tra i modi per ottenere il segnale in questione vi è quello di trasmetterlo unitamente al segnale che convoglia l'informazione (da demultiplare), ma si tratta di un approccio al problema che contrasta con un principio generale, da applicare ogniqualvolta si progetti un sistema trasmissivo nel rispetto della teoria dell'informazione.

Per motivi di ottimizzazione nell'impiego della capacità dei canali trasmissivi (contenimento della banda impiegata e/o della potenza in trasmissione, riduzione del carico degli amplificatori ecc.) è infatti opportuno, già in trasmissione, depurare il segnale dei contenuti informativi che non corrispondono a "vere" informazioni da trasferire, nel senso che non ne possiedono l'imprevedibilità, misurata come è noto dall'entropia.

Ora, tra i contenuti informativi, importantissimi ma di limitata entità, vi sono appunto quelli che permettono la ricostruzione, in ricezione, dei cronosegnali i quali, d'altra parte, sono indispensabili per il processo di rivelazione, ma non sono necessariamente specifici di ogni flusso informativo; essi sono, di fatto, (o possono essere) comuni a tutti i flussi che escono da un nodo trasmissivo o di commutazione, cosicché l'informazione temporale può essere più convenientemente trasferita, con la tecnica del "canale comune".

Sulla base di queste considerazioni di fondo, dai primi multiplex FDM *a portante trasmessa*, si è rapidamente passati a quelli *a portante soppressa*, affidando a generatori *plesiocroni* (cioè operanti a frequenza molto prossima a quella nominale: plesios=vicino) il compito di riprodurre, per la demodulazione, le frequenze portanti molto prossime a quelle utilizzate nella modulazione. E' peraltro da osservare che anche i segnali di sistemi a portante soppressa contengono ancora gli elementi informativi relativi alla portante; l'estrazione di tali informazioni non era però agevole al tempo in cui il problema della sincronizzazione si presentò per i sistemi in questione. Da ciò l'opzione per i generatori plesiocroni.

Malgrado i particolari accorgimenti impiegati per contenere alcuni degli effetti della non perfetta identità di frequenza dei generatori plesiocroni sui segnali demodulati, i più evoluti sistemi FDM, operanti a frequenze massime di 60 MHz ed oltre, richiedevano precisioni irraggiungibili in ambito plesiocrono (con i generatori master commerciali dei primi anni 70), e spingevano all'adozione di altre tecniche; è anche da ricordare, a tal proposito, che le esigenze di sincronia, derivanti dal crescente impiego dei canali telefonici per la trasmissione dei dati, possono essere significativamente più onerose di quelle relative alla trasmissione dei segnali telefonici.

Le tecniche verso cui ci si orientò concretamente si fondano sulla distribuzione di cronosegnali di riferimento, o sulla estrazione delle portanti dai segnali modulati, attuata mediante procedure di stima statistica; entrambi i citati approcci implicano, comunque, l'adozione di disposizioni circuitali piuttosto complesse, in cui è centrale il PLL (phase locked loop).

Mentre con l'evoluzione delle tecniche di trasmissione aumentavano le esigenze in materia di cronosegnali, si registravano anche sostanziali progressi nella realizzazione dei generatori primari.

L'evoluzione dei generatori primari -cioè non asserviti ad altri- è stata molto notevole negli ultimi quaranta anni, non solamente per le precisioni e le stabilità raggiunte in assoluto, ma anche perché prestazioni molto elevate possono essere fornite da campioni primari "commerciali", di costo e dimensioni relativamente ridotte, anche trasportabili. Ciò ha consentito, tra l'altro, l'esecuzione, con metodi piuttosto speditivi, del confronto di frequenza e di fase tra campioni primari situati in punti lontani della terra⁽³⁾.

Tale problema ha relazione con le telecomunicazioni, in quanto esse hanno fornito gli strumenti per effettuare i confronti in questione; marginale in questo contesto, ha comunque attinenza con quegli aspetti della sincronizzazione dei nodi di una rete in cui sono rilevanti i tempi di trasferimento dei segnali determinati dai mezzi trasmissivi.

Sulla caratterizzazione dei campioni primari si tornerà in altri testi del ciclo di articoli sulla sincronizzazione; è però utile richiamare, sin d'ora, che questo argomento richiede un'analisi fine dei segnali, in relazione ai dispositivi che possono fisicamente produrli o trattarli.

3.2 *La sincronizzazione di bit nella trasmissione numerica*

Tornando nel campo della trasmissione numerica, il problema di disporre di cronosegnali di adeguate caratteristiche in tutti i punti di rigenerazione di un collegamento trasmissivo, fa propendere, nei sistemi numerici per portanti metallici, verso l'estrazione dei cronosegnali dal flusso numerico che convoglia l'informazione, con l'eventuale ricorso a dispositivi non lineari, nel caso in cui il flusso dei segnali di linea non contenga *di per sé* la componente spettrale corrispondente alla frequenza di bit-rate, o a multipli o sottomultipli di essa.

(3) Il confronto tra campioni lontani, fino alla metà degli anni '60 (quando si rese appunto possibile il trasporto fisico dei generatori primari) veniva effettuato via radio; ma la velocità di propagazione finita delle onde elettromagnetiche, unitamente agli affetti della mutevole distanza dalla terra degli strati ionizzati rifrangenti le onde, costringevano ad adottare particolari previdenze come l'impiego di onde molto lunghe, di prolungati e tediosissimi confronti di fase, e l'integrazione con misure astronomiche. E ciò ottenendo risultati spesso insufficienti in relazione alle esigenze dei laboratori scientifici e, in qualche caso, anche a quelle più modeste, ma pur sempre rilevanti, dei sistemi di Tlc.

L'estrazione dei cronosegnali dal flusso di dati, oltre a non comportare, in prima approssimazione, alcun aggravio in termini di ingombro del canale trasmissivo, può facilitare il processo di decisione nei rigeneratori, poiché questo processo è influenzato dal jitter. In altri termini, ai fini della trasmissione numerica, un cronosegnale estratto dal segnale di informazione può presentare alcuni vantaggi, tra cui quello di non imporre vincoli stringenti nei generatori di ritmo dei terminali trasmettenti.

I vantaggi in questione sono però limitati al singolo sistema di trasmissione, cioè non considerato nel contesto di una gerarchia numerica di ordine superiore, né in quello, ancor più ampio, di una rete numerica commutata. In alcuni di questi casi il jitter può essere compensato o riassorbito con dispositivi piuttosto semplici. Ben più complesso è, invece, il problema dell'affasciamento, in un flusso trasmissivo di ordine superiore, di flussi tributari caratterizzati da cronosegnali diversi.

3.3 Dalla trasmissione numerica alla rete numerica

Come verrà chiarito in alcuni testi di questo ciclo, limitatamente ad una moltiplicazione numerica, il problema dell'affasciamento di flussi tributari *eterocroni* può essere risolto con l'ingegnosa tecnica del <pulse stuffing>, e/o con le modalità caratteristiche dell'SDH (Synchronous Digital Hierarchy).

Queste tecniche però, al termine del tronco trasmissivo, restituiscono i flussi tributari ciascuno con il proprio ritmo. Volendo pertanto operare su tali flussi funzioni di commutazione senza decodificare i segnali, è necessario accettare che il commutatore, che opera con un proprio ritmo, dopo aver accumulato nelle "memorie tampone" le massime dissincronie che queste possono contenere, si sganci dal flusso numerico ed avvii una fase di risincronizzazione a livello logico (utilizzando cioè i sincrosegnali). E' piuttosto ovvio che, se i cronosegnali dei tributari sono plesiocroni rispetto a quello del commutatore, la situazione critica di perdita e di recupero del sincronismo si ripeterà di rado. Ma non così nel caso di sistemi trasmissivi tributari, in cui i generatori di cronosegnali siano di modeste prestazioni in termini di approssimazione (di breve e di lungo periodo) dei valori nominali della frequenza.

Queste considerazioni fanno emergere la esigenza di utilizzare in una rete numerica complessa *quanto meno* generatori plesiocroni; o, meglio ancora, di disporre in tutta la rete di un unico cronosegnale. Si deve però osservare che, anche in quest'ultimo caso, vari fattori determinano spostamenti di fase dei diversi flussi numerici attorno al valore nominale, con riflessi sui valori non più costanti delle frequenze di cifra: i ritmi rispettano, di fatto, solo mediamente la frequenza del cronosegnale primario e la rete, nel suo complesso, risulterà non sincrona ma <mesocrona>.

Per assicurare la presenza dei cronosegnali in tutti i punti in cui sono necessari, si è già accennato alla possibilità di configurare due strategie di soluzione del problema: generarli indipendentemente in punti diversi della rete, con generatori di alta precisione e stabilità, o trasferirli da un nodo all'altro della rete. Quest'ultima strategia può a, sua volta, fondarsi su due diversi criteri: convogliare i cronosegnali sullo stesso flusso di dati che trasporta l'informazione, ovvero instradarli su appositi canali trasmissivi ad essi riservati.

L'esigenza di sincronizzazione reciproca può riguardare processi che si attuano in più di due punti della rete, come tipicamente nelle strutture "a maglia"; in tal caso, il carattere *relativo* della sincronizzazione continua a sussistere, purché si assuma che il cronosegnale generato in uno dei punti sia quello *di riferimento*, al quale tutti gli altri punti si devono uniformare. Con facili ragionamenti si perviene alla conclusione che tutti i nodi di una grande rete nazionale devono essere reciprocamente sincronizzati secondo una strategia di tipo gerarchico.

All'adozione di una *disciplina gerarchica* di sincronizzazione, consegue una concentrazione della responsabilità di funzionamento dell'intero sistema sui generatori (primari) di più alto livello, e la necessità di contenere gli effetti di eventuali guasti di essi. Tale necessità è soddisfatta operando secondo criteri di *asservimento condizionato*, e dotando i generatori asserviti *anche* di capacità di funzionamento indipendente.

Deve essere chiarito, a questo punto, che ogni soluzione da adottare va esaminata non solo con riferimento ad una situazione di funzionamento normale, ma anche in condizioni di avaria di qualche parte della struttura di sincronizzazione: il problema si proietta, quindi, in una dimensione sostanzialmente *quantitativa* con la necessità, tra l'altro, di valutare comparativamente grandezze eterogenee, e di attingere in molti casi ai criteri della Ricerca Operativa. I fattori di cui tenere conto si collocano in diverse aree: la frequenza e la gravità dei disservizi in relazione alla perdita di sincronismo, la perdita di sincronismo in relazione a situazioni di funzionamento normale o anomalo degli apparati, l'affidabilità dei sistemi di generazione e di distribuzione dei cronosegnali, i costi dei provvedimenti per contenere gli effetti indesiderati ecc.

3.4 La trasmissione dei cronosegnali

Uno dei motivi per cui si è inizialmente introdotta una classificazione dei segnali di tempo in *cronosegnali* e *sincrosegnali*, è nel diverso tipo di informazione che essi devono convogliare. I segnali di sincronismo-cronosegnali e sincrosegnali- hanno struttura e funzioni differenti, che influiscono, tra l'altro, sul modo con cui si possono trasferire, data la diversa sensibilità alle caratteristiche dei mezzi trasmissivi che li convogliano.

Si è accennato alle modalità con cui trasferire i cronosegnali. Si ritiene però utile integrare quanto già esposto richiamando alcuni aspetti rilevanti del problema del trasferimento di tali segnali:

- i cronosegnali sono segnali di ritmo, e sono pertanto caratterizzati, in linea di principio, da una sola frequenza; la trasmissione di essi non è influenzata da un sistema trasmissivo lineare tempo-invariante (in senso stretto) e causale; il ritardo di fase determinato dal sistema trasmissivo, se costante nel tempo, è ininfluente ai fini della sincronizzazione dei ritmi dei processi dei due nodi della rete da sincronizzare; può invece divenire importante se cambia nel tempo, data la ben nota relazione differenziale tra fase e frequenza di un segnale;
- poiché in molti casi il mezzo trasmissivo risente di fattori esterni (ad esempio della temperatura) una rete, anche se totalmente sincronizzata mediante cronosegnali di un'unica origine, risulta non sincrona ma *mesocrona*: in essa sono identici solamente i valori medi delle frequenze dei segnali di ritmo. Gli effetti (sui processi) sono ovviamente differenti in relazione alla specifica sensibilità dei processi.

4. I sincrosegnali

Da quanto esposto con riferimento alla struttura ed alla funzione dei sincrosegnali, si intuisce come la scelta delle sequenze di simboli, in cui si concretizzano, sia fortemente influenzata dalle modalità con le quali, nell'ambito delle diverse funzioni di rete (moltiplicazione, estrazione e reimmissione di tributari, commutazione) si intende attuare la sincronizzazione logica dei processi. I fattori in gioco sono molteplici e da correlare, in particolare, alle caratteristiche di qualità, e di disponibilità dei mezzi trasmissivi su cui i sincrosegnali sono convogliati, ed agli effetti della perdita di sincronismo sulla qualità dei servizi. Gran parte della materia è stata oggetto di studi approfonditi che hanno portato, in molti casi, alla standardizzazione dei sincrosegnali o a specifiche riguardanti i massimi effetti tollerabili sui servizi erogati.

L'argomento, che mal si presta ad una trattazione generale, si specializza nei diversi sistemi di telecomunicazione, con riferimento ai quali sarà ripreso nell'ambito di questo ciclo di articoli ed in contributi più specifici. Nell'economia di una presentazione preliminare sulla sincronizzazione, si ritiene comunque opportuno richiamare alcuni aspetti legati alla loro trasmissione:

- i sincrosegnali sono simboli o gruppi di simboli inseriti nei flussi numerici, caratterizzati da particolari sequenze o ricorrenze statistiche che ne rendono possibile l'identificazione nell'ambito del flusso e, di conseguenza, l'identificazione del valore semantico di ciascun bit del flusso;

- il trasferimento dei sincrosegnali è, per lo più, automaticamente assicurato dalla trasmissione dei flussi di simboli in cui sono inseriti; non è un problema diverso dalla trasmissione dei dati che convogliano l'informazione;
- la trasmissione influisce tipicamente su di essi in termini di possibile equivocazione;
- i sincrosegnali sono relativi agli specifici processi; essi sono perciò numerosi nel caso di flussi ottenuti da molti tributari e per moltiplicazioni successive. Da ciò l'opportunità di inviarli attraverso gli stessi flussi di dati che convogliano le informazioni.

5. La sincronizzazione al di là degli aspetti logici

L'obiettivo della sincronizzazione potrebbe essere teoricamente raggiunto se, in ogni nodo, si disponesse di generatori di cronosegnali perfetti, operanti esattamente e stabilmente su una determinata frequenza nominale. Ma tale ipotetica situazione può essere ovviamente solo approssimata, data la pratica impossibilità di realizzare il <generatore teorico>; alla luce di questa constatazione si comprende come l'impiego di generatori indipendenti conduca di fatto a strutture *plesiocrone*.

Se, per converso, si intende attuare la sincronizzazione *trasferendo* i cronosegnali, i tempi di propagazione "non costanti" lungo il collegamento di interconnessione, renderanno i due cronosegnali, disponibili in due nodi collegati, identici nei valori medi delle frequenze ma di fase diversa e variabile (con effetti anche sulla frequenza istantanea).

Prescindendo però dalle valutazioni di natura tecnico-economica, in base alle quali orientare concretamente le scelte riguardanti le strategie di sincronizzazione di una rete complessa, le considerazioni sin qui sviluppate fanno emergere un carattere essenziale del problema della sincronizzazione. Anche quando ci si riferisca ad una rete totalmente numerica, ed a dispositivi operanti totalmente secondo criteri logici ed attraverso stati finiti, la sincronizzazione si sviluppa *nel continuo* nei suoi aspetti causali; gli effetti di dissincronie si tradurranno, alla fine, in inconvenienti rilevabili nel funzionamento logico e discreto dei dispositivi numerici, ma ciò non significa che appartengano all'area della logica numerica gli strumenti con cui trattare concretamente il problema, e la stessa metrologia per qualificare i dispositivi impiegati.

Il concetto esposto risulta forse più chiaro considerando gli effetti delle dissincronie: se, come si è detto, gli inconvenienti sopra richiamati hanno carattere discreto in relazione alla natura numerica dei dispositivi interessati, la frequenza, con cui tali inconvenienti si verificano, varia *nel continuo*, rispecchiando appunto il carattere di continuità delle cause, cioè l'entità effettiva delle dissincronie.

Questa osservazione riconduce agli aspetti ingegneristici della sincronizzazione, con le ovvie implicazioni tecnico economiche già in parte richiamate; e ciò soprattutto con riferimento ai sistemi plesiocroni.

Il quesito che il progettista si deve porre concretamente è: qual è l'entità della dissincronia accettabile, e la conseguente frequenza di perdita di sincronismo logico che ne deriva, ai fini della funzionalità complessiva del singolo impianto e dell'intera rete? E' infatti ben evidente che una perdita di sincronismo, che causi un deterioramento della qualità dei servizi forniti di *media gravità*, può essere accettabile se si verifica, ad esempio, con frequenza media annuale, ma non se ha frequenza mensile o oraria!

Il problema, a questo punto, si amplia ulteriormente, dato che l'attenzione del progettista è richiamata dal carattere aleatorio dei fenomeni di dissincronia e delle relative cause, ma anche dalla combinazione probabilistica di tali fenomeni nell'ambito di una rete, in cui i punti, nei quali si originano dissincronie, possono essere numerosi.

Su questo argomento si tornerà successivamente e nei testi del ciclo; è invece necessario chiarire subito che, quanto esposto, è valido pur prescindendo dall'analisi fine delle forme d'onda che materializzano i cronosegnali primari (cioè quelli da cui si parte per ricostruire gli intervalli di tempo e le relative trame per il riconoscimento del valore dei simboli ricevuti) e i sincronosegnali utilizzati nei processi specifici e ancora, ovviamente, i segnali da trattare.

6. L'ingegneria della sincronizzazione: effetti delle perdite di sincronismo, sensibilità dei processi, tempi di recupero

Con le tecniche numeriche -nella trasmissione e nella commutazione- il problema della sincronizzazione, oltre ad interessare gli apparati di trasmissione acquisisce fondamentale importanza nella commutazione, data la necessità di operare su flussi numerici originati in vari punti della rete, con modalità non necessariamente sincrone. Così, mentre l'esigenza di sincronizzazione si generalizza investendo parti sempre più ampie della rete, ed al limite l'intera rete (non solo quella nazionale!), diviene rilevante, data la natura dei segnali numerici, il controllo del posizionamento temporale dei segnali -tipicamente impulsivi- dei vari flussi numerici che pervengono in un apparato di moltiplicazione o di commutazione. Tale controllo, come si è detto, può essere necessario anche quando la frequenza di cifra dei diversi flussi sia *mediamente* identica nel tempo.

Si è anche osservato che, nelle tecniche in questione, il *significato* di ogni singolo bit dipende dalla sua collocazione in gruppi di bit, in *parole*, ed il significato di queste dalla relativa collocazione in *trame*. Da ciò la

necessità di riconoscere tale collocazione. A questo proposito è da ricordare, per inciso, un caso, di notevole interesse pratico, in cui il riconoscimento del significato del bit è forse ancor più complesso: quello di flussi ottenuti con la tecnica di moltiplicazione *bit per bit*, in cui deve essere possibile risalire al flusso tributario cui ciascun bit appartiene.

Si è ritenuto opportuno tornare su questi concetti, già anticipati, per trarre il problema della sincronizzazione di una rete numerica dal punto di vista dell'ingegneria dei sistemi. In questo quadro, infatti, è necessario tenere conto della sensibilità dei singoli processi alla inevitabile imperfezione dei cronosegnali ma, simultaneamente, dei problemi tecnico-economici collegati al miglioramento dei generatori ed alla realizzazione dei dispositivi interessati alla sincronizzazione. Come in ogni progettazione di sistema è necessario ripartire tra vari dispositivi tolleranze e ed affidabilità, tenendo conto, in questo caso, della criticità *sui servizi* delle perdite di sincronismo, e dei tempi di recupero quando, per qualche motivo, si sia perduto.

In sintesi, nei sistemi numerici, il problema della sincronizzazione è molto più articolato che nei sistemi analogici; non si esaurisce infatti nel controllo della frequenza e della fase dei cronosegnali, secondo le esigenze dei vari livelli gerarchici di moltiplicazione, ma interessa anche strutture numeriche (cioè gruppi di bit) più complesse; ciò ha, tra l'altro, significativi effetti sui tempi di recupero dell'allineamento -di trama, supertrama ecc.- nel caso in cui si sia perduta la sincronizzazione. Ed è ovvio che la *tollerabilità* di tali tempi debba essere valutata, come già accennato, in relazione agli effetti sui servizi erogati e sulle altre funzioni presenti nella rete, e necessarie alla sua operatività.

Prescindendo da un esame storico della evoluzione delle esigenze di sincronizzazione nelle reti di telecomunicazioni, è oggi possibile impostare il problema della sincronizzazione in termini piuttosto generali, fermo restando che sarà poi comunque necessario studiarlo analiticamente nei singoli aspetti, per la predisposizione delle specifiche soluzioni.

7. La sincronizzazione delle reti: probabilità ed entità di cause e di effetti

Si è detto che i cronosegnali dei sistemi plesiocroni devono essere così *prossimi* ai valori nominali da determinare, nella cooperazione tra sistemi diversi, inconvenienti di dissincronia molto rari e/o tollerabili.

In questa osservazione, e in generale in quanto esposto, ricorrono aggettivi come adeguato, sufficiente, tollerabile, che potrebbero dare adito ad un'interpretazione qualitativa del problema della sincronizzazione. Ogni trattamento

richiede, invece, un cronosegnale di specifiche caratteristiche che devono essere individuate e descritte in termini quantitativi: presenta, in altre parole, una *sua tolleranza* rispetto alle caratteristiche in questione.

D'altra parte, le esigenze di ogni singolo trattamento, nei confronti dei cronosegnali, devono essere stabilite anche in relazione agli effetti sul funzionamento di una rete complessa, cui concorrono altri trattamenti ed altri cronosegnali: non è difficile immaginare come le caratteristiche statistiche dei cronosegnali e dei trattamenti interagiscano tra loro, combinandosi con modalità statistiche: ciò può avvenire non solo lungo *catene di tolleranza* ma anche in *loop chiusi*, con il possibile manifestarsi di fenomeni di instabilità.

In sostanza il problema della sincronizzazione, oltre ad essere tipicamente quantitativo, richiede per lo più un approccio probabilistico, strettamente correlato ai caratteri aleatori dei segnali, così come sono realmente prodotti da *sorgenti fisiche*, o come divengono in relazione alle modalità con cui sono processati; caratteri che possono essere individuati e valutati solamente attraverso un'analisi fine, ricorrendo, tipicamente, alla configurazione di modelli rappresentativi dei fenomeni reali anche sotto un profilo statistico.

Secondo una logica generale, le caratteristiche dei cronosegnali *necessarie*, e dei dispositivi atti a generarli ed a trasferirli, devono essere fissate in relazione alle prestazioni finali da raggiungere, che sono quelle complessive della rete di telecomunicazione; di fatto, dalle prestazioni *da raggiungere*, quelle dei *servizi erogati*, si deve risalire alle caratteristiche dei cronosegnali.

Si deve tuttavia osservare che, a causa del carattere complesso ed integrato delle reti di TLC, possono essere piuttosto numerosi gli elementi delle reti stesse che il processo logico sopra richiamato deve considerare e che, con le proprie caratteristiche statistiche, partecipano alla determinazione del risultato finale (ad esempio alla qualità del segnale restituito dal trasduttore di ricezione); inoltre, non è per lo più univoco il modo con cui ripartire, tra i vari elementi, i margini di tolleranza complessivamente disponibili in relazione al risultato da ottenere.

Per questo motivo, le scelte da operare non possono fondarsi solamente su considerazioni di topologia e di architettura, ma investono l'area delle tecnologie, con le ovvie implicazioni di natura tecnico-economica. Si deve far riferimento, tra l'altro, a dispositivi e macchine *fisicamente realizzate* di cui si possano anche ipotizzare, con ragionevole approssimazione, i costi industriali.

Pur prescindendo da questo complesso problema, è da sottolineare che, per risalire alle esigenze di sincronizzazione di un sistema di TLC, si deve esaminare ogni parte in cui si presentino criticità nei confronti della sincronizzazione stessa. Le criticità in questione, per quanto richiamato, non sono solamente

quelle che si manifestano nello specifico dispositivo in esame, ma anche quelle indotte nel suo funzionamento dalle altre parti del sistema interessate al trattamento del segnale.

Nel complesso, è da rilevare che lo studio di una strategia di sincronizzazione si gioca simultaneamente su catene di grandezze definite statisticamente e modellizzabili in processi aleatori, e su fattori tecnico-economici: uno studio di alta sistemistica, che si complica ulteriormente quando si debba tenere conto dei *fenomeni transitori*: in pratica, delle fasi di perdita e di recupero della sincronizzazione (e degli effetti sui servizi erogati), e dei problemi di indisponibilità conseguenti a guasti; è questo un aspetto di importanza primaria, e crescente al crescere della centralizzazione nella generazione dei cronosegnali, fortemente influenzato dalla affidabilità delle reti di distribuzione (dei cronosegnali) sino ai punti di utilizzazione.

8. Conclusione

In relazione alla potenza ed alla varietà degli strumenti tecnologici oggi disponibili, può sorprendere il fatto che un problema non nuovo, come la sincronizzazione di una rete numerica di dimensione nazionale, sia ancora argomento di studio sul piano delle metodologie, ed oggetto di programmi di realizzazione.

Una spiegazione di tale situazione emerge da un'analisi dei vari problemi specifici in cui l'argomento si concretizza, e che si è appunto cercato di richiamare.

Il quadro è così complesso che un tentativo di descriverlo secondo una classificazione ragionevolmente sintetica appare inadeguato; ciò dà ragione di una certa varietà tassonomica che si riscontra anche nei testi specialistici.

Nel ciclo di articoli che prende avvio in questo numero del Notiziario si farà riferimento ai vari problemi richiamati, nell'intesa che, successivamente, alcuni testi saranno dedicati allo specifico argomento delle interrelazioni tra tali problemi nell'ambito di una rete complessa; questo argomento è, di fatto, tra i meno approfonditi nella letteratura tecnica consolidata.

La sincronizzazione nelle reti numeriche di trasporto

Parte prima: le reti di trasporto negli anni '80⁽¹⁾

M. Carbonelli, D. De Seta, D. Perucchini (*)

A partire dalla seconda metà degli anni '80 si è sviluppata un'intensa attività di ricerca e standardizzazione riguardante i sistemi e le reti basate sulla gerarchia di moltiplicazione numerica sincrona SDH. Tale sviluppo è stato favorito dalla concreta necessità di dare risposta alle crescenti esigenze, manifestate sia dall'utenza sia dai gestori di reti e servizi, di maggiore capacità e flessibilità delle reti numeriche pubbliche e private. Il progetto della nuova gerarchia di moltiplicazione ha dato l'occasione di riconsiderare, sia in ambito scientifico che in ambito normativo, i problemi connessi con la distribuzione dei segnali di sincronismo: in questo articolo, suddiviso in tre parti, si vuole tentare una panoramica dei più importanti aspetti connessi con la sincronizzazione delle reti numeriche, al fine di fornire ad un lettore non specialista in materia gli strumenti indispensabili per una buona comprensione dell'argomento. In questa prima parte, dopo aver richiamato la terminologia e i concetti fondamentali riguardanti i segnali numerici, si affrontano, anche da un punto di vista storico, i principali problemi connessi con la sincronizzazione delle reti numeriche di trasporto basate sull'attuale gerarchia plesiocrona. Nella seconda parte, l'esame di tali problematiche di sincronizzazione è esteso alle future reti basate sulla gerarchia numerica sincrona. Nella terza parte, che verrà pubblicata in un prossimo numero del Notiziario, vengono presentati, ad un livello molto generale, lo stato dell'arte delle strategie di sincronizzazione e i metodi per la caratterizzazione dei dispositivi di generazione e rigenerazione di cronosegnali di elevata qualità; infine, dopo la descrizione della strategia di sincronizzazione delle centrali numeriche di commutazione della rete nazionale, vengono forniti alcuni cenni sullo stato della normativa sulla sincronizzazione delle reti di trasporto basate sull'SDH.

1. Introduzione

Nel corso degli ultimi anni si è sviluppata una formidabile attività di ricerca e standardizzazione riguardante i sistemi e le reti basate sulla gerarchia di moltiplicazione numerica sincrona, universalmente nota come SDH (Synchronous Digital Hierarchy). Tale sviluppo è stato favorito dalla concreta necessità di dare risposta alle crescenti esigenze, manifestate sia dall'utenza sia dai gestori di reti e servizi, di maggiore capacità e flessibilità delle reti numeriche pubbliche e private, a fronte di evidenti e sostanziali limitazioni intrinseche delle strutture degli apparati e delle potenzialità gestionali delle attuali reti.

Il progetto dell'SDH ha dato l'occasione di riconsiderare tutta una serie di problemi inerenti al funzionamento dei sistemi numerici di trasmissione, e di concepire nuove soluzioni che sfruttano le conoscenze raggiunte in anni di esperienza in campo sul comportamento di apparati e sistemi. Tanto per citare uno degli aspetti fondamentali della nuova gerarchia, che riflette l'intento di fare tesoro dei passati errori, ricordiamo la grande profusione di byte di overhead previsti nella trama SDH in misura largamente superiore alle necessità individuate al momento del progetto, al fine di poter fronteggiare future nuove richieste di canali di servizio, funzionali alle esigenze del gestore o dell'utente.

(*) ing. Marco Carbonelli, ing. Domenico De Seta, ing. Daniele Perucchini - Fondazione Ugo Bordoni - Roma

(1) Lavoro eseguito nell'ambito della Convenzione in atto tra l'Amministrazione delle Poste e Telecomunicazioni e la Fondazione Ugo Bordoni.

Attualmente uno dei più importanti problemi allo studio riguardanti l'introduzione dell'SDH è il progetto della rete di sincronizzazione richiesta al fine di utilizzare pienamente le potenzialità della nuova gerarchia. In particolare occorre scegliere un'opportuna strategia di sincronizzazione, tenendo anche conto dei problemi connessi con la transizione dalle attuali reti di trasporto basate sui sistemi plesiocroni a quelle basate sull'SDH: ed è appunto tale legame tra PDH (Plesiochronous Digital Hierarchy) ed SDH a porre alcuni seri problemi sia al progettista degli apparati sia al progettista della rete di sincronizzazione.

In questo articolo, suddiviso in tre parti, traendo spunto dal rinnovato interesse per i problemi di sincronizzazione portato dall'introduzione dell'SDH, si vuole tentare una panoramica dei più importanti aspetti connessi con la sincronizzazione delle reti numeriche, al fine di fornire ad un lettore non specialista in materia gli strumenti indispensabili per una buona comprensione dell'argomento. Questa prima parte è dedicata alla rete numerica di trasporto che si è venuta sviluppando nel corso degli anni '80: dopo un breve richiamo dei concetti fondamentali e della terminologia riguardante i segnali numerici (par.2), con particolare enfasi sulle proprietà degli associati cronosegnali, vengono analizzati i problemi di sincronizzazione nelle reti numeriche di trasporto basate sulla gerarchia di moltiplicazione plesiocrona PDH (par.3). Nella seconda parte, sono illustrate le principali caratteristiche della nuova gerarchia di moltiplicazione sincrona SDH e degli apparati basati su tale standard: successivamente, viene descritta la struttura tipica di una rete di trasporto sincrona con particolare attenzione ai problemi legati alla sua sincronizzazione e alla prevista interconnessione con la rete PDH. Nell'ultima parte, che verrà pubblicata in un prossimo numero del Notiziario, vengono presi in esame gli aspetti più importanti del progetto di una rete di distribuzione di cronosegnali per la sincronizzazione di reti di telecomunicazione, considerando le possibili strategie e i dispositivi per mezzo dei quali si realizza la distribuzione del cronosegnale; infine, vengono introdotti i parametri e i criteri comunemente utilizzati per valutare la qualità degli orologi e della rete di sincronizzazione, dando qualche cenno sugli attuali orientamenti della normativa internazionale.

2. Richiami sui segnali numerici

In questo articolo, dedicato come si è detto alla sincronizzazione delle reti numeriche, si farà frequentemente riferimento a segnali numerici. Pur ritenendo il lettore familiare con il concetto di segnale numerico, è opportuno tentare una definizione che aiuti a meglio impostare le necessarie considerazioni sui sistemi di trasmissione e di commutazione che trattano segnali numerici.

Un *segnale numerico* è costituito dalla somma di segnali elementari che si susseguono nel tempo, a ciascuno dei quali è associata un'informazione attraverso una delle caratteristiche del segnale elementare stesso (ampiezza, durata, forma d'onda, etc.): aspetto essenziale dei segnali numerici è la natura *discreta* dell'informazione associata. Esempi di segnali numerici generati da sorgenti intrinsecamente discrete sono quelli scambiati tra calcolatori, i segnali di tipo telex (teletext, videotel, televideo, etc.), la segnalazione telefonica: in tutti questi casi l'insieme dei simboli che la sorgente può emettere è limitato e noto. Laddove l'effettiva sorgente di informazione sia intrinsecamente analogica e tempo-continua è necessario, al fine di generare un segnale numerico, prevedere le operazioni di campionamento e quantizzazione. Con l'operazione di *campionamento* il segnale generato dalla sorgente viene osservato in corrispondenza ad istanti che si susseguono nel tempo (*istanti di campionamento*) al fine di generare una sequenza di valori (*campioni del segnale*): si realizza così la discretizzazione del segnale sull'asse dei tempi. La successiva *quantizzazione* dei campioni così ottenuti consiste nel rappresentare tutti i valori compresi in un assegnato intervallo dell'asse delle ampiezze con un unico valore convenzionale, detto *livello di quantizzazione*, realizzando così la discretizzazione del segnale sull'asse delle ampiezze. Esempi ben noti di segnali numerici generati a partire da sorgenti intrinsecamente analogiche sono il segnale telefonico numerico a 64 kbit/s, i segnali prodotti dai codec televisivi a 34 Mbit/s, e dai codec musicali a 384 kbit/s.

Occorre sottolineare come per un segnale numerico sia di fondamentale importanza l'individuazione degli istanti caratteristici in cui hanno significato i valori del segnale ai fini dell'associazione e del riconoscimento dell'informazione in esso contenuta. Da questo punto di vista un segnale numerico può essere classificato, con riferimento alla tipologia di sincronismo in esso presente, come segnale isocrono o segnale anisocrono. Un segnale numerico è detto *isocrono* quando gli intervalli di tempo fra istanti caratteristici consecutivi hanno la stessa durata o durate che sono multipli interi della durata più breve: in pratica sono ammesse variazioni, di entità massima specificata, della durata di detti intervalli. Quando tali condizioni non sono necessariamente verificate il segnale è detto *anisocrono*. I segnali isocroni sono di prevalente interesse nell'ambito delle tecniche numeriche di trasmissione e commutazione, e in particolare si considerano quasi esclusivamente quelli in cui gli intervalli tra istanti caratteristici consecutivi hanno la stessa durata. Al fine di rappresentare in maniera esplicita la successione degli istanti caratteristici di un segnale numerico isocrono è utile introdurre il concetto di *cronosegnale*: idealmente questo è un segnale avente una forma d'onda particolarmente semplice, in

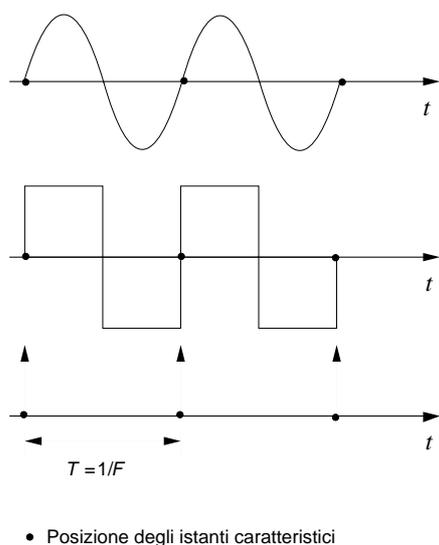


Figura 1 Possibili realizzazioni di un cronosegnale: a) cronosegnale sinusoidale; b) cronosegnale ad onda quadra; c) cronosegnale impulsivo

cui gli istanti caratteristici sono associati ad “eventi” facilmente individuabili sul segnale. In Figura 1 sono mostrate tre possibili realizzazioni di un cronosegnale, tra le più comunemente utilizzate nell’analisi dei sistemi numerici: il *cronosegnale sinusoidale*, in cui gli istanti caratteristici sono rappresentati dagli attraversamenti dello zero con pendenza positiva (o negativa) di una sinusoide; il *cronosegnale ad onda quadra*, nel quale gli istanti significativi cadono in corrispondenza al fronte di salita (o discesa); il *cronosegnale impulsivo*, i cui istanti caratteristici sono rappresentati dalla posizione di impulsi matematici.

La grandezza fondamentale che caratterizza un cronosegnale è la frequenza nominale $F=1/T$ dove T è il periodo di tempo che intercorre tra istanti caratteristici successivi: in un cronosegnale ideale la frequenza è costante nel tempo, cioè il cronosegnale è rigorosamente periodico.

Nel campo delle telecomunicazioni si indica col termine *orologio* un dispositivo in grado di generare i cronosegnali necessari al funzionamento degli apparati: nella terza parte dell’articolo verranno dati dei brevi cenni sulle tipologie di orologi utilizzati.

Gli orologi reali sono in grado di generare solo cronosegnali pseudo-periodici. Un modello matematico che descrive un cronosegnale sinusoidale $s(t)$ pseudo-periodico è

$$s(t) = A \sin(\Phi(t)),$$

dove $\Phi(t)$ è la fase totale istantanea. Per definizione, la *frequenza istantanea* $f(t)$ del cronosegnale è data da

$$f(t) = \frac{1}{2\pi} \frac{d\Phi(t)}{dt}.$$

Tale modello, come vedremo in seguito, risulta di grande utilità per caratterizzare la precisione e la stabilità degli orologi nel dominio della frequenza e nel dominio del tempo.

Nella pratica i cronosegnali che vengono utilizzati nei sistemi numerici sono inevitabilmente affetti da fluttuazioni della posizione temporale degli istanti significativi: tale fenomeno è denominato *jitter* con riferimento alle fluttuazioni veloci, e *wander* per quelle lente, laddove il confine che separa i due fenomeni è fissato in maniera convenzionale intorno alla frequenza di 10 Hz. Il wander costituisce, come vedremo in seguito, un problema particolarmente importante nella sincronizzazione delle reti numeriche e richiede metodi d’analisi e contromisure “ad hoc”, che si differenziano profondamente da quelle adottate per il jitter.

Quanto detto sinora fa riferimento alla tipologia di sincronismo di un singolo segnale numerico: importanti considerazioni nascono nell’esaminare la relazione tra i cronosegnali associati a due diversi segnali numerici. A tale riguardo si possono individuare quattro importanti casi, descritti come segue:

- segnali *sincroni*: due segnali numerici si dicono sincroni quando la fase tra i corrispondenti istanti caratteristici dei cronosegnali associati è costante;
- segnali *mesocroni*: due segnali numerici si dicono mesocroni se i rispettivi cronosegnali associati hanno esattamente la stessa frequenza media;
- segnali *plesiocroni*: due segnali numerici si dicono plesiocroni se i rispettivi cronosegnali associati hanno frequenze che differiscono da un assegnato valore, detto frequenza nominale, entro limiti specificati;
- segnali *eterocroni*: due segnali numerici si dicono eterocroni quando i rispettivi cronosegnali associati hanno frequenze nominali diverse.

Per completezza d’informazione sulla terminologia corrente occorre ricordare che con il termine *asincrono* si intende genericamente riferire una condizione di assenza di sincronismo tra segnali.

Nel campo delle reti numeriche, i due tipi di segnale di maggiore interesse pratico sono i segnali mesocroni e quelli plesiocroni. Infatti, la condizione di perfetto sincronismo implicita nella definizione data al punto a) non è praticamente realizzabile tra segnali trasportati in una rete geograficamente estesa, a causa delle variazioni di fase accumulate lungo i collegamenti tra i nodi della rete. Quindi, segnali che nascono in un nodo di rete rigorosamente sincroni, generati cioè a partire da un stesso orologio, possono al termine del loro percorso nella rete risultare fortemente degradati dal punto di vista delle relazioni di sincronismo, pur mantenendo nel lungo termine la caratteristica di avere la stessa frequenza media. I principali meccanismi che causano tali degradazioni nelle reti numeriche sono:

- le fluttuazioni del tempo di propagazione dovute ad escursione termica dei mezzi trasmissivi;
- i meccanismi di giustificazione presenti nelle tecniche di multiplazione;
- i meccanismi di estrazione della temporizzazione richiesti per la rigenerazione dei segnali numerici;
- le instabilità intrinseche nel funzionamento degli orologi asserviti.

In una rete numerica i segnali comunemente denominati 'sincroni' devono perciò intendersi più propriamente 'mesocroni'.

3. Sincronizzazione nella rete di trasporto degli anni '80

La struttura raggiunta dalle attuali reti pubbliche di telecomunicazione nei paesi a più elevato sviluppo economico e tecnologico è il risultato di una lunga evoluzione iniziata alla fine del secolo scorso. Senza alcuna pretesa di voler tentare una storia della telefonia, le tappe di questa evoluzione possono essere individuate nei seguenti punti:

- 1880 realizzazione dei primi collegamenti telefonici a breve distanza su linee aeree in rame, commutati manualmente da operatore mediante l'uso di "cordoni telefonici";
- 1892 prima centrale automatica di commutazione in grado di sostituire l'operatore per l'istadamento del traffico locale;
- 1900 estensione geografica delle reti telefoniche, resa possibile dall'uso della pupinizzazione dei cavi;
- 1915 impiego dell'amplificazione, che consente la copertura geografica dei continenti;
- 1930 introduzione della tecnica di multiplazione a divisione di frequenza (FDM) per la trasmissione simultanea di più canali su uno stesso portante nella rete di giunzione a media e lunga distanza;
- 1957 introduzione dei cavi sottomarini amplificati e dei satelliti, che rendono possibile la rete mondiale;
- 1962 introduzione commerciale delle tecniche di multiplazione a divisione di tempo (TDM) e sviluppo dei primi sistemi di trasmissione basati sulla modulazione ad impulsi codificati (PCM);
- 1975 introduzione delle prime centrali numeriche di commutazione e sviluppo della gerarchia numerica di multiplazione plesiocrona (PDH).

L'introduzione dei cosiddetti ripartitori automatici numerici (RED), già avviata a partire dalla seconda metà degli anni ottanta, e il successivo sviluppo della gerarchia numerica sincrona (SDH) renderanno possibile l'automatizzazione delle operazioni di permutazione dei circuiti numerici dedicati e dei flussi numerici della rete pubblica commutata. Nel

seguito verranno approfonditamente analizzate le caratteristiche e la struttura della rete di trasporto focalizzando l'attenzione, in particolare, sulle problematiche di sincronizzazione identificabili nelle diverse fasi dell'evoluzione che tale rete ha subito negli ultimi venti anni.

Da questa breve rassegna delle innovazioni tecnologiche che hanno determinato l'enorme sviluppo delle reti telefoniche, emerge come tale processo evolutivo abbia sempre ruotato intorno alle due funzioni fondamentali di una rete telefonica: la trasmissione e la commutazione. La *trasmissione* è l'insieme delle tecniche che consentono di trasferire informazione tra due nodi di rete posti a sensibile distanza l'uno dall'altro. La *commutazione* è l'insieme delle tecniche che consentono di assegnare in modo dinamico i canali trasmissivi di una rete al fine di interconnettere gli utenti sulla base delle loro richieste.

Prima dell'introduzione delle tecniche numeriche l'impatto di ogni innovazione ha di volta in volta interessato una o l'altra di tali funzioni primarie, mantenendone chiaramente distinte la natura e l'utilizzazione negli apparati presenti in rete: l'introduzione delle tecniche FDM, per citare un esempio, ha enormemente potenziato le capacità trasmissive delle reti di giunzione, non influenzando affatto le caratteristiche e le funzionalità delle centrali di commutazione. Con l'avvento delle tecniche numeriche e particolarmente con l'impiego della multiplazione a divisione di tempo dei canali telefonici, introdotta per meglio utilizzare le capacità trasmissive proprie dei portanti metallici e hertziani, ha avuto inizio un processo di graduale integrazione delle tecniche di trasmissione e commutazione. La struttura di trama del segnale multiplo PCM a 2,048 Mbit/s, costituita dall'interallacciamento di trentadue ottetti, di cui trenta dedicati al trasporto di canali telefonici a 64 kbit/s, rende possibile sfruttare anche per le operazioni di commutazione il principio della divisione di tempo. Nasce così l'*autocommutatore numerico* che fonda il suo principio di funzionamento sull'operazione di trasferimento di un campione di segnale, in forma di otetto di cifre binarie (*byte*), da un intervallo di tempo (*time slot*) ad un altro di due trame distinte: una trama entrante e una uscente dall'autocommutatore, rispettivamente. Una descrizione dettagliata delle tecniche utilizzate negli autocommutatori numerici esula dagli scopi di questo articolo: ci limiteremo a richiamare solo alcuni aspetti fondamentali che hanno un impatto determinante sulle prestazioni all'interfaccia con i sistemi trasmissivi.

Per poter effettuare la commutazione numerica occorre portare tutti i flussi entranti ad avere una stessa frequenza di simbolo, e quindi una stessa frequenza di otetto, derivata dall'orologio dell'autocommutatore

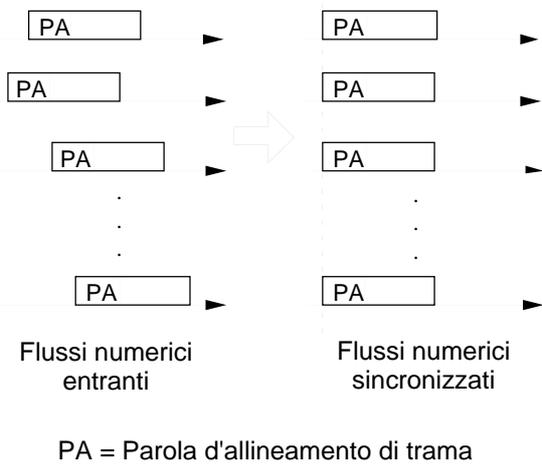


Figura 2 Sincronizzazione di trama nelle unità di linea delle centrali numeriche

(orologio di centrale o locale), cosicché sia possibile rendere le trame temporalmente allineate (in fase), come illustrato in Figura 2.

Tali operazioni vengono effettuate in un sistema di terminazione della centrale denominato *unità di linea* (o anche *terminale di centrale*): lo schema a blocchi esemplificativo delle funzioni svolte nell'unità di linea è mostrato in Figura 3.

Al lato trasmittente, sul flusso binario di ottetti forniti in uscita dal multiplexore di centrale, vengono effettuate le seguenti funzioni:

- 1) inserzione della segnalazione nel time slot 16;
- 2) inserzione della parola di allineamento nel time slot 0;
- 3) codifica di linea (da binario unipolare a HDB3).

Al lato di ricezione il segnale di linea a 2,048 Mbit/s è sottoposto alle seguenti operazioni:

- 4) decodifica di linea (da HDB3 a binario unipolare);
- 5) estrazione della temporizzazione di bit;
- 6) allineamento e sincronizzazione di trama;
- 7) estrazione della segnalazione⁽²⁾.

Senza entrare nel merito delle singole funzioni, possiamo limitarci, per le considerazioni che interessano la sincronizzazione di rete, ad esaminare più in dettaglio i punti 5) e 6).

3.1 Sincronizzazione nelle unità di linea delle centrali numeriche

Come si è detto, le funzioni di elaborazione svolte negli autocommutatori numerici sono controllate da un cronosegno di centrale fornito da un orologio locale. Perché la trasmissione e la commutazione dei segnali numerici afferenti alla centrale abbiano luogo in maniera da garantire l'integrità dell'informazione da essi trasportata è necessario che i cronosegnali del multiplex e della centrale siano mesocroni, cioè, secondo la definizione adottata in CCITT, sia verificata l'uguaglianza delle rispettive frequenze medie a lungo termine. Per chiarire meglio il significato e le implicazioni di tale requisito consideriamo dapprima la semplice situazione schematizzata in Figura 4, in cui una coppia di multiplex PCM è collegata "back-to-back" tramite un sistema di linea numerico bidirezionale.

(2) L'inserzione e l'estrazione della segnalazione può non essere richiesta nel caso d'impiego della segnalazione su canale comune (CCITT n. 7).

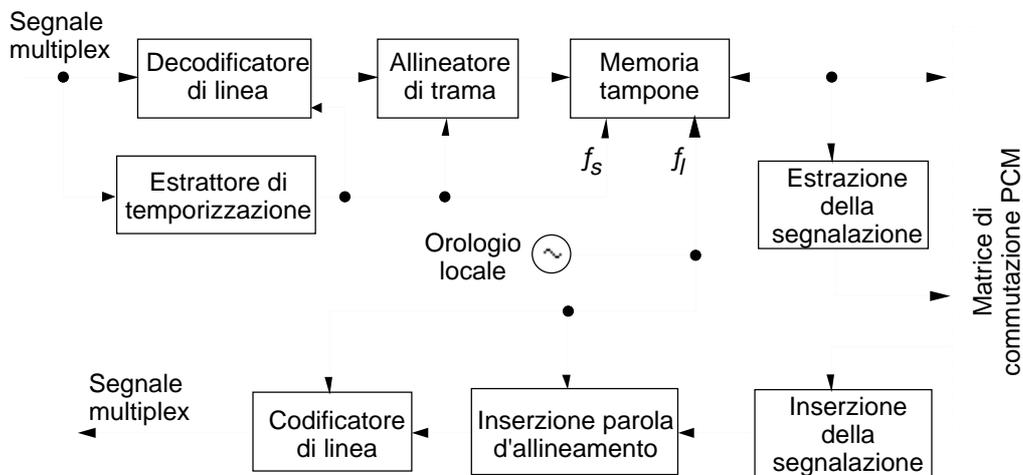


Figura 3 Schema a blocchi esemplificativo dell'unità di linea di una centrale numerica

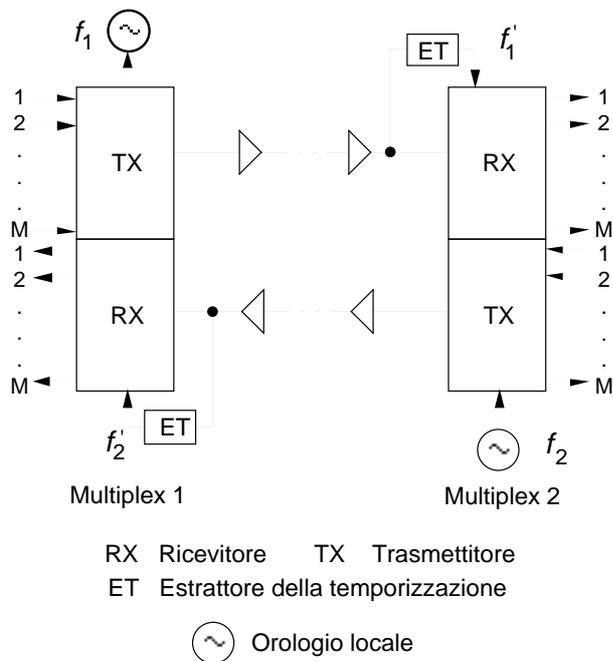


Figura 4 Schema di temporizzazione dei multiplex PCM

Ognuno dei due multiplex utilizza un suo orologio locale per generare le temporizzazioni necessarie alla moltiplicazione, mentre ricava dal segnale multiplex proveniente dal terminale lontano quelle richieste per

demoltiplicare i tributari in ricezione: in questo schema, quindi, le sezioni di moltiplicazione e demoltiplicazione possono funzionare correttamente in maniera plesiocrona, cioè con due orologi indipendenti aventi frequenze f_1 e f_2 prossime alla frequenza nominale f_0 entro una assegnata tolleranza Δf . Se si passa a considerare il caso di Figura 5, in cui i due multiplex sono interconnessi tramite una centrale numerica per consentire la commutazione dei tributari affasciati, si osserva subito che un problema di sincronizzazione nasce all'interfaccia tra multiplex e centrale, perché, come si è detto, nella centrale occorre utilizzare un unico cronosegnaie (a frequenza f_c) per temporizzare i flussi in ingresso e in uscita alla rete di commutazione.

Una prima esigenza che si manifesta è quella di prevedere un'adeguata memorizzazione dei flussi entranti al fine di assorbire le inevitabili differenze di fase con cui le relative trame giungono in centrale, sia a causa dei differenti percorsi di propagazione sia perché nei multiplex non c'è in generale alcuna preassegnata relazione di fase fra le trame in ricezione e in trasmissione. Nello schema a blocchi di principio dell'unità di linea mostrato in Figura 3 questa operazione di "rifasatura" delle trame entranti è realizzata nella memoria tampone (*buffer*), che, come indicato, viene scritta con una frequenza f_s derivata dal cronosegnaie estratto dal flusso ricevuto e letta con una frequenza f_l ricavata dall'orologio di centrale. Una sia pur piccola differenza tra le frequenze f_s e f_l causa una periodica perdita o ripetizione di ottetti in transito nella memoria

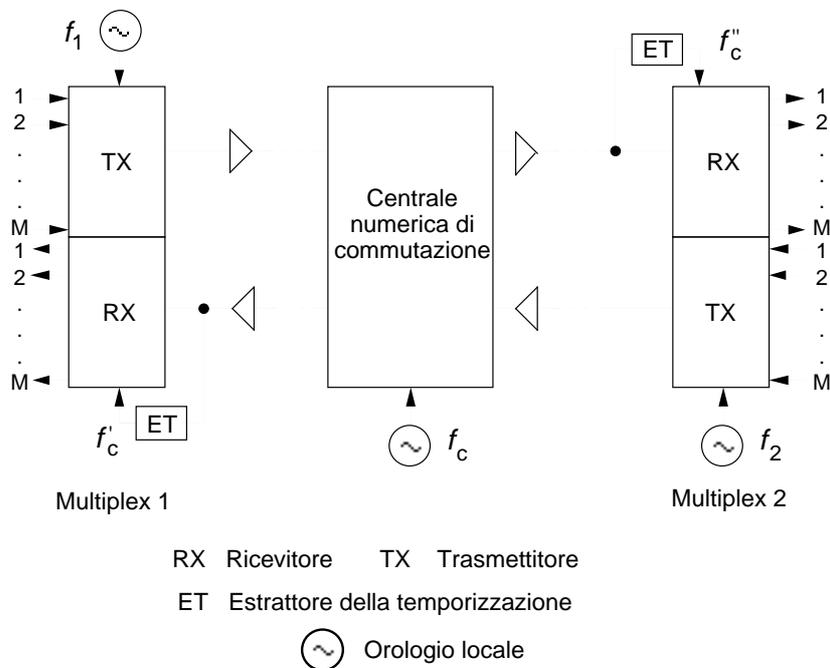


Figura 5 Temporizzazione di multiplex collegati tramite una centrale numerica

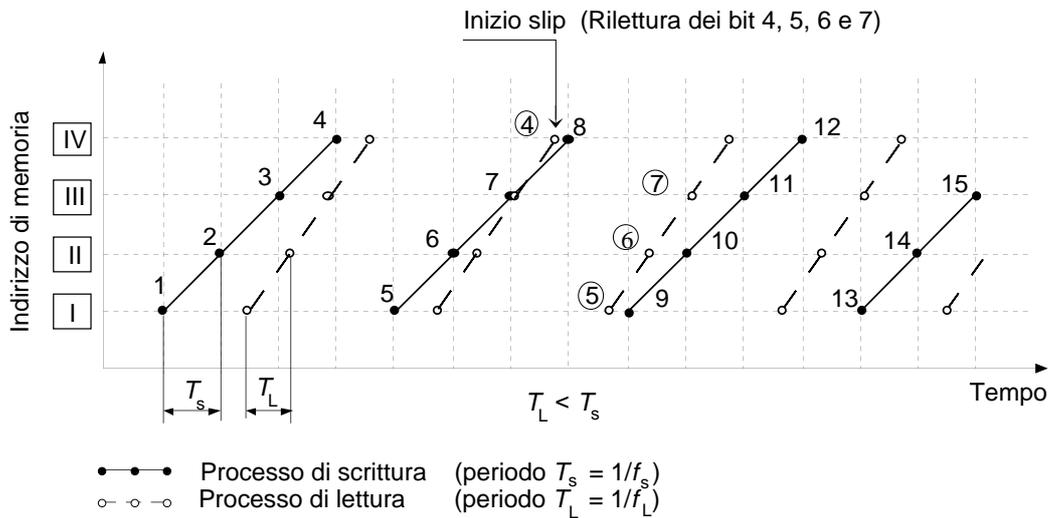


Figura 6 Diagramma esemplificativo del meccanismo di formazione degli slip in una memoria elastica a quattro celle

tampone: questi eventi prendono il nome di *slip di sincronizzazione*. In Figura 6 è illustrato il caso di svuotamento di una memoria tampone a quattro celle (bit) in cui l'orologio di lettura è più veloce di quello di scrittura, cosicché periodicamente si verifica la rilettura di quattro bit consecutivi.

Il tasso di slip F_{slip} è funzione del numero di N di celle (bit) di memoria e della differenza tra le frequenze di lettura e di scrittura. Esprimendo queste ultime in hertz, la relazione che fornisce il tasso di slip è data da

$$F_{slip} = 86400 \frac{|f_1 - f_2|}{N} \quad (\text{slip / giorno})$$

Se per gli orologi di scrittura e lettura sono specificate le tolleranze relative Δf_s e Δf_l rispetto alla frequenza nominale f_0 , si può facilmente ricavare che il tasso di slip F_{slip} è limitato dalla disuguaglianza

$$F_{slip} \leq 86400 \frac{(\Delta f_s + \Delta f_l) f_0}{N} \quad (\text{slip / giorno})$$

L'impiego di una memoria sufficientemente grande consente di ridurre il tasso di slip entro limiti prefissati, per ogni assegnata precisione degli orologi. Se si assume che l'orologio di centrale abbia caratteristiche di precisione molto migliori di quelle dei multiplex ($\Delta f_j \ll \Delta f_s$), per i quali la tolleranza è fissata a 50 parti per milione ($\Delta f_s = 50 \cdot 10^{-6}$) nella Racc. G. 703 del CCITT e si prevede di memorizzare una sola trama⁽³⁾ ($N = 256$ bit), risulta che per $f_0 = 2,048$ MHz il tasso di

(3) Una dimensione della memoria elastica pari a una o più trame consente di ottenere *slip controllati*, cioè ripetizioni o perdite di un numero intero di trame: tale meccanismo garantisce il mantenimento dell'allineamento di trama, limitando le degradazioni conseguenti al verificarsi degli slip.

slip può raggiungere valori fino a 24 slip al minuto. Tale valore è assolutamente inaccettabile anche per i servizi che meglio sopportano la presenza di slip di centrale: tanto per fissare le idee, i requisiti di rete indicati come obiettivo dal CCITT stabiliscono, nella Racc. G.822 [10], una soglia di 5 slip/giorno, mentre un funzionamento caratterizzato da valori maggiori di 30 slip/ora è valutato inaccettabile.

Per ovviare a tale situazione, nel caso di Figura 5, è necessario sincronizzare i flussi numerici in ingresso al nodo di commutazione, asservendo all'orologio di centrale gli orologi presenti nelle sezioni di trasmissione dei multiplex ad essa collegate. Questo in effetti è solo il primo provvedimento di sincronizzazione richiesto per il corretto funzionamento di una rete integrata: è evidente infatti che problemi del tutto analoghi si presentano con riferimento alle relazioni mutue di fase tra più centrali interconnesse mediante sistemi numerici di trasmissione. Questo è un problema certamente ancor più complesso la cui soluzione costituisce in sostanza l'obiettivo primario delle tecniche di sincronizzazione della rete di commutazione.

3.2 Sincronizzazione nelle reti di trasporto basate sulla PDH

L'esigenza di sfruttare i supporti fisici dell'esistente rete di giunzione tra centrali di commutazione ha portato a sviluppare la cosiddetta gerarchia di moltiplicazione plesiocrona, che, partendo dal segnale multiplo a 2,048 Mbit/s, si è progressivamente articolata su tre ulteriori livelli (a 8,448 Mbit/s, 34,368 Mbit/s e 139,264 Mbit/s). Questo sviluppo della rete numerica di *trasporto*, termine con cui possiamo designare l'insieme delle risorse trasmissive destinate a

realizzare il trasferimento di informazione mediante flussi numerici con frequenza di cifra normalizzata, ha comportato al suo nascere una fondamentale scelta tra la tecnica di *multiplazione sincrona* e la tecnica di *multiplazione plesiocrona*. Nella prima si assume che i segnali tributari e il segnale multiplo siano sincroni tra loro: solo in questa condizione è possibile garantire l'integrità delle informazioni trasportate, limitare il tempo di ritardo attraverso il moltiplicatore-demoltiplicatore e ridurre la complessità dell'apparato stesso. Tuttavia la condizione di sincronismo ideale non è praticamente realizzabile anche laddove si utilizzi una rete di distribuzione della sincronizzazione: come vedremo nel seguito, quando si considererà la gerarchia numerica sincrona, il meglio che si riesce ad ottenere è una situazione di mesocronia tra i segnali numerici presenti in rete. Perciò, all'epoca in cui vennero sviluppati i primi sistemi di multiplazione di ordine superiore al primo livello gerarchico la complessità dei problemi connessi con la sincronizzazione di rete richiese dalle tecniche sincrone indusse a sviluppare la tecnica di multiplazione plesiocrona basata sull'adozione del meccanismo di giustificazione positiva. Come è noto, nella tecnica plesiocrona si ammette che i cronosegnali associati ai tributari e al segnale multiplo possano avere frequenze diverse dal loro valore nominale: prima di affasciare bit a bit i tributari è necessario perciò portarli a una stessa frequenza, superiore a quella nominale e sincrona

con il segnale multiplo, mediante un'operazione di ritemporizzazione svolta nel cosiddetto *sincronizzatore*. Questo dispositivo, come mostrato in Figura 7, comprende un buffer per ciascun tributario nel quale i bit sono scritti alla frequenza del cronosegnale di tributario e letti a una frequenza maggiore di quella di scrittura: per evitare lo svuotamento del buffer che si avrebbe in conseguenza di tale differenza di frequenza, è necessario inibire, in corrispondenza di prefissati punti della trama d'uscita (opportunità di giustificazione positiva), la lettura della memoria e inserire nel segnale multiplo un bit che non trasporta informazione di tributario (bit di riempimento). La giustificazione è preannunciata mediante opportuna segnalazione (bit di segnalazione di giustificazione) presente nella trama del segnale multiplo. Da quanto detto risulta che nella trama del segnale multiplo, accanto ai bit dei tributari, sono presenti bit di informazione ausiliaria che permettono la segnalazione di giustificazione: come è noto altri bit ausiliari sono destinati ad altri scopi come, ad esempio, l'indicazione dell'inizio di trama (parola di allineamento) e la gestione degli allarmi. L'insieme dei bit ausiliari è comunemente denominato *overhead* della trama.

Al lato ricezione, il riconoscimento della parola d'allineamento consente d'individuare la posizione dei bit di overhead, in particolare quelli che segnalano la giustificazione dei tributari affasciati, e la posizione dei bit di tributario: il compito di ricostruire ciascun segnale

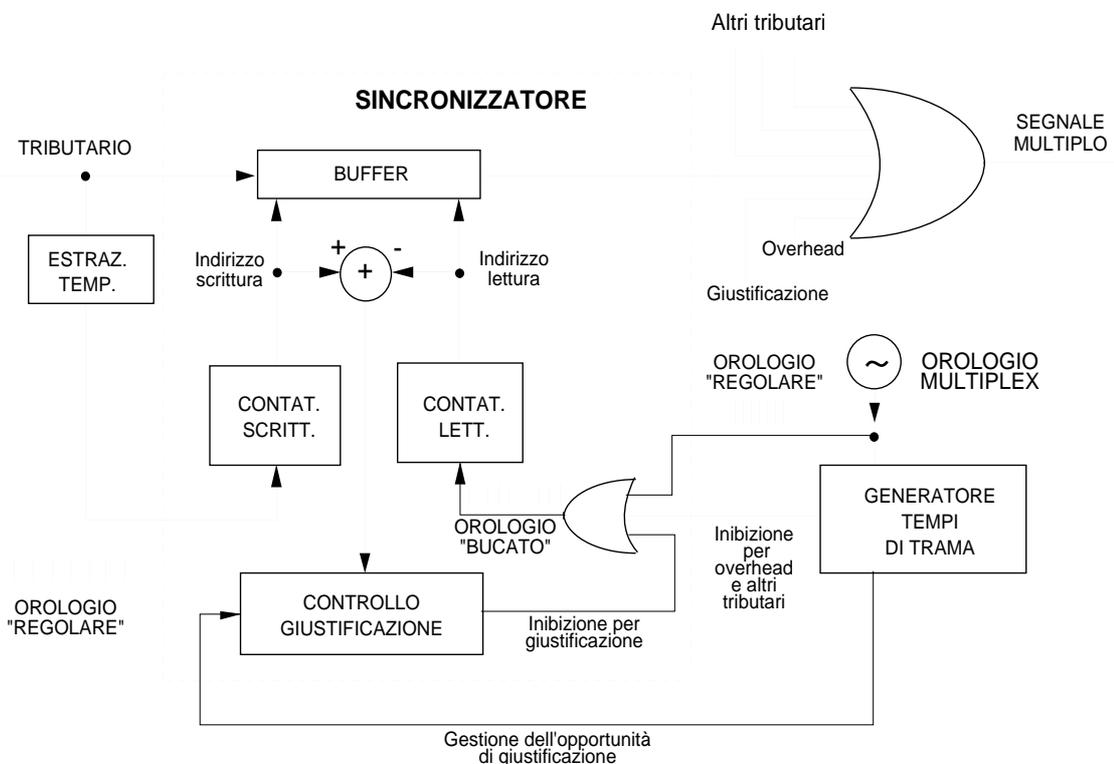


Figura 7 Schema funzionale dell'inserimento di un tributario plesiocrono nel segnale multiplo

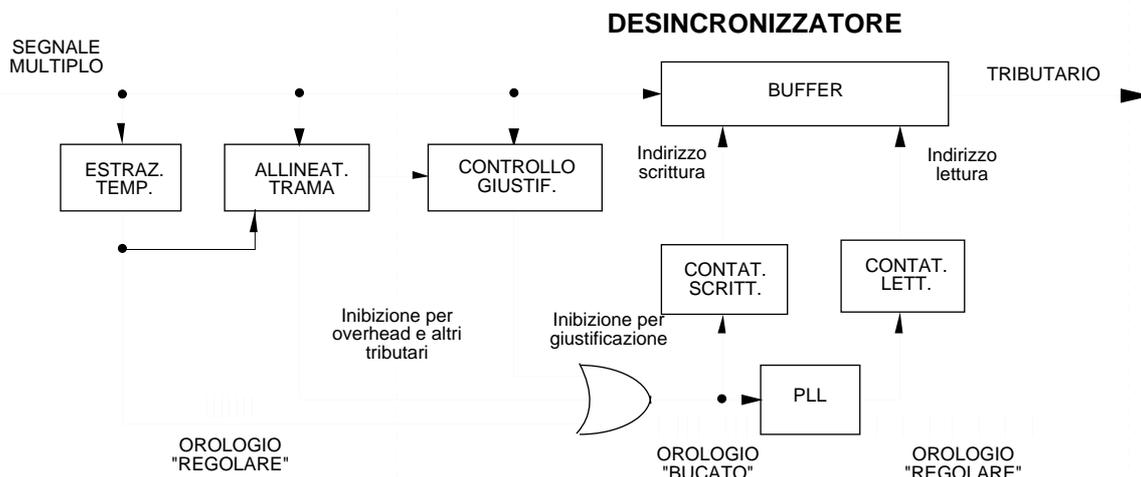


Figura 8 Schema funzionale dell'estrazione di un tributario plesiocrono dal segnale multiplo

tributario è svolto nel relativo *desincronizzatore* il cui schema di principio è mostrato in Figura 8.

I bit di tributario individuati nella trama multipla vengono scritti in un buffer sotto il controllo di un orologio "bucato", ottenuto inibendo il cronosegnale estratto dal segnale multiplo negli istanti caratteristici pertinenti sia all'overhead, sia ai bit di giustificazione, sia agli altri tributari: tale orologio viene anche utilizzato in un circuito ad aggancio di fase (PLL = Phase Locked Loop) per ricavare l'orologio di lettura del buffer (orologio regolare) e deve perciò riprodurre con la massima fedeltà possibile il cronosegnale originario del tributario. Puntando l'attenzione sugli aspetti di sincronizzazione delle operazioni sopra descritte è importante sottolineare come il processo di moltiplicazione-demoltiplicazione plesiocrona consente di trasportare, insieme al contenuto informativo dei flussi tributari, anche l'associata temporizzazione. In pratica, mentre per il contenuto informativo la qualità del trasporto è influenzata esclusivamente dalle degradazioni introdotte dal mezzo trasmissivo che connette moltiplicatore e demoltiplicatore, il trasferimento dell'informazione di temporizzazione associata a ciascun tributario è affetto anche dalle fluttuazioni di fase legate al meccanismo di giustificazione, che il PLL non riesce a filtrare completamente (jitter da tempo d'attesa). Al fine di minimizzare l'ampiezza di tali fluttuazioni di fase residue, occorre individuare opportuni valori per la frequenza media nominale di lettura del buffer del sincronizzatore e per le tolleranze di frequenza dei tributari e del segnale multiplo. Per la gerarchia plesiocrona europea, basata come si è detto sul segnale primario a 2,048 Mbit/s, i valori delle tolleranze di frequenza ammesse sono riportati in Tabella 1.

L'introduzione dei vari livelli della gerarchia plesiocrona ha reso possibile il graduale sviluppo di una rete di giunzione numerica basata sull'impiego di moltiplicatori e sistemi di linea alle diverse frequenze di cifra normalizzate. Tale rete ha assunto configurazioni progressivamente più complesse via via che la disponibilità di nuovi apparati ha consentito un maggior numero di tipologie di collegamento: solo a titolo d'esempio, in Figura 9 sono schematizzati alcuni possibili collegamenti numerici basati sulla PDH.

Al fine di contenere le degradazioni del segnale numerico trasportato in una rete PDH è stato necessario specificare un certo numero di parametri di interfaccia atti a caratterizzare il segnale stesso, e progettare gli apparati in modo tale da soddisfare i requisiti di qualità all'interfaccia di ingresso e di uscita, prevedendo se necessario l'adozione di dispositivi capaci di ridurre le

Frequenza nominale (MHz)	Tolleranza (ppm)
2,048	±50
8,448	±30
34,368	±20
139,264	±15

ppm = parti per milione

Tabella 1 Valori delle tolleranze di frequenza ammesse per la gerarchia plesiocrona europea (ppm = parti per milione)

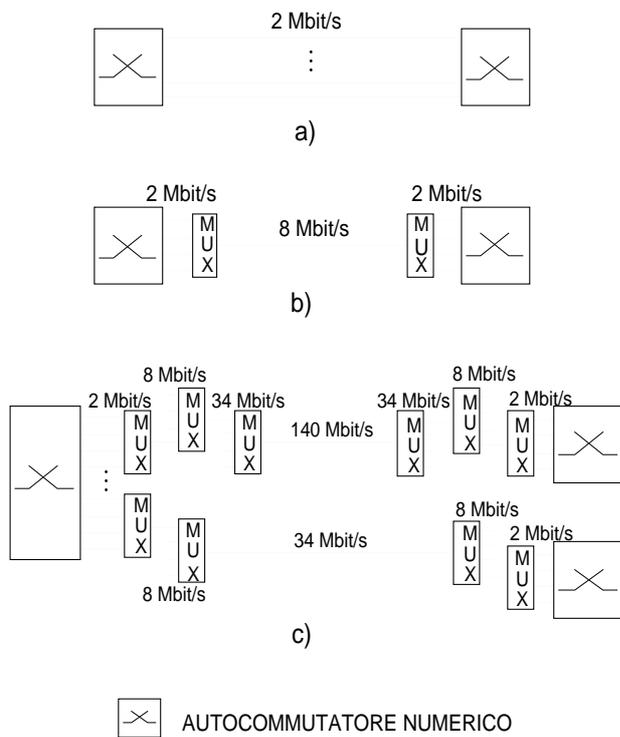


Figura 9 Esempi di connessione tra autocommutatori numerici: a) collegamenti a 2 Mbit/s; b) collegamento a 8 Mbit/s; c) collegamenti a 34 e 140 Mbit/s

degradazioni in eccesso. Riferendoci in particolare agli aspetti di temporizzazione degli apparati trasmissivi, ricordiamo che le due più importanti degradazioni cui vanno soggetti i cronosegnali sono il jitter e il wander: entrambi sono costituiti da variazioni della posizione degli istanti significativi, e si differenziano tra loro per la velocità di tali variazioni, che sono rapide nel jitter e lente nel wander, e per le cause da cui hanno origine. Il jitter è essenzialmente prodotto dai circuiti di estrazione della temporizzazione e da quelli di rigenerazione del segnale di linea, dall'impiego di tecniche di giustificazione nei moltiplicatori, dal rumore di fase presente negli oscillatori degli orologi d'apparato; il wander è causato dalle fluttuazioni di temperatura dei portanti fisici (che si traducono in variazioni del tempo di propagazione), dalla deriva della lunghezza d'onda dei laser nei sistemi ottici (variazioni del tempo di propagazione dovute alla dispersione cromatica delle fibre), dalle tecniche di giustificazione. Jitter e wander sono a loro volta causa di importanti degradazioni della qualità del trasporto quali: incremento del tasso d'errore, slip nelle memorie elastiche presenti all'ingresso di molti apparati numerici, distorsione dei segnali analogici ricostruiti in ricezione dopo la conversione numerico/analogico (ad esempio, per i servizi hi-fi e video).

3.3 Strategie di controllo del jitter e del wander nelle reti numeriche

Per meglio comprendere gli aspetti quantitativi del controllo del jitter che verranno considerati in questo paragrafo, è utile richiamare brevemente il modello di riferimento che consente di definire il jitter come segnale. Le deviazioni temporali degli istanti caratteristici di un segnale numerico, rispetto alle loro posizioni nominali, costituiscono una sequenza tempo-discreta di valori analogici, definiti cioè in un insieme continuo: tale sequenza può essere interpretata come la versione campionata di un sottostante segnale analogico tempo-continuo che rappresenta la modulazione di fase del cronosegno sinusoidale associato al segnale numerico. Quando si parla di ampiezza, frequenza, forma d'onda, etc., del jitter ci si riferisce alle corrispondenti grandezze associate alla versione tempo-discreta o a quella tempo continua, in dipendenza dalle tecniche di rivelazione impiegate: queste, nel primo caso sono basate sulla misura delle variazioni degli intervalli di tempo tra istanti significativi consecutivi, mentre nel secondo caso sono basate sulla demodulazione di fase applicata al cronosegno estratto dal segnale numerico.

Per attuare una strategia di controllo del jitter e del wander sono stati introdotti nelle raccomandazioni del CCITT, e in particolare nella G.823 [11], i seguenti criteri:

- 1) identificazione dei limiti di rete che non debbono essere superati in corrispondenza a qualsiasi interfaccia gerarchica;
- 2) regolamentazione delle prestazioni dei singoli apparati numerici;
- 3) individuazioni di linee-guida per la valutazione delle prestazioni in presenza di apparati in cascata (problemi di accumulazione) e per l'identificazione di metodologie affidabili per la misura dei parametri.

Gli obiettivi di un simile approccio sono quelli di rendere possibile una facile integrazione in rete di nuovi apparati e una gestione flessibile della rete numerica.

3.3.1 Limiti di rete per il jitter/wander

Con riferimento al primo punto ricordiamo che per misurare e caratterizzare il jitter e il wander presenti alle interfacce gerarchiche si sono scelte due grandezze diverse: il jitter picco-picco misurato in una opportuna banda di frequenza, nel caso del limite di rete per il jitter; il Maximum Time Interval Error (MTIE) nel caso del limite di rete per il wander: per la definizione di tale grandezza si rimanda il lettore alla terza parte dell'articolo. La scelta di due grandezze distinte è strettamente legata alla diversità delle tecniche di misura disponibili per caratterizzare i due fenomeni.

Per quanto riguarda il limite di rete per il jitter all'interfaccia gerarchica a 2,048 Mbit/s è prescritto che il jitter picco-picco non superi 1,5 UI (Unit Interval = periodo di simbolo) quando è misurato nella banda 20 Hz ÷ 100 kHz e 0,2 UI quando è misurato nella banda 18 kHz ÷ 100 kHz: analoghe specifiche sono date nella G.823 per gli altri livelli gerarchici.

Il limite di rete per il wander è stato finora fissato solo per le interfacce ai nodi di rete: l'MTIE su un periodo d'osservazione di $S > 10000$ s non deve eccedere il valore $(10^{-2}S + 10000)$ ns. Tale specifica tiene conto delle caratteristiche fondamentali di propagazione dei mezzi trasmissivi, dei fenomeni di instabilità degli oscillatori presenti negli orologi e delle caratteristiche di stabilità a lungo termine del master della rete.

3.3.2 Specifiche di jitter/wander per gli apparati

Per quanto riguarda le specifiche di jitter e wander per gli apparati, l'approccio seguito è basato sull'identificazione di tre aspetti fondamentali: l'accettazione di jitter/wander in ingresso, la generazione di jitter/wander in uscita in assenza di jitter/wander in ingresso, la funzione di trasferimento del jitter/wander.

Esaminiamo brevemente ciascuno di tali aspetti. Al fine di assicurare che un apparato possa essere connesso ad ogni interfaccia gerarchica all'interno di una rete numerica è necessario prevedere che le porte d'ingresso dell'apparato stesso possano accettare livelli di jitter/wander pari almeno al massimo valore presente in rete, cioè al limite di rete per il jitter/wander definito al punto 1). Tuttavia, per consentire l'impiego di agevoli metodologie di prova degli apparati, le

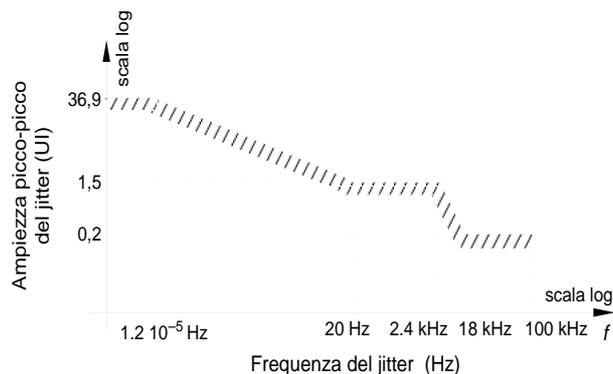


Figura 10 Maschera di accettazione del jitter sinusoidale alle porte di ingresso a 2,048 Mbit/s degli apparati

specifiche di accettazione del jitter sono definite in termini di ampiezza picco-picco a ciascuna frequenza del jitter sinusoidale che, se presente in ingresso all'apparato, non deve provocare alcuna significativa degradazione di funzionamento. A titolo di esempio in Figura 10 è riportata la maschera di accettazione del jitter/wander sinusoidale per le porte d'ingresso a 2,048 Mbit/s degli apparati. La congruenza di tale maschera con le specifiche del limite di rete per il jitter è assicurata dall'ipotesi ragionevole che il jitter picco-picco presente in rete a ciascuna frequenza sia comunque inferiore al jitter picco-picco in una banda che comprende la frequenza considerata.

Le specifiche sul jitter generato dagli apparati in assenza di jitter in ingresso sono date in termini di ampiezza picco-picco misurata in una assegnata banda di frequenza: per esempio per il multiplex PCM primario a 2,048 Mbit/s il valore massimo consentito è di 0,05 UI picco-picco quando misurato nella banda 20 Hz ÷ 100 kHz (Racc. G.732) [8].

Per quanto riguarda infine la funzione di trasferimento del jitter, definita come rapporto tra le ampiezze del jitter in uscita e di quello in ingresso all'apparato, ci limitiamo a riportare che le specifiche sono date per mezzo di una maschera guadagno/frequenza del tipo mostrato in Figura 11 con riferimento al moltiplicatore numerico 2/8 Mbit/s.

3.3.3 Il problema della valutazione del jitter/wander accumulato nelle reti numeriche

La variabilità delle configurazioni di rete impedisce la considerazione di tutti i possibili casi. Per analizzare una particolare configurazione di rete è necessario

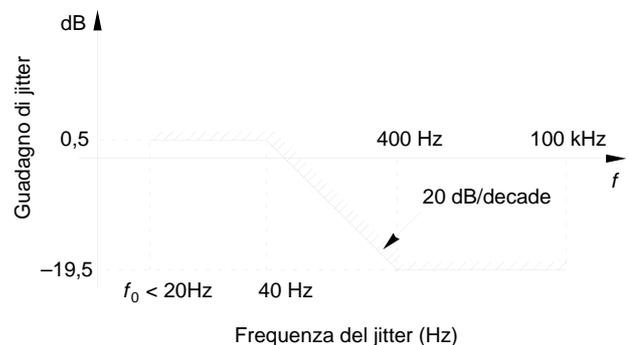


Figura 11 Caratteristica di trasferimento del jitter del moltiplicatore del secondo livello gerarchico a 8,448 Mbit/s

usare le informazioni disponibili sulle caratteristiche di jitter/wander degli apparati che ne fanno parte unitamente ad appropriati modelli di accumulazione. Quest'ultimo aspetto è stato oggetto di numerose indagini teoriche che hanno via via preso in considerazione vari aspetti del problema: i differenti tipi di apparati (rigeneratori, multiplatori, etc.), i differenti supporti fisici (cavi coassiali, ponti radio, fibre ottiche), le diverse caratteristiche di jitter (sistematico, non sistematico). I più significativi risultati al riguardo possono essere reperiti nell'Annesso B della racc. G.823.

4. Note conclusive

A conclusione di questa prima parte dedicata ad una panoramica dei problemi di sincronizzazione nelle reti numeriche basate sulla PDH è utile riportare l'attenzione sugli aspetti storico-evolutivi legati all'introduzione e al grande sviluppo della rete numerica di trasmissione. Da un impiego "sparso" dei primi multiplatori e sistemi di linea numerici si è passati a una progressiva diffusione di nuovi apparati, accompagnata dal progetto e dalla realizzazione di grandi strutture portanti nella rete a lunga distanza (dorsali in cavo coassiale e in fibra ottica). La rete numerica di trasmissione ha costituito perciò il nucleo iniziale del processo di numerizzazione e ha progressivamente incluso gli autocommutatori numerici, via via che questi sono stati introdotti a fianco o in sostituzione delle tradizionali centrali elettromeccaniche, avviando quella integrazione di trasmissione e commutazione che avrebbe dovuto portare alla realizzazione della cosiddetta rete numerica integrata nelle tecniche (IDN=Integrated Digital Network). Tale processo ha oggi raggiunto un elevato grado di penetrazione nella rete pubblica nazionale, con una evoluzione analoga a quella seguita nella gran parte degli altri Paesi della comunità internazionale a più elevato sviluppo tecnologico. Come è stato descritto nelle linee essenziali in questo articolo, le scelte di base che hanno caratterizzato lo sviluppo delle attuali reti di trasporto sono chiaramente individuabili nell'impiego degli apparati trasmissivi plesiocroni da un lato, e, dall'altro, nella progressiva introduzione di apparati di commutazione tra di loro sincronizzati al fine di assicurare elevati livelli di qualità dei servizi offerti. Tuttavia, come vedremo nella seconda parte dell'articolo, una nuova rivoluzione si è andata innestando su tale processo, non ancora giunto al suo completamento ma già caratterizzato da importanti segni di inadeguatezza rispetto alle crescenti esigenze dell'utenza e dei gestori di rete: la risposta a tali esigenze è stata individuata nelle potenzialità offerte dalla nuova gerarchia di multiplazione sincrona.

Bibliografia

- [1] Stover, H.A.: *Timing and synchronization of switched digital communications networks*. International Conference on Communications, 1976.
- [2] Decina, M.; Pietromarchi, A.; Bovo, A.; Musumeci, L.: *Development of network synchronization techniques in Italy*. International Conference on Communications, 1976.
- [3] Decina, M.; de Julio, U.: *Performance of integrated digital networks: international standards*. International Conference on Communications, 1982.
- [4] Decina, M.; Roveri, A.: *Introduzione alle reti telefoniche analogiche e numeriche*. La Goliardica, 1985.
- [5] Trischitta, P.R.; Varma, E.L.: *Jitter in digital transmission systems*. Artech House, New York, 1989.
- [6] Angelini, G. et. al.: *Manuale di reti di telecomunicazioni e trasmissione dati*. Calderini, Bologna, 1990.
- [7] *Physical/electrical characteristics of hierarchical digital interfaces*. Rac. CCITT G.703, Blue Book, Melbourne, 1988.
- [8] *Characteristics of primary PCM multiplex equipment operating at 2048 kbit/s*. Rac. CCITT G.732, Blue book, vol.III, fascicolo III.4, Melbourne, 1988.
- [9] *Controlled slip rate objectives on an international digital connection*. Rac. CCITT G.821, Ginevra, giugno 1992.
- [10] *Controlled slip rate objectives on an international digital connection*. Rac. CCITT G.822, Blue book, vol.III, fascicolo III.5, Melbourne, 1988.
- [11] *The control of jitter and wander within digital networks which are based on the 2048 kbit/s hierarchy*. Rac. CCITT G.823, Blue book, vol.III, fascicolo III.5, Melbourne, 1988.

La sincronizzazione nelle reti numeriche di trasporto

Parte seconda: le future reti di trasporto basate sull'SDH⁽¹⁾

M. Carbonelli, D. De Seta, D. Perucchini (*)

In questa seconda parte dell'articolo dedicato alla sincronizzazione delle reti numeriche di trasporto, vengono considerate le problematiche connesse con lo sviluppo delle future reti basate sulla gerarchia numerica sincrona SDH. In particolare, si illustrano le principali caratteristiche della nuova gerarchia e degli apparati basati su tale standard; successivamente, viene descritta la struttura tipica di una rete sincrona di trasporto con particolare attenzione ai problemi legati alla sua sincronizzazione e alla prevista interconnessione con la rete PDH.

1. Introduzione

La rete di trasporto degli anni '80 è stata caratterizzata, come analizzato nella prima parte di questo articolo, dall'impiego dei sistemi di trasmissione basati sulla gerarchia numerica plesiocrona PDH. Questa rete, mentre da un lato assicurava un'elevata capacità trasmissiva attraverso l'utilizzazione dei livelli gerarchici superiori, dall'altro non permetteva di operare con un sufficiente grado di rapidità per quello che riguarda le riconfigurazioni che, sia per ragioni di traffico sia per il verificarsi di guasti, sono richieste con frequenza non trascurabile nei nodi della rete stessa. In particolare, in questo tipo di rete le riconfigurazioni vengono attuate operando su delle strutture meccaniche, dette *ripartitori*, che consentono di attestare fisicamente le varie interfacce di moltiplicazione e di linea presenti in un nodo in modo tale da interconnettere gli apparati in maniera flessibile. In pratica si può affermare che le connessioni realizzate nel ripartitore costituiscono dei collegamenti semipermanenti. Le operazioni necessarie per ottenere una riconfigurazione vengono realizzate permutando opportunamente i cavi che si attestano al ripartitore. Un esempio del principio di utilizzazione di un ripartitore meccanico è schematizzato in Figura 1: in questo caso è possibile, operando manualmente sui cavi che si attestano al ripartitore, indirizzare i 2 Mbit/s in ingresso verso il desiderato moltiplicatore posto a valle della struttura meccanica.

La rigidità intrinsecamente connessa con la natura "manuale" delle operazioni di permutazione può essere superata attraverso la sostituzione dei ripartitori meccanici con dei *ripartitori numerici automatici*. Tali dispositivi, denominati RED (Ripartitori Elettronici Digitali), integrano le funzioni di moltiplicazione e demoltiplicazione con la funzione di permutazione permettendo di attuare in tempi estremamente brevi le operazioni di riconfigurazione di rete. L'elemento fondamentale che caratterizza il funzionamento dei RED è la matrice di connessione elettronica che, utilizzando le informazioni trasmesse da un centro di controllo, attua le permutazioni sui flussi numerici che afferiscono al ripartitore. In Figura 2 è mostrato lo schema di principio delle operazioni svolte in un RED che ha in ingresso e in uscita flussi del terzo livello gerarchico PDH (34 Mbit/s) e che attua la permutazione sui flussi del primo livello (2 Mbit/s): in tale configurazione il ripartitore è detto RED 3/1. Osservando la figura si nota come all'ingresso e all'uscita dell'apparato è necessario utilizzare rispettivamente dei moltiplicatori e dei demoltiplicatori PDH.

Questa necessità è conseguenza diretta della struttura dei segnali multipli plesiocroni che, impiegando la tecnica di giustificazione e l'interallacciamento bit a bit, non rendono possibile un accesso immediato ai segnali tributari: per esemplificare, questo significa che

(*) ing. Marco Carbonelli, ing. Domenico De Seta, ing. Daniele Perucchini - Fondazione Ugo Bordoni - Roma

(1) Lavoro eseguito nell'ambito della Convenzione in atto tra l'Amministrazione delle Poste e Telecomunicazioni e la Fondazione Ugo Bordoni.

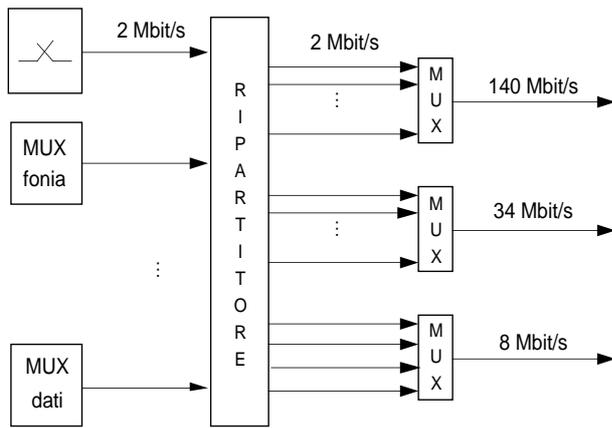


Figura 1 Schema di impiego di un ripartitore per la permutazione di flussi a 2,048 Mbit/s

se si è interessati alla permutazione di flussi a 2 Mbit/s da un segnale multiplo a 34 Mbit/s è indispensabile attuare prima la demultiplazione $34 \rightarrow 8$ Mbit/s e successivamente quella $8 \rightarrow 2$ Mbit/s (come mostrato in Figura 2). Il ripartitore presenta quindi caratteristiche di notevole complessità ed elevato costo. Queste limitazioni, unitamente all'esigenza dei gestori delle telecomunicazioni di poter disporre di reti sempre più flessibili, potenti ed economiche, hanno favorito l'avvio di una vera e propria rivoluzione nel campo dei sistemi e delle reti numeriche: tale rivoluzione è stata resa possibile dai progressi conseguiti nel campo della microelettronica, nell'ottica e nella produzione di software, e si è concretizzata nella standardizzazione di una nuova gerarchia di multiplazione, la gerarchia numerica sincrona SDH.

L'avvento della gerarchia sincrona ha di fatto congelato le ricerche e i progetti che erano in atto per

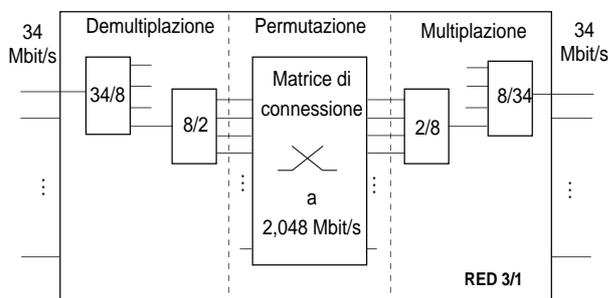


Figura 2 Schema di principio delle operazioni svolte in un ripartitore automatico (RED3/1) che ha in ingresso e uscita flussi a 34 Mbit/s e permuta i flussi a 2 Mbit/s

i sistemi plesiocroni, in particolare per quel che riguarda i sistemi di linea in fibra ottica a velocità superiore a 565 Mbit/s e i RED con interfacce d'accesso PDH, dando un fortissimo impulso allo studio e allo sviluppo di apparati di nuova concezione, quali i terminali sincroni, i multiplatori add-drop e i cross-connect, che consentiranno di utilizzare, sia nell'area di distribuzione che in quella di giunzione, topologie di rete ad anello caratterizzate da un'alta affidabilità ed efficienza.

In questa seconda parte dell'articolo dedicato alla sincronizzazione delle reti numeriche, sono illustrate le principali caratteristiche della nuova gerarchia di multiplazione sincrona SDH (par.2) e degli apparati (par.3) basati su tale standard. La descrizione sarà necessariamente per grandi linee e indirizzata alla individuazione dei problemi connessi con la sincronizzazione: un'analisi più approfondita è contenuta in articoli già pubblicati su questo Notiziario [1,2]. Successivamente, viene descritta (par.4) la struttura tipica di una rete di trasporto sincrona, con particolare attenzione ai problemi legati alla sua sincronizzazione e alla prevista interconnessione con la rete PDH.

2. Caratteristiche fondamentali dell'SDH

La standardizzazione della gerarchia numerica sincrona è il frutto di una intensa attività di ricerca e di studio che ha caratterizzato, a partire dalla metà degli anni '80, i lavori degli enti di normativa internazionali (CCITT, ETSI, ANSI). Mentre per quel che riguarda la struttura della trama dei vari livelli gerarchici quasi tutto è stato oramai concordato, resta ancora del lavoro da svolgere, soprattutto con riguardo al controllo della qualità di trasmissione (in termini di prestazioni d'errore, di jitter e di wander), della sincronizzazione e della gestione di rete.

L'SDH è organizzata su tre livelli gerarchici detti *moduli di trasporto sincrono a livello N* (STM-N = Synchronous Transport Module at level N), in cui il parametro N assume i valori 1, 4 e 16 per il primo, secondo e terzo livello rispettivamente. La frequenza di cifra F_N associata ad ogni livello gerarchico è calcolabile a partire dalla frequenza del livello primario $F_1 = 155,52$ Mbit/s, utilizzando la relazione $F_N = F_1 \cdot N$. Tutti i livelli gerarchici sono organizzati in trame della durata di 125 μ s strutturate ad ottetti: queste caratteristiche consentono all'SDH di trasportare in modo trasparente gli ottetti fonici PCM, rendendo possibile il futuro impiego di interfacce a 155 Mbit/s anziché a 2 Mbit/s negli autocommutatori numerici. Il segnale STM-N si ottiene per interallacciamento di ottetti a partire da N tributari STM-1 opportunamente allineati in fase. Per la rappresentazione della trama del segnale STM-N ci si avvale, come mostrato in

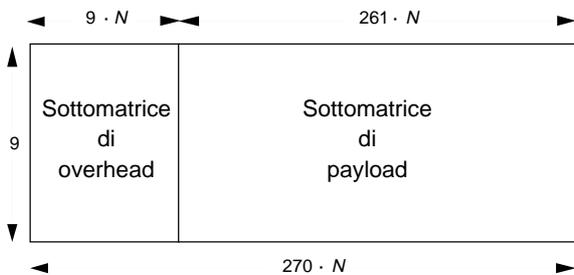


Figura 3 Rappresentazione in forma di matrice della trama del segnale STM-N

Figura 3, di una matrice a 9 righe e 270·N colonne, in cui ogni elemento identifica l'ubicazione di un ottetto: l'ordine di successione cronologica degli ottetti è da sinistra verso destra e dall'alto verso il basso.

La matrice di Figura 3 è logicamente suddivisa in due sottomatrici:

- la sottomatrice di overhead (9x(9·N)), che contiene informazioni ausiliarie richieste per la realizzazione delle funzioni di allineamento di trama, telesorveglianza e gestione;
- la sottomatrice di payload (9x(261·N)), che contiene le informazioni relative ai tributari (carico utile).

Per realizzare il trasporto sincrono, le informazioni di carico utile vengono inserite, all'atto della moltiplicazione, in strutture numeriche sincrone che rappresentano i "mattoni" necessari per costruire i segnali STM-N. Le strutture numeriche vengono a loro volta inserite nella sottomatrice di payload e sono organizzate in modo da poter trasportare sia i flussi delle esistenti gerarchie plesiocrone (europea e americana), sia i flussi generati a partire da celle ATM (Asynchronous Transfer Mode) e DQDB (Distributed Queue Dual Bus). Le strutture numeriche costituiscono una delle innovazioni più importanti introdotte con la gerarchia sincrona.

E' importante notare che, a dispetto del termine 'sincrona' presente nella denominazione della gerarchia, la moltiplicazione adottata nell'SDH è di tipo mesocrono poiché, nel breve termine o in condizioni di cattivo funzionamento della rete di sincronizzazione, i cronosegnali delle strutture numeriche possono presentare sia ampie fluttuazioni di fase sia vere e proprie caratteristiche di plesiocronia⁽²⁾. Per fronteggiare queste ultime situazioni, anche nella gerarchia sincrona è stato necessario prevedere il meccanismo della giustificazione.

Tra le strutture numeriche previste nell'architettura di moltiplicazione sincrona per la costruzione di un segnale

Contenitori Virtuali	Ritmo binario dei VC (Mbit/s)	Ritmo binario dei segnali PDH trasportati (Mbit/s)
VC-11	1,664	1,544
VC-12	2,24	1,544 2,048
VC-2	6,848	6,312
VC-3	48,96	34,368 44,736
VC-4	150,336	139,264

Tabella 1 Contenitori virtuali per i tributari della gerarchia PDH

STM-N, i contenitori virtuali VC (Virtual Container) costituiscono, nello stesso tempo, la struttura base per ottenere tutte le altre strutture e l'elemento concettualmente innovativo rispetto alla tradizionale moltiplicazione plesiocrona. Infatti i VC sono destinati a trasportare, all'interno della rete SDH, i flussi tributari che si presentano all'ingresso della rete: per ogni tipo di tributario della gerarchia PDH è stato predisposto un VC opportuno, come indicato in Tabella 1. L'introduzione dei contenitori virtuali consente di trasportare individualmente ciascun tributario, in modo indipendente dal trasporto di ogni altro tributario.

I VC, che hanno una struttura di lunghezza predefinita, vengono inseriti nella trama STM-N e sono *singolarmente e indipendentemente* individuabili attraverso una informazione che è a loro associata all'atto della moltiplicazione: questa informazione, detta *puntatore*, è allocata in una posizione fissa all'interno della trama STM-N. Il puntatore è costituito dall'insieme di dieci cifre binarie il cui equivalente decimale permette di determinare, nella sottomatrice di payload, l'allocazione del primo byte del VC e quindi, di accedere all'intero VC senza intervenire sugli altri VC presenti nella trama. Questa possibilità rende estremamente più semplici le operazioni di inserimento, estrazione e permutazione dei flussi tributari nei segnali multipli. Infatti, mentre nei sistemi plesiocroni per individuare un tributario di un dato livello gerarchico è necessario demoltiplicare il segnale multiplo fino al raggiungimento del livello di interesse, eseguendo di fatto tutte le demoltiplicazioni anche su tutti gli altri flussi tributari presenti nel segnale di partenza, con l'introduzione del puntatore e della struttura VC si rende possibile l'elaborazione indipendente dei vari segnali tributari: in particolare si può affermare che, all'interno della rete SDH, i VC costituiscono dei *tributari sincroni* con cui si costruiscono i segnali multipli STM-N. Ogni VC, una volta che è stato costruito in un nodo di origine, non viene più né elaborato né alterato, per quello che riguarda il suo contenuto informativo, sino all'arrivo nel nodo di destinazione: l'unica azione che può essere su di esso attuata nei nodi in cui semplicemente transita è quella di uno spostamento all'interno della trama del segnale STM-N in uscita dal nodo. Come vedremo nel par.4, questa possibilità

(2) Nonostante questa precisazione al fine di essere coerenti con le denominazioni utilizzate a livello internazionale identificheremo con il termine "moltiplicazione sincrona" la "moltiplicazione mesocrona".

che un VC ha di fluttuare nel segnale multiplo sincrono consente anche di fronteggiare le inevitabili variazioni di fase tra il cronosegnale ad esso associato e il cronosegnale che governa il funzionamento dei nodi di rete attraversati dal VC.

3. Gli apparati dell'SDH

Per meglio comprendere le considerazioni che saranno svolte nel par.4 è necessario descrivere preliminarmente le più importanti tipologie e i principi di funzionamento degli apparati dell'SDH. In generale si possono identificare tre diversi tipi d'apparato: il moltiplicatore terminale, il moltiplicatore add-drop e il cross-connect.

I *moltiplicatori terminali* TM (Terminal Multiplexer) costituiscono gli apparati più semplici: con essi è possibile costruire i flussi STM-N, a partire dai segnali tributari che devono essere trasportati nella rete SDH. Il moltiplicatore terminale inserisce ciascun flusso tributario d'ingresso in un opportuno VC operando un primo "adattamento" tra il cronosegnale del tributario in ingresso e quello del VC, derivato dall'orologio del moltiplicatore. Dopo aver costruito i VC, il moltiplicatore è in grado di generare il segnale STM-N che, al suo interno e in posizioni fisse, trasporterà le informazioni di puntatore associate (nella fase di moltiplicazione) ad ogni contenitore virtuale.

Nella fase di demoltiplicazione, dopo aver attuato l'allineamento di trama sul segnale sincrono STM-N, è possibile leggere le informazioni di puntatore e quindi risalire alle posizioni dei VC all'interno della trama: a partire dai VC sono individuabili i bit del tributario che, in fase di moltiplicazione, erano stati inseriti nel segnale sincrono; e quindi è possibile ricostruire nuovamente il flusso informativo tributario dopo aver operato un nuovo "adattamento" tra la temporizzazione del VC e quella del segnale tributario trasportato al suo interno. Le modalità con cui è possibile realizzare i suddetti adattamenti fra cronosegnali saranno illustrate in dettaglio nel par.4.

L'*Add-Drop Multiplexer* (ADM) è un apparato che realizza le funzioni di estrazione e di inserimento di flussi

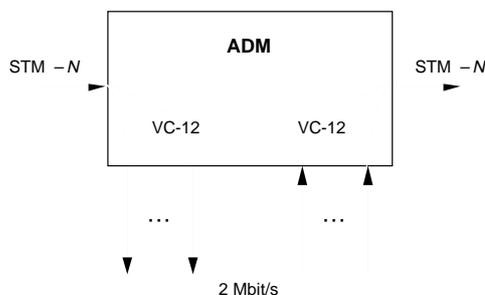


Figura 4 Rappresentazione schematica dell'inserzione e dell'estrazione di flussi a 2 Mbit/s in un ADM

tributari in un segnale multiplo SDH. In Figura 4 è schematicamente rappresentata l'inserzione e l'estrazione di flussi a 2 Mbit/s in un ADM avente interfacce di ingresso e uscita STM-N.

Il funzionamento dell'ADM beneficia enormemente delle potenzialità offerte dalla moltiplicazione sincrona che sono legate, come si è detto, all'impiego del meccanismo di puntatore. Infatti, i contenitori virtuali presenti nel segnale STM-N in ingresso vengono identificati nella trama attraverso il puntatore a loro associato: i VC che devono essere estratti vengono elaborati come avverrebbe in un TM, mentre quelli che devono transitare nel nodo vengono direttamente inseriti nel segnale STM-N d'uscita. La capacità resa disponibile dall'estrazione di tributari consente ora l'inserimento di nuovi tributari, che inizieranno così il loro trasporto nella rete SDH. Le operazioni di inserimento e estrazione vengono attuate tutte a livello di contenitori virtuali che, come già accennato in precedenza, costituiscono la struttura base su cui si opera all'interno della rete sincrona.

I *Digital Cross-Connect* (DXC) costituiscono i ripartitori della rete SDH: talora, per identificare questi apparati, si utilizza anche il termine RED. I DXC permettono di operare la permutazione dei VC presenti nei segnali STM-N d'ingresso, come illustrato in Figura 5. In virtù delle funzioni tipiche del DXC, i VC che viaggiano su uno stesso STM-N in ingresso al DXC possono essere separati e inseriti, per motivi legati alla gestione e configurazione della rete, su STM-N d'uscita diversi. Le operazioni di permutazione vengono attuate sui VC attraverso l'utilizzazione di una matrice di connessione che è direttamente controllata dall'amministratore della rete.

Sia i DXC che gli ADM introducono una notevole flessibilità nelle operazioni di configurazione di rete e consentono di progettare reti di trasporto basate su topologie ad anello, che risultano più affidabili delle topologie a stella attualmente adottate: la struttura ad anello, infatti, consente di ovviare a guasti che colpiscono un collegamento o un apparato dell'anello, reinstrandando il traffico su collegamenti alternativi predisposti nell'ambito di opportune strategie di protezione.

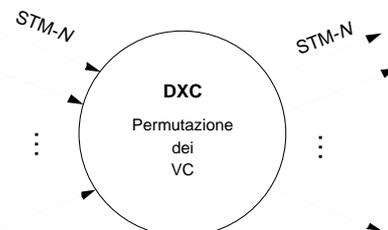


Figura 5 Funzione svolta da un DXC, con interfacce d'ingresso e d'uscita STM-N, in cui si opera la permutazione automatica dei flussi VC trasportati nelle trame sincrone

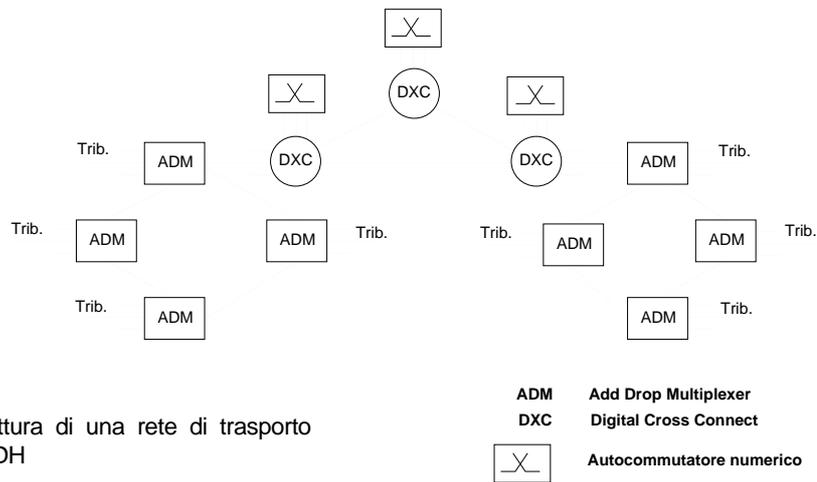


Figura 6 Possibile struttura di una rete di trasporto basata sull'SDH

4. Sincronizzazione nelle reti di trasporto basate sull'SDH

L'introduzione dei sistemi SDH nella rete di trasporto comporterà una vera e propria rivoluzione per quel che riguarda la filosofia di gestione della rete, la manutenzione e le prestazioni che verranno offerte. Una possibile struttura di una porzione di rete basata sull'SDH è mostrata in Figura 6, in cui è prevista la realizzazione di anelli di DXC nell'area di giunzione e di anelli di ADM nell'area di accesso.

Per analizzare approfonditamente le problematiche di sincronizzazione consideriamo la Figura 7 in cui è stata rappresentata una porzione della rete, mettendo in evidenza

la tipologia degli apparati e le sorgenti di sincronizzazione presenti. Tutti i nodi di rete richiedono l'impiego di un orologio che, in condizioni di normale funzionamento della rete di distribuzione della sincronizzazione, è agganciato all'orologio primario di rete (si veda la terza parte dell'articolo). In condizioni di malfunzionamento della rete di sincronizzazione, l'orologio di ciascun nodo deve essere in grado di fornire un cronosegnale di assegnata qualità per un tempo sufficiente a consentire la riconfigurazione della rete di sincronizzazione o la riparazione del guasto che ha causato il malfunzionamento. I flussi tributari in ingresso agli ADM provengono tipicamente dalla rete PDH e non sono perciò sincroni con gli orologi dei nodi SDH: l'adattamento del cronosegnale di tributario, come vedremo nel paragrafo successivo, è realizzato in un dispositivo, denominato sincronizzatore, presente negli apparati SDH e il cui funzionamento è schematizzato in Figura 8.

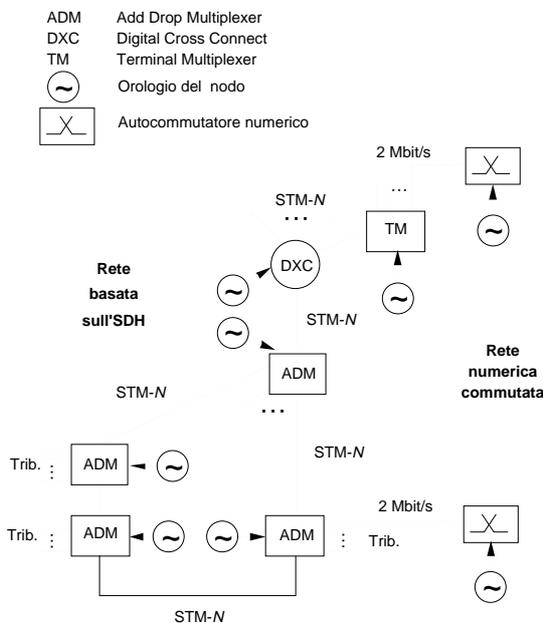


Figura 7 Porzione di una rete basata sull'SDH in cui sono state evidenziate le sorgenti di sincronizzazione presenti nei vari nodi

4.1 Inserimento di tributari plesiocroni nella trama SDH

Il dispositivo che realizza nei moltiplicatori terminali la funzione di inserimento dei tributari plesiocroni nella trama SDH è analogo a quello già analizzato per il caso della moltiplicazione plesiocrona e consente di gestire, attraverso la tecnica della giustificazione di bit, le differenze di frequenza tra i segnali tributari e i VC. Nella struttura VC sono presenti, oltre ai bit dedicati al trasporto del tributario e all'overhead di percorso POH (Path OverHead), altri bit ausiliari che consentono di segnalare e attuare, a seconda del tipo di contenitore considerato⁽³⁾, la giustificazione

(3) E' bene notare come la giustificazione prevista sia di tipo positivo-nullo-negativo nel caso del VC-11, del VC-12, del VC-2 e del VC-3, quando quest'ultimo trasporta un 34Mbit/s, e di tipo positivo nel caso del VC-4 e del VC-3 che trasporta un 45 Mbit/s della PDH americana.

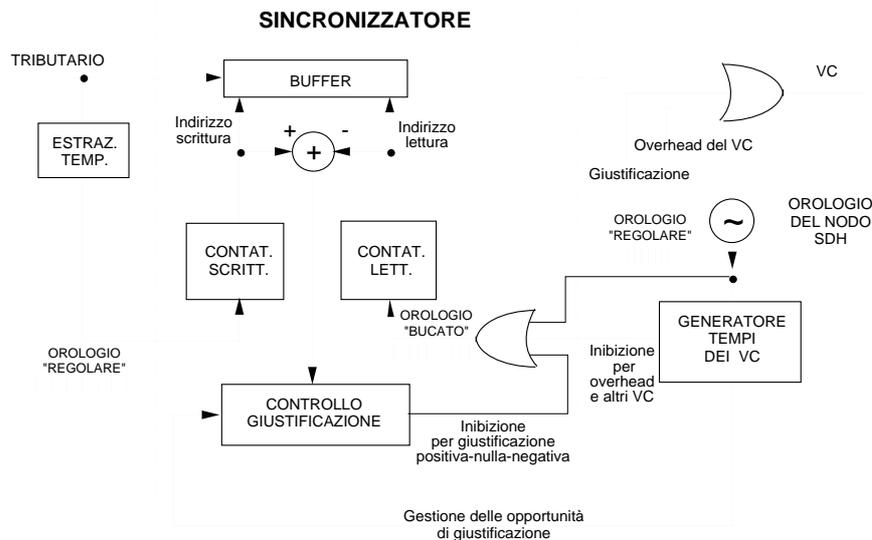


Figura 8 Schema funzionale dell'inserimento di un tributario plesiocrono in un contenitore virtuale dell'SDH

positiva, negativa o nulla. In pratica il sincronizzatore SDH si differenzia da quello PDH per il fatto che, mentre nel plesiocrono la giustificazione è solo positiva, nel sincrono può essere anche negativa: questo perché, in conseguenza della struttura e delle caratteristiche della trama sincrona, la frequenza media di lettura del buffer non è sempre maggiore della frequenza media di scrittura e, quindi, oltre a presentarsi condizioni di svuotamento, già analizzate nel par. 3.1 della prima parte dell'articolo, si verificano anche condizioni di riempimento del buffer stesso. Questi riempimenti, che potrebbero condurre alla comparsa di uno slip, sono fronteggiati utilizzando per il trasporto del tributario d'ingresso le *opportunità di giustificazione negativa*, presenti nella struttura dei VC, che normalmente trasportano bit non significativi: l'uso di questi bit aggiuntivi consente di aumentare la frequenza media di lettura del buffer, evitando così il fenomeno del trabocco. La decisione sull'opportunità di attuare o meno una giustificazione viene presa confrontando lo stato di riempimento del buffer (Figura 8) con due valori di soglia posti a distanza prefissata. Se lo stato di riempimento è tale da essere contenuto all'interno dei valori assunti dalle due soglie, si attua la giustificazione nulla (cioè nell'opportunità di giustificazione positiva viene trasportato un bit di carico utile e in quella negativa un bit di riempimento); se lo stato di riempimento si riduce a tal punto da oltrepassare la soglia inferiore, si attua la giustificazione positiva (sia l'opportunità di giustificazione positiva che quella negativa trasportano bit di riempimento); infine, se lo stato di riempimento è tale da oltrepassare la soglia superiore, si attua la giustificazione negativa (sia l'opportunità di giustificazione positiva che quella negativa trasportano bit di carico utile). La giustificazione operata per inserire un tributario in un VC viene comunemente denominata *giustificazione da inserimento o da mapping*: questa

giustificazione viene eseguita solo all'atto della costituzione di un VC, e viene rielaborata unicamente in fase di desincronizzazione, quando cioè si deve ricostruire il tributario trasportato nell'SDH. Come già osservato, una volta che si sono costruiti i contenitori virtuali, si è in grado di generare i flussi STM-N. Per approfondire ora quali operazioni di sincronizzazione si attuano sui VC, una volta che questi arrivano in un nodo intermedio di rete, è necessario fare alcune considerazioni sul funzionamento interno degli apparati SDH e sulla strategia di sincronizzazione di rete.

4.2 Risincronizzazione dei VC nei nodi intermedi della rete SDH

Osservando la Figura 7 si può notare che un nodo SDH riceve uno o più flussi STM-N, ognuno dei quali è caratterizzato dalla temporizzazione del nodo SDH che lo ha generato, eventualmente degradata da jitter e wander accumulati durante la trasmissione. L'apparecchiatura SDH *risincronizza* all'orologio che controlla il funzionamento del nodo (orologio locale) tutti i VC che afferiscono attraverso i flussi STM-N d'ingresso: la risincronizzazione deve essere operata sia nei nodi di transito, in cui si effettuano usualmente operazioni di permutazione, sia nei nodi terminali in cui si ricostruiscono i tributari. Per attuare la risincronizzazione si utilizzano delle memorie elastiche che vengono scritte utilizzando la temporizzazione estratta dai flussi entranti e vengono lette con la temporizzazione dell'orologio locale. A causa delle inevitabili differenze tra la frequenza di lettura e quella di scrittura delle memorie elastiche, differenze riconducibili sia alle caratteristiche di funzionamento reale della rete di sincronizzazione sia alla presenza di

jitter e wander sui segnali STM-N afferenti al nodo, si devono prevedere meccanismi di giustificazione positiva-nulla-negativa dei contenitori virtuali da inserire nei flussi uscenti dal nodo. Nell'SDH tali giustificazioni vengono effettuate utilizzando il puntatore di VC con le seguenti modalità:

- la segnalazione di giustificazione è effettuata complementando, nella configurazione di bit del puntatore della trama precedente, le cifre d'ordine pari o dispari, a seconda che sia richiesta la giustificazione positiva o negativa;
- la giustificazione viene attuata in corrispondenza alla prima opportunità di giustificazione, e un incremento di +1 o -1 del valore del puntatore non complementato si attua nella trama successiva (questi incrementi non possono susseguirsi ad intervalli di durata inferiore a quella di quattro trame).

Per chiarire il funzionamento del meccanismo di puntatore, esaminiamo in dettaglio quali sono le operazioni a cui è sottoposto un VC che transita in un ADM o in un DXC; anche se per altri contenitori (VC-12, VC-11, VC-2) si rendono necessarie due elaborazioni di puntatore in cascata, per semplicità considereremo il caso di un VC che subisce un'unica elaborazione di puntatore (come avviene per il VC-4 e il VC-3). I cronosegnali in gioco nell'operazione di ritemporizzazione del flusso associato al VC in ingresso sono, come mostrato in Figura 9, quello estratto dal flusso STM-N d'ingresso e quello locale del nodo SDH.

I due cronosegnali, come già osservato, non si trovano mai in relazione di 'perfetto sincronismo' ma, praticamente, sono mesocroni, quando la rete di sincronizzazione funziona correttamente, o plesiocroni, quando, per esempio, il nodo di rete che trasmette il flusso sincrono non è più asservito all'orologio primario della rete di sincronizzazione e funziona in oscillazione libera: in tutte e due queste situazioni sarà necessario attuare,

anche se con cadenze diverse, degli aggiustamenti di puntatore sul flusso STM-N in uscita. La necessità di questi aggiustamenti viene determinata in funzione dello stato di riempimento del buffer indicato in Figura 9. Come nel caso del sincronizzatore di Figura 8, attuando le opportune giustificazioni (positive, nulle o negative) si compensano le eventuali differenze tra l'orologio estratto in ingresso e quello generato localmente nel nodo. Si può affermare, quindi, che il meccanismo di puntatore permette alla rete numerica basata sull'SDH di operare anche in condizioni di plesiocronia tra i nodi della rete stessa: in questa condizione, però, si risconterà un'altissima attività di puntatore (centinaia di aggiustamenti al secondo).

4.3 Estrazione dei tributari plesiocroni da un segnale STM-N

E' evidente che, quanto più ci si avvicina a una condizione di funzionamento della rete di sincronizzazione in cui sono assenti degradazioni sui segnali di sincronismo, tanto più l'intervallo di tempo tra successivi aggiustamenti di puntatore tende a crescere: è bene sottolineare che questa condizione di funzionamento è un obiettivo primario nel progetto e nella gestione della rete poiché migliora la qualità, espressa usualmente in termini di jitter e wander, dei flussi tributari che vengono ricostruiti dopo essere stati trasportati nella rete sincrona. Infatti, i salti di fase prodotti dagli aggiustamenti di puntatore, che risultano di ampiezza pari a uno o tre ottetti a seconda del VC considerato, impongono elevate prestazioni, e quindi elevata complessità realizzativa, agli apparati di desincronizzazione presenti nei demultiplicatori sincroni. Il desincronizzatore SDH, il cui schema funzionale è mostrato in Figura 10, ha una funzione analoga a quella esaminata nel caso plesiocrono, ed è progettato in modo da ridistribuire opportunamente nel tempo i bruschi salti di fase che si presentano sul

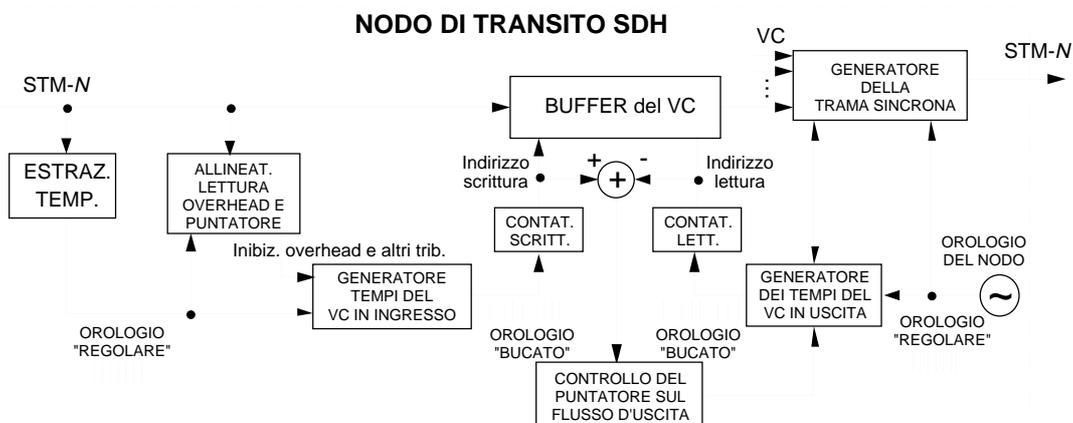


Figura 9 Schema funzionale delle operazioni attuate su un VC nel transito di un nodo SDH

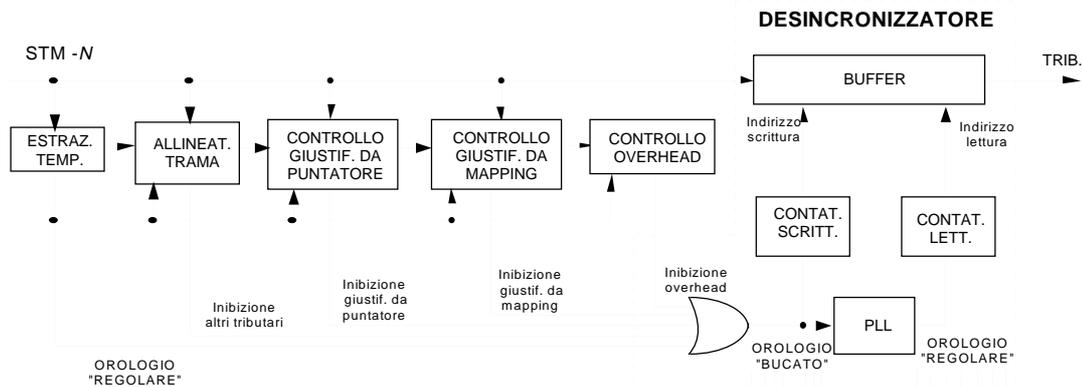


Figura 10 Schema funzionale dell'estrazione di un segnale tributario da un segnale STM-N

segnale di temporizzazione in ingresso al buffer (orologio bucato). Tale temporizzazione è ottenuta inibendo tutti gli istanti caratteristici del flusso STM-N che non sono corrispondenti a bit del segnale tributario da ricostruire. Mentre i salti di fase corrispondenti a bit di overhead si presentano in modo deterministico e producono sui tributari ricostruiti un jitter caratterizzato da un contenuto spettrale con componenti nell'intorno della frequenza di trama, i salti di puntatore e quelli connessi con la giustificazione attuata nel sincronizzatore si presentano casualmente nel tempo e producono prevalentemente jitter con componenti spettrali a frequenza molto bassa (wander). Il desincronizzatore, che si comporta come un filtro passa basso per le instabilità di fase presenti al suo ingresso, deve essere in grado di eliminare le componenti di instabilità ad elevata frequenza (connesse con l'overhead) e di distribuire su tempi sufficientemente lunghi i salti di fase occasionali connessi con le giustificazioni. Per quantificare le ampiezze dei fenomeni in gioco, si consideri che, per i flussi tributari PDH a standard europeo, i salti di fase che il desincronizzatore si trova a dover gestire in conseguenza a un singolo aggiustamento di puntatore hanno un'ampiezza⁽⁴⁾ che varia da un minimo di 160 ns (nel caso del VC-4) a un massimo di 3,6 μ s (nel caso del VC-12). Tali instabilità, se non opportunamente distribuite nel tempo sul segnale in uscita dalla rete SDH, possono risultare di gran lunga maggiori di quelle accettate alle interfacce degli apparati della gerarchia plesiocrona, rendendo di fatto incompatibile l'utilizzazione contemporanea in rete degli apparati delle due gerarchie. Per questo motivo, la rete di sincronizzazione che si adotterà per temporizzare gli apparati sincroni dovrà assicurare caratteristiche di elevata affidabilità e stabilità nel breve e nel lungo termine.

(4) Si noti che i salti di fase, che per definizione sono espressi in radianti, possono essere ricondotti, una volta definito il periodo del cronosegnale cui sono associati, a corrispondenti salti nel tempo e, quindi, misurati in secondi.

Bibliografia

- [1] Mariconda, A.; Misul, R.C.; Parente, F.; Pietroiusti, R.: *La nuova gerarchia di moltiplicazione sincrona (SDH)*. «Notiziario Tecnico SIP», n.1, anno 1, luglio 1992, pp.38-55.
- [2] Mariconda, A.; Misul, R.C.; Parente, F.; Pietroiusti, R.: *Apparati per la rete SDH*. «Notiziario Tecnico SIP», n.1, anno 2, aprile 1993, pp.49-63.
- [3] Decina, M.; Roveri, A.: *Introduzione alle reti telefoniche analogiche e numeriche*. La Goliardica, 1985.
- [4] Angelini, G. et. al.: *Manuale di reti di telecomunicazioni e trasmissione dati*. Calderini, Bologna, 1990.
- [5] Brugia, O.; Carbonelli, M.; De Seta, D.; Perucchini, D.: *Gerarchia numerica sincrona: una risposta alle esigenze della nuova rete di trasporto*. 39° Convegno internazionale delle comunicazioni, Genova, 1991.
- [6] Carbonelli, M.; De Seta, D.; Perucchini, D.; Ruffini, S.: *Synchronization of SDH networks: slave clock model and stability measures*. Proc. Globecom '92, Orlando, dicembre 1992, pp.829-833.
- [7] Sexton, M.; Reid, E.: *Transmission networking: SONET and the Synchronous Digital Hierarchy*. Artech House, Boston, 1992.
- [8] *Synchronous digital hierarchy bit rates*. Rac. CCITT G.707, Blue book, vol.III, fascicolo III.4, Melbourne, 1988.
- [9] *Network node interface for the synchronous digital hierarchy*. Rac. CCITT G.708, Blue book, vol.III, fascicolo III.4, Melbourne, 1988.
- [10] *Synchronous multiplexing structure*. Rac. CCITT G.709, Blue book, vol.III, fascicolo III.4, Melbourne, 1988.
- [11] *Characteristics of synchronous digital hierarchy (SDH) equipment functional blocks*. Rac. CCITT G.783, Ginevra, novembre 1992.
- [12] *Architectures of transport networks based on the synchronous digital hierarchy (SDH)*. Rac. CCITT G.803, Ginevra, giugno 1992.

Misure ed analisi di traffico

G. Giacobbo Scavo (*)

Nell'articolo viene sottolineato come le misure di traffico siano di crescente importanza per i gestori in relazione all'evoluzione del quadro giuridico-istituzionale in cui operano, ed alle caratteristiche della fase che il processo di sviluppo delle Telecomunicazioni sta attraversando. Per alcuni servizi è infatti superato il periodo di rapida espansione e l'ulteriore crescita appare condizionata da vari fattori, la cui valutazione prospettica è piuttosto aleatoria. Ciò rende problematica l'elaborazione di previsioni attendibili che, d'altra parte, assumono particolare criticità, soprattutto per i servizi in cui è aperta la competizione tra più soggetti e si vanno pertanto riducendo i margini di profitto.

In concreto, una vera sfida si sviluppa sul piano del costo e su quello del grado di servizio; attraverso il timone costituito appunto dalle misure di traffico è possibile governare la rotta senza che l'abbassamento del primo comporti necessariamente il peggioramento del secondo o viceversa.

Dopo una breve descrizione delle principali finalità delle misure di traffico, ci si sofferma su quelle necessarie per la progettazione e la valutazione delle prestazioni della rete nonché sulle problematiche aperte in questo campo dall'introduzione dei nuovi servizi.

1. Introduzione

Nei decenni passati la maggior parte dei gestori europei ha operato in un quadro giuridico ed economico notevolmente vincolato.

Oggi che questo scenario è in rapida evoluzione verso la competizione fra più gestori e l'applicazione del criterio del "price-cap"⁽¹⁾ per la determinazione delle tariffe, diventa essenziale dimensionare correttamente le risorse della rete e ridurre i margini temporali fra progettazione, realizzazione, attivazione e utilizzo a traffico degli impianti. Fondamento ineludibile per raggiungere questi obiettivi è la completa, continua, pronta conoscenza della domanda di traffico e dello stato della rete.

Se, negli anni trascorsi, si poteva sostenere che un sovradimensionamento della rete comportasse solo un anticipo di capitale⁽²⁾ perché le risorse in sovrannumero sarebbero state comunque utilizzate nel seguito, oggi, in relazione alla flessione della crescita di utenza e di traffico ed alla presenza di altri gestori concorrenti nel segmento di utenza a più alto traffico, un siffatto approccio sarebbe estremamente pericoloso, perché le risorse rischierebbero di non essere mai utilizzate.

Inoltre, soprattutto in questa fase di introduzione di Nuovi Servizi, è necessario analizzare il comportamento dell'utente e determinare le caratteristiche tecniche per ogni servizio, al fine di mettere a punto i modelli di previsione, validare gli algoritmi di dimensionamento, determinare il costo del singolo servizio.

Per i suddetti motivi, soprattutto nell'attuale situazione di congiuntura dell'economia nazionale, è vitale che un'azienda che vende servizi per Telecomunicazioni conosca puntualmente la domanda di traffico che i propri clienti esprimono, al fine di modulare oculatamente le previsioni e quindi gli investimenti.

2. Finalità delle misure

Per progettare e gestire la rete di Telecomunicazione bisogna, ovviamente, conoscere le caratteristiche quantitative e qualitative del traffico offerto alla rete.

(1) Con questo nome viene chiamato il sistema introdotto in Gran Bretagna dalla Oftel, controllore del servizio di Telecomunicazioni, per stabilire l'incremento complessivo delle tariffe in modo indipendente dagli introiti del gestore: tale incremento viene fissato pari all'aumento dei prezzi al consumo diminuito di una percentuale prestabilita che serve da stimolo al miglioramento dell'efficienza del gestore.

(2) Non è poco se si considera l'alto costo del denaro.

(*) ing. Giuseppe Giacobbo Scavo - SIP DG - Roma

Più puntuale è la conoscenza della variabilità dei diversi flussi di traffico nel corso del giorno, della settimana e dell'anno, minori sono i margini di sicurezza da adottare nelle previsioni. Senza questa premessa, ogni strategia tesa a ridurre i costi della rete rispettando gli obiettivi di grado di servizio, come l'adozione di sofisticati algoritmi di dimensionamento, o la scelta di instradamenti più evoluti⁽³⁾, sarebbe inutile, velleitaria, pericolosa.

Le attività tecniche che maggiormente necessitano delle misure di traffico sono: la pianificazione e progettazione della rete, nonché la valutazione del grado di servizio, l'esercizio e la gestione del traffico. I flussi di dati e di azioni che queste attività si scambiano sono schematizzati in fig. 1 attraverso tre anelli, ognuno dei quali sta in un suo strato temporale, ma si fonda su entità comuni alle altre attività: le misure di traffico e i parametri da queste ultime derivabili.

Nel primo anello le misure di traffico devono essere effettuate con continuità, i dati devono essere forniti e analizzati in tempo reale perché da essi derivano le azioni di "Network Management"⁽⁴⁾, come il reinstradamento di certi flussi o il filtraggio di certi traffici. Queste azioni sono finalizzate alla massimizzazione della capacità di smaltimento di traffico della rete, poiché, essendo l'orizzonte temporale di questi interventi dell'ordine dei minuti, il numero di risorse disponibili non è gestibile: è quello fissato in sede di progetto. La brevità di questi tempi è imposta dalla rapidità con cui, in presenza di sovraccarichi o guasti, la quantità di traffico smaltito dalla rete diminuisce, soprattutto a causa dell'azione di reciproco contrasto fra le sorgenti di traffico (utenti) e il sistema di servizio (autocommutatori).

Infatti l'utente, in situazione di congestione, ripresenta al sistema il suo tentativo operando una successione di chiamate ripetute; l'autocommutatore, diagnosticato lo stato di congestione, respinge per "time out"⁽⁵⁾ una percentuale sempre maggiore di tentativi. Poiché esso non è in grado di identificare la causa della congestione o il punto della rete in cui si manifesta, respinge le chiamate in modo indiscriminato, e non solo quelle che sarebbero comunque bloccate dalle strozzature della rete a valle dell'autocommutatore; in tal modo estende la congestione ad altri flussi di traffico, provocando la reazione di altri utenti e perciò la diffusione del fenomeno

delle chiamate ripetute. Questo circolo vizioso può portare la rete al degrado più totale in un tempo dell'ordine delle decine di minuti.

Il secondo anello, che viene percorso in un tempo dell'ordine delle ore e dei giorni, rappresenta le attività di Esercizio e Manutenzione della rete. Tali attività richiedono misure continue in modo da rilevare per tempo i guasti.

Il terzo anello, poiché comporta la realizzazione di ampliamenti della rete, viene percorso in tempi molto più lunghi. I dati di misura utilizzati in questa fase servono essenzialmente come base per la progettazione e la pianificazione della rete, attività che traggono orizzonti temporali rispettivamente dell'ordine dei mesi o degli anni.

Questa schematizzazione dei più frequenti utilizzi delle misure di traffico⁽⁶⁾ evidenzia che *l'attività di misura deve essere tale da fornire e conservare i dati utili in tutti e tre gli anelli, individuando quelle misure che sono comuni a tutti (la maggior parte) e specializzando invece, in funzione dell'utilizzazione, i periodi di osservazione, il tempo di analisi, il livello di sintesi dei risultati.*

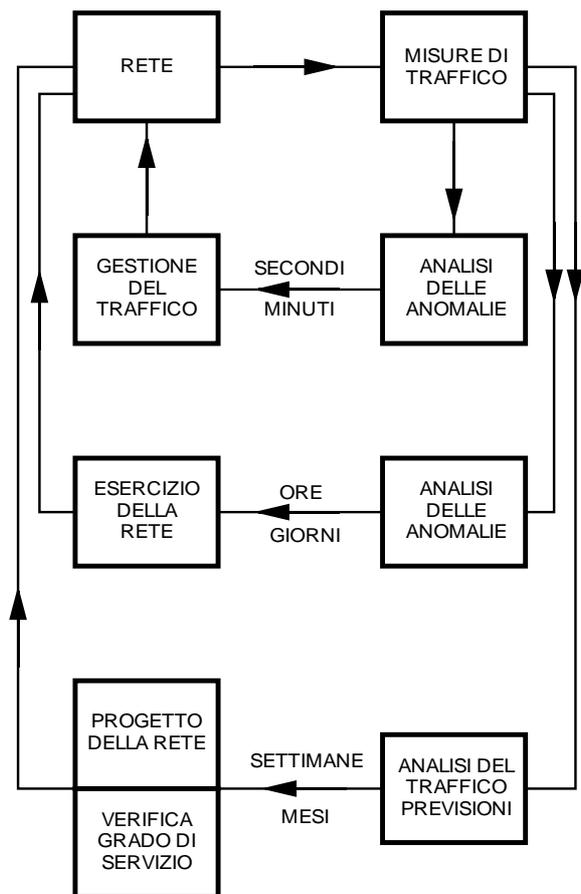


Figura 1 Attività che necessitano di misure di traffico e scale temporali delle relative azioni

(3) Non gerarchici, dipendenti dal tempo, adattativi [2].

(4) Gestione della rete e del traffico.

(5) Quando una chiamata presenta una richiesta di instradamento ad un autocommutatore, questo, se non riesce a soddisfarla entro un prefissato intervallo di tempo ("time out"), la blocca, cercando così di alleggerire il proprio carico.

(6) Ci sono altri utilizzatori delle misure di traffico, come: il marketing, la politica tariffaria, la valutazione degli introiti.

3. Tipologia delle misure

Ogni flusso di traffico che interessa un autocommutatore può essere associato ad un ingresso e/o ad una uscita. Considerato il caso generale di un autocommutatore con funzioni sia di centrale locale che di transito, i diversi flussi possono essere schematizzati, a livello globale, come in fig. 2.

In funzione del tipo di utilizzo viene richiesto un diverso grado di dettaglio; si possono così distinguere le seguenti principali categorie di misure:

- *misure globali*: forniscono i parametri caratteristici dei flussi di traffico illustrati in fig. 2, come il numero di impegni ed il numero di tentativi perduti per congestione per ognuna delle aliquote di traffico utente-utente, utente-giunzione, giunzione-utente e giunzione-giunzione;
- *misure sui fasci di circuiti*: facendo riferimento ad un particolare fascio, vengono forniti i valori relativi al traffico smaltito, al numero di impegni, al numero di tentativi persi, al numero di circuiti realmente in servizio;
- *misure sui "call record"*: per ogni chiamata vengono fornite le entità che la caratterizzano, come: la sorgente, l'istante di impegno, le cifre selezionate, gli esiti; questo tipo di misura, poiché costituisce un non trascurabile carico per l'autocommutatore, viene effettuata per campionamento. Per quanto riguarda la progettazione della rete, la misura sui "call record" porta un notevole contributo soprattutto in fase di cambiamento della struttura della rete e degli instradamenti del traffico, perché fornisce gli interessi di traffico fra i vari nodi e, perciò, offre gli elementi per istituire fasci nuovi e per ridimensionare quelli esistenti dopo un cambiamento di struttura; *quando la rete è stabile e molto magliata, le misure sui "call record" assumono minore importanza, perché le misure sui fasci sono sufficienti a "monitorare" lo stato della rete, determinando periodicamente gli eventuali adeguamenti necessari; inoltre, i dati forniti dalle misure di intensità di traffico sui fasci consentono*

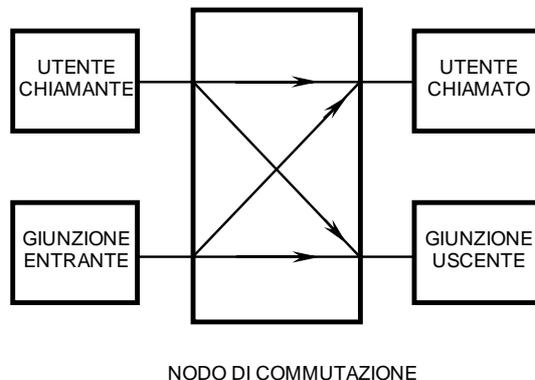


Figura 2 Flussi di traffico in un nodo di commutazione

di definire un sistema di tante equazioni quanti sono i fasci della rete e tante incognite quante sono le relazioni di traffico e, quindi, *di determinare anche gli interessi di traffico fra i nodi senza dover ricorrere alle misure di analisi*; in queste equazioni, i traffici smaltiti sono i termini noti, e i gradi di perdita misurati sui fasci sono le costanti; se la rete è completamente magliata, allora il sistema di equazioni è eguale al numero di incognite, cioè il numero dei fasci è eguale a quello dei flussi di traffico da nodo origine a nodo destinazione; perciò il problema è facilmente e totalmente risolvibile; altrimenti, bisogna ricorrere ad un algoritmo più sofisticato, che non fornisce i valori veri delle incognite, perché, in tal caso, il numero di incognite è maggiore del numero di equazioni, ma dà risultati che stanno in un intorno del 7% del valore vero [4, Racc. E.501 dell'ITU-TS] (per maggiori dettagli v. Appendice). Le misure sui "call record" assumono ancora importanza nella fase attuale, caratterizzata da un orientamento verso una rete integrata e cioè dall'utilizzo di risorse comuni per più servizi. Poiché è presumibile che questi ultimi abbiano fattori di crescita diversi, per essere in grado di formulare previsioni corrette è necessario conoscere il traffico verso ognuno dei vari servizi;

- *misure su ISDN*: c'è un'analoga tipologia di misure: globali, sui fasci, sui "call record", con in più le misure per ciascun servizio, le quali forniscono il numero di tentativi, il numero di chiamate che hanno ottenuto il segnale di risposta, il traffico smaltito.

4. Determinazione del periodo di misura e del carico della rete

Tradizionalmente le misure di traffico sono quelle che forniscono il traffico smaltito da un fascio di giunzioni. I criteri esistenti in ambito internazionale per la definizione del periodo di punta e del carico della rete si riferiscono unicamente a questo tipo di misure e, quindi, al singolo fascio, per il quale *l'ora di punta è definita come quella composta dai quattro quarti d'ora consecutivi la cui somma dei traffici è massima nel periodo osservato* (Racc. E.500 dell'ITU-TS).

Supponendo di tenere sotto osservazione un fascio per un giorno, la definizione del periodo di punta rispetto a cui dimensionare il fascio e valutare il grado di servizio ha una sua logica; ma se consideriamo di tenere il fascio sotto osservazione continuamente per un anno, la precedente osservazione risulterebbe un po' azzardata, perché rischieremo di dimensionare in base ad un evento sporadico nell'anno, per es. la mezzanotte del 31 dicembre.

Per questo motivo, per cautelarsi rispetto ad eventi eccezionali, si effettua una media su più giorni. Il metodo più usato, detto TCBH (*Time Consistent Busy Hour*), consiste appunto nell'elaborare il profilo di traffico del giorno medio, sommando i traffici dei giorni

esaminati quarto d'ora per quarto d'ora e dividendo il risultato per il numero dei giorni. Detti:

n il numero di giorni presi in esame;

A_{ij} il traffico smaltito nel i -esimo quarto d'ora ($i = 1, \dots, 96$) del giorno j ($j = 1, \dots, n$);

l'intensità media M_i del traffico smaltito nel i -esimo quarto d'ora del giorno, con riferimento agli n presi in esame, è:

$$M_i = \frac{1}{n} \sum_{j=1}^n A_{ij} .$$

La successione degli $M_i, i = 1, \dots, 96$, costituisce perciò il profilo di traffico del giorno medio; applicando ad esso la definizione di ora di punta, cioè identificando i quattro quarti d'ora consecutivi la cui somma dei traffici è massima nella giornata, si determina la cosiddetta TCBH, cioè l'ora di punta del giorno medio del periodo considerato.

Questo metodo, che dagli albori della telefonia veniva adottato per determinare il numero di operatori necessari nei vari turni del giorno, è basato sull'assunzione di una certa stabilità giornaliera dei profili di traffico. Il risultato ottenuto è l'appiattimento dei picchi dei vari giorni.

Un metodo diverso, adottato da Deutsche Bundespost Telekom, consiste nel determinare la punta di ciascun giorno e nel fare poi la media fra queste punte, indipendentemente dalla loro contemporaneità. Per questo motivo, l'ora così definita è detta ADPH (Average Daily Peak Hour).

Questo criterio, consentendo di conservare per ciascun giorno solo il valore di traffico dell'ora di punta, richiede una minore quantità di dati da memorizzare ed una più semplice elaborazione.

Il traffico nella ADPH, che risulta sempre maggiore o eguale a quello della TCBH, è più rappresentativo dei profili di traffico considerati, quando questi ultimi sono instabili, vale a dire quando i profili giornalieri differiscono tra loro in forma e volume.

La normativa internazionale citata (Racc. E.500 dell'ITU-TS), ai fini del dimensionamento dei fasci della rete, definisce "carico normale" quello dei 30 giorni a più alto traffico dell'anno, mentre, ai fini del dimensionamento degli autocommutatori, definisce quello dei 10 giorni a più alto traffico⁽⁷⁾.

Per quanto riguarda le misure per la pianificazione della rete, la questione più dibattuta, soprattutto nel passato, è stata quella relativa alla *continuità delle misure*. Infatti, in ambiente analogico, il costo delle misure era sensibile, perché bisognava usare apparecchiature esterne all'autocommutatore. Con il criterio della discontinuità delle misure, vale a dire prefissando i periodi di misura per ciascun autocommutatore, era possibile usare un numero limitato di apparecchiature spostandole, quando necessario, da un posto all'altro.

L'introduzione di autocommutatori SPC (Stored Program Control) ha consentito di superare l'onerosità del rilievo della misura, ma si è creato un collo di bottiglia nella fase di elaborazione. Così, anche per la rete numerica ci si trova ancora a doverla analizzare, progettare e pianificare sulla base di un numero limitato di giorni all'anno.

Perciò si adotta una metodologia statistica per stimare il traffico del periodo di punta dell'anno. Essa consiste nell'assumere gli N giorni⁽⁸⁾ di misura come campioni casuali della popolazione di giorni (periodo base) candidati a contenere i 30 giorni di punta dell'anno. La media dei 30 giorni a più alto traffico (M) viene calcolata dalla media (m) e dalla deviazione standard (σ) dei valori degli N giorni campionati, utilizzando un fattore moltiplicativo k che si desume dalla statistica:

$$M = m + k\sigma$$

dove, detto A_i il traffico dell'ora TCBH del giorno i ($i = 1, \dots, N$), la media m del traffico nei giorni campionati è data da

$$m = \frac{1}{N} \sum_{i=1}^N A_i ,$$

mentre la deviazione standard σ è espressa da

$$\sigma = \left[\frac{1}{N-1} \sum_{i=1}^N (A_i - m)^2 \right]^{1/2} .$$

Il fattore k , che si ricava dalle tavole di statistica riferentesi alla distribuzione normale, e che è riportato nella Racc. E.500 dell'ITU-TS, è funzione del numero dei giorni a più alto traffico che si ricercano (30 o 10) ed anche della lunghezza del periodo base, cioè, come detto, del periodo candidato a contenere i giorni di punta. La lunghezza di tale periodo può variare da un mese per le aree con traffico stagionale a tutto l'anno; in corrispondenza il valore di k da utilizzare per calcolare il traffico dei 30 giorni a più alto traffico dell'anno varierà tra 2 e 0. Bisogna ancora aggiungere che, affinché la stima sia affidabile, il numero di giorni misurati non deve essere inferiore a 30.

(7) La stessa Raccomandazione definisce anche un "carico alto" per i fasci e per gli autocommutatori: quello dei 5 giorni a più alto traffico dell'anno. Il grado di servizio che la rete deve rispettare è, ovviamente, meno stringente in condizioni di carico alto che in quelle di carico normale, ma la sua determinazione è stata rinviata, in ambito internazionale, a studi futuri (Racc. E. 721 dell'ITU-TS).

(8) Affinché il campione sia significativo N dovrebbe essere maggiore o eguale a 30.

5. Problematiche aperte

Il problema della limitatezza del numero di giorni di misura nell'arco dell'anno e delle sue conseguenze è stato affrontato dalla Compagnia Telefonica di Helsinki, la quale ha svolto uno studio molto approfondito [3], mettendo sotto osservazione i 2700 fasci della rete per 2 anni in modo continuo.

Irisultati indicano che *le misure discontinue forniscono dei valori in media più bassi di quelli reali del 7.6 %*. Questa valutazione è stata effettuata individuando per ogni fascio le due settimane di maggior traffico nel primo dei due anni analizzati e calcolando il traffico nelle stesse due settimane del secondo anno. Si è potuto constatare che, solo per l'11 % dei fasci, le due settimane di punta erano le stesse dell'anno precedente. Se si fossero volute individuare le due settimane realmente di punta per almeno il 50 % dei fasci, sarebbe stato necessario estendere il periodo di misura a tre mesi prima e tre mesi dopo le due settimane di punta del primo anno.

Ciò fa sì che la dispersione intorno al valore del 7.6% (sottostima media del traffico) sia ampia e tale che, se si richiedesse di non sottostimare il traffico del periodo di punta per almeno il 90% dei fasci, bisognerebbe incrementare di almeno il 20% il valore trovato nelle due settimane.

Per quanto riguarda la rete SIP, non sono state mai effettuate misure in modo continuativo per un intero anno; quindi, specialmente quando il periodo di misura era mediamente di soli 10 giorni l'anno (1991) per SGU (*Stadio di Gruppo Urbano*), esso veniva necessariamente posizionato nell'anno in base a sensazioni, ... premonizioni.

In quest'ultimo periodo di tempo si sta dando un notevole impulso all'attività di misura ai fini della progettazione, cercando di illustrarne gli scopi e di diffondere la sensibilità per questo tipo di problemi⁽⁹⁾. Si è così passati da un numero medio di 10 giorni di misura per SGU nel '91, a 30 giorni nel '92 e ci si aspetta un altro significativo miglioramento nel '93.

Inoltre, *la stessa definizione di ora di punta è molto datata*, infatti la durata del periodo di riferimento deve essere tale da fornire un'indicazione del grado di servizio corrispondente, in una qualche misura, alle impressioni dell'utente. Per esempio, garantire un determinato grado di servizio medio in un periodo di 24 ore non risulterebbe molto significativo e, soprattutto, non corrisponderebbe alle esigenze dell'utente. Perciò, fin dagli inizi del secolo, il periodo di riferimento è stato fissato in un'ora, perché questa durata era ritenuta proporzionata alle durate di conversazione (circa 20 volte il tempo di conversazione) e ai tempi di attesa per instaurare la chiamata. Oggi giorno, in rete numerica, i tempi di attesa sono praticamente nulli, e anche i tempi di impegno delle chiamate completate⁽¹⁰⁾ sono, in certe applicazioni, notevolmente inferiori ai normali tempi di conversazione.

Perciò, *per quanto riguarda il servizio di base, bisognerebbe, attraverso misure ed analisi ad hoc dei dati, valutare l'opportunità di passare al quarto d'ora. Per quanto riguarda i Nuovi Servizi, invece, almeno per alcuni di essi, l'ora non ha più alcun senso; bisognerebbe determinare, servizio per servizio, le condizioni di riferimento più opportune (durata del periodo di punta, grado di servizio, grandezze caratterizzanti il servizio).*

Si considerino ad esempio le videoconferenze, il cui tempo medio di impegno supera le due ore o il *televoting*, la cui durata media è, al contrario, brevissima. Per quest'ultimo servizio viene riportata in fig. 3 la misura del numero di chiamate effettuate su un campione della rete del British Telecom in un'ora e mezza e su un periodo di riferimento di 5, 15, 30 minuti.

Guardando il diagramma in cui i valori vengono mediati ad intervalli di 5 minuti, si può notare che da un intervallo all'altro il traffico cresce di un fattore 9 e decresce di un fattore 5. Ciò fa sì che, effettuando le misure ogni 15 minuti, come fanno attualmente gli autocommutatori, si misura, nel quarto d'ora di punta, un numero di tentativi (30) che è circa la metà di quelli manifestatisi nei 5 minuti di punta (72). Se poi si analizzano i dati ad intervalli di 30 minuti⁽¹¹⁾, si trova una punta (17) che è circa un quarto di quella dei cinque minuti.

Dalla definizione di ora di punta e di carico della rete, si potrebbe desumere che la progettazione della rete sia effettuata dimensionando ciascun fascio separatamente dagli altri. In realtà, il problema non è quello di dimensionare alcuni singoli elementi della rete, ma la rete nel suo insieme, garantendo lo smaltimento di tutti i flussi di traffico. Progettare la rete in base al periodo di punta di ciascuna relazione di traffico conduce ad un sovradimensionamento, perché ci sono elementi della rete condivisi da più flussi di traffico (il fascio XD di fig. 4, che viene impegnato sia dalla relazione XB che dalla XC), i quali possono andare in punta in periodi diversi. Progettare invece in base alla punta di ogni singolo nodo può portare ad un sottodimensionamento, come si può vedere dall'esempio della fig. 4.

Nella rete schematizzata, il traffico uscente dal nodo X e diretto a B e C viene smaltito in prima scelta sul fascio XB e XC e, in seconda scelta, attraverso il nodo di transito D.

Sia di 20 e 90 erlang il traffico offerto da X rispettivamente a B e C nell'ora 1 e di 70 e 10 erlang quello dell'ora 2. Perciò l'ora di punta dell'autocommutatore X è la prima, in cui si originano 110 erlang. Se si dimensiona per il traffico di

(9) E' quanto si sta facendo anche con questo articolo.

(10) Per chiamata completata si intende una chiamata che abbia ottenuto risposta.

(11) Come detto, oggi lo si fa ogni 60 minuti.

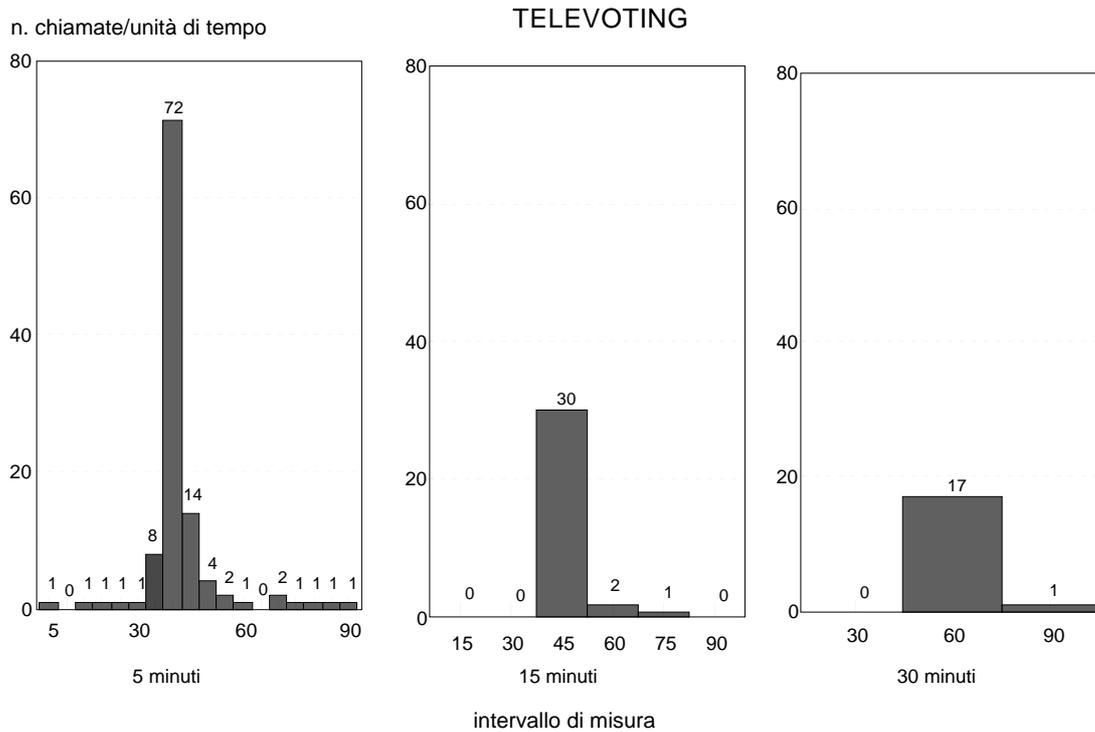


Figura 3 Andamento del numero di chiamate effettuate nell'arco di un'ora e mezza con periodi di riferimento di 5, 15 e 30 minuti rispettivamente (da un campione della rete del British Telecom)

quest'ora, supponendo di fissare una perdita del 20 % sui fasci diretti e dell'1 % per quelli di ultima scelta (il fascio XD), avremo il dimensionamento di fig. 4 a, cioè di 19 e 76 circuiti rispettivamente sui fasci XB e XC ed un trabocco di 3.8 e 17.5 erlang sul fascio XD, che viene perciò dimensionato con 43 circuiti.

Verificando il comportamento di questa rete nell'ora 2, in cui il traffico originato da X è globalmente inferiore, ma con interessi diversi, si osserva che dal fascio XB traboccano 51 erlang, che, offerti al fascio di ultima scelta

di 43 circuiti, danno luogo ad una perdita del 22 %, con grave danno per la relazione XB.

Come si vede, quindi, la definizione di carico rispetto a cui dimensionare la rete è più complessa e non esistono normative al riguardo. Gli studi effettuati suggeriscono di considerare tutte le ore in cui ci sono sensibili variazioni di interessi.

Gran parte dei problemi evidenziati sono conseguenza del fatto che, sia nel campo delle misure che in quello della progettazione, la cultura dominante è quella

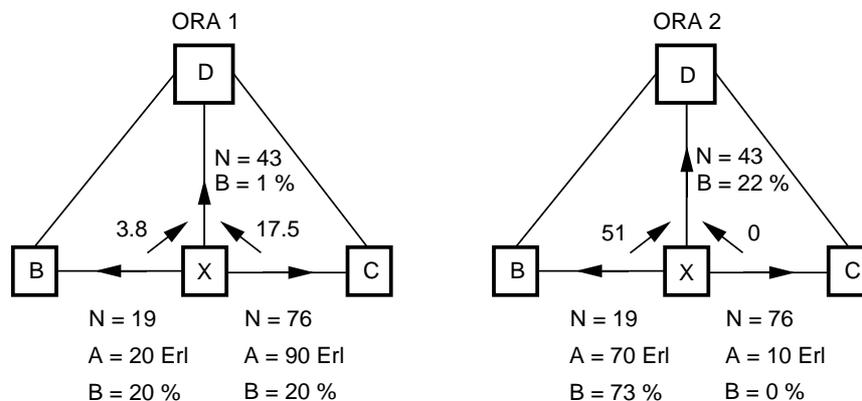


Figura 4 Schemi di rete utilizzati nell'esempio di dimensionamento: a) interessi di traffico dell'ora 1 (utilizzati nel dimensionamento); b) interessi di traffico nell'ora 2 (A = intensità media del traffico offerto (supposto poissoniano); N = numero di circuiti componenti ciascun fascio; B = grado di perdita del fascio)

“dell'autocommutatore”. Tale cultura tende ad analizzare nel dettaglio i singoli elementi della rete, affrontando separatamente i vari problemi. Questo approccio, pur avendo l'innegabile pregio della semplicità, da quando si è introdotta nelle reti una maggiore complessità di instradamento, risulta limitativo o del tutto errato, perché rischia di fare qualcosa di analogo a ciò che fanno i pianificatori delle città in cui viviamo: autostrade che terminano su viottoli, colli di bottiglia che, man mano che vengono effettuati lavori sulla rete, si spostano da un punto all'altro, congestioni che si diffondono rapidamente alle vie limitrofe. Questa cultura deve essere integrata da una più sistemistica, che gestisca la rete nel suo complesso, definendone i periodi di punta e valutandone le grandezze fondamentali caratterizzanti il comportamento della rete da origine a destinazione.

6. Conclusioni

Fino a poco tempo fa il processo di progettazione si sviluppava in tempi così lunghi da far sì che tra la prima fase (misura di traffico) e l'ultima (attivazione dei circuiti) passassero quattro anni; perciò le misure di traffico non potevano essere molto utili ai fini della progettazione, se non per evidenziare qualche isolato stato di sofferenza della rete che richiedeva interventi urgenti.

Per il suddetto motivo e per ragioni di altro genere, quali lo scenario monopolistico delle Telecomunicazioni e la tecnologia di commutazione nella quale era ancora presente l'analogico (difficoltà e costo delle misure, che dovevano essere effettuate con apparecchiature esterne agli autocommutatori), si è sviluppato un circolo involutivo: poche misure, rete sovradimensionata, perché progettata in base a previsioni di medio termine che, non potendo tenere in gran conto le misure effettuate sporadicamente quattro anni prima, dovevano ammettere un congruo margine di sicurezza; scarsa utilità delle misure, perché esse, in tale situazione, ben difficilmente rilevavano delle perdite e perciò servivano a poco.

Questa spirale viene oggi sostituita da un circolo virtuoso: misure puntuali e frequenti (periodicità mensile), previsioni di traffico a breve termine e quindi molto affidabili, dimensionamento ad hoc, con un ottimo grado di servizio, ma senza alcun margine se non quello dovuto alla modularità delle risorse. Per attuare con successo questa strategia bisogna disporre di misure frequenti, al fine di cogliere per tempo ogni possibile cambiamento nel comportamento dell'utente, per caratterizzare correttamente ogni Nuovo Servizio, per aggiornare le previsioni ed eventualmente le metodologie di dimensionamento, per modificare la rete con ampliamenti ed anche con diminuzioni, senza paura di riconoscere eventuali errori di misura, di previsione, di progettazione.

Bibliografia

- [1] Albert, A.: *Regression and the Moore-Penrose Pseudoinverse*. Academic Press, New York, 1972.
- [2] Giacobbo Scavo, G.: *Evoluzione dei criteri di instradamento del traffico nella rete*. «Notiziario Tecnico SIP», Vol. 1, n. 2, dicembre 1992, pp. 65-70.
- [3] Parviala, A.: *The Stability of Traffic Intensity Profiles and Its Influence on Measurement Schedules and Dimensioning*. International Teletraffic Congress 11, Kyoto, 1985.
- [4] Tu, M.: *Estimation of origin-destination traffic demand in dynamic routing networks*. International Teletraffic Congress 13, Copenhagen, giugno 1991, pp. 471-476.

Appendice

In questa Appendice si espone il metodo per la determinazione del traffico offerto origine-destinazione quando sono disponibili solo misure di intensità sui fasci⁽¹²⁾. Detti:

- F il numero dei fasci;
- R il numero di relazioni di traffico;
- M numero massimo di percorsi a disposizione di una generica relazione di traffico;
- a_i traffico offerto dalla relazione i , $i=1, \dots, R$;
- b_q traffico smaltito dal fascio q , $q=1, \dots, F$;
- O_{ij} traffico offerto al j -esimo percorso della relazione i , $j=1, \dots, M$, $i=1, \dots, R$;
- S_{ij} traffico smaltito dal j -esimo percorso della relazione i , $j=1, \dots, M$, $i=1, \dots, R$;
- B_{ij} grado di perdita del j -esimo percorso della relazione i , $j=1, \dots, M$, $i=1, \dots, R$;
- L_{ijk} grado di perdita del k -esimo ramo del j -esimo percorso per la relazione i , $k=1, 2$, $j=1, \dots, M$, $i=1, \dots, R$;
- $f(ijk)$ indice del k -esimo fascio che compone il percorso j -esimo della relazione i , $k=1, 2$, $j=1, \dots, M$, $i=1, \dots, R$; si conviene di porre $f(ijk)=0$ se le regole di instradamento prevedono meno di j percorsi alternativi per la relazione i ovvero se il j -esimo percorso è composto da meno di k rami; risulta perciò $0 \leq f(ijk) \leq F$;
- F'_q insieme delle coppie relazione-percorso che utilizzano il fascio q ($q=1, \dots, F$); formalmente $F'_q = \{(i, j) \mid f(ij1)=q \text{ o } f(ij2)=q, 1 \leq i \leq R, 1 \leq j \leq M\}$.

(12) L'illustrazione del metodo è ripresa dalla Racc. E.501 dell'ITU-TS.

Si noti che ad ogni collegamento ijk corrisponde un unico fascio q ($q=1, \dots, F$), ma ad un fascio q può corrispondere un certo numero di percorsi appartenenti in generale a diverse relazioni di traffico. Questa corrispondenza viene fornita dal valore assunto dalla funzione $f(ijk)$.

In base alle definizioni precedenti, risulta

$$S_{ij} = O_{ij}[1-B_{ij}]. \quad (\text{A.1})$$

Assumendo l'indipendenza delle probabilità di blocco dei due fasci componenti un percorso, si può inoltre scrivere

$$B_{ij} = L_{ij1} + L_{ij2} - L_{ij1}L_{ij2}, \quad (\text{A.2})$$

perciò, in presenza di "crankback"⁽¹³⁾, si ha:

$$O_{ij} = a_i \prod_{t=1}^{j-1} B_{it}. \quad (\text{A.3})$$

Dalle (A.1) e (A.3) si ricava

$$S_{ij} = a_i [1-B_{ij}] \prod_{t=1}^{j-1} B_{it} = a_i s_{ij}, \quad (\text{A.4})$$

dove

$$s_{ij} = [1-B_{ij}] \prod_{t=1}^{j-1} B_{it}. \quad (\text{A.5})$$

Perciò il traffico smaltito su ciascun fascio q sarà:

$$b_q = \sum_{(i,j) \in F_q} S_{ij} = \sum_{(i,j) \in F_q} s_{ij} a_i. \quad (\text{A.6})$$

Il traffico offerto da ogni relazione i può essere derivato dal sistema di equazioni (A.6). Possono essere applicati vari metodi per risolvere il sistema di equazioni, che può essere scritto in forma matriciale:

$$\mathbf{Z} \mathbf{a} = \mathbf{b}, \quad (\text{A.7})$$

dove

$$\mathbf{b} = [b_1, \dots, b_F]^T$$

$$\mathbf{a} = [a_1, \dots, a_R]^T$$

$$\mathbf{Z} = \begin{bmatrix} z_{11} & \cdots & z_{1R} \\ \vdots & & \vdots \\ z_{F1} & \cdots & z_{FR} \end{bmatrix},$$

dove per $q=1, \dots, F$ e $i=1, \dots, R$ si ha:

$$z_{qi} = \begin{cases} s_{ij} & \text{se il fascio } q \text{ è utilizzato dal} \\ & j\text{-esimo percorso della} \\ & \text{relazione di traffico } i \\ 0 & \text{altrimenti} \end{cases}.$$

Se, per risolvere il sistema di equazioni, viene usato il metodo della matrice pseudo-inversa [1], la soluzione del sistema (4) può essere rappresentata da:

$$\hat{\mathbf{a}} = \mathbf{Z}^+ \mathbf{b},$$

dove $\hat{\mathbf{a}}$ è il traffico di relazione stimato e \mathbf{Z}^+ è la matrice pseudo-inversa di \mathbf{Z} . Si può dimostrare che $\hat{\mathbf{a}}$ è il vettore che rende minimo l'errore quadratico nella risoluzione del sistema (A.7), espresso mediante la norma euclidea del vettore dei residui $\mathbf{b} - \mathbf{Z} \hat{\mathbf{a}}$. Se il sistema è quadrato, cioè se il numero di equazione è eguale al numero di incognite, la soluzione è univocamente determinata; infatti:

$$\mathbf{Z}^+ = \mathbf{Z}^{-1}$$

Per le reti non completamente connesse, cioè quando il numero di equazioni è inferiore al numero delle incognite, il sistema di equazioni non ha un'unica soluzione, perciò il traffico offerto per ogni relazione deve essere stimato, introducendo un errore che è ovviamente tanto maggiore quanto più diminuisce il numero dei fasci. Nel caso in cui le equazioni scritte in accordo al sistema (A.7), pur in numero minore delle incognite, siano tuttavia linearmente indipendenti, si ha

$$\mathbf{Z}^+ = \mathbf{Z}^T (\mathbf{Z} \cdot \mathbf{Z}^T)^{-1}$$

In tale caso, il metodo della pseudo-inversa coincide con quello dei minimi quadrati.

(13) Si tratta di una prestazione di instradamento evoluto, che può essere fornita in un ambiente costituito da autocommutatori numerici e rete di segnalazione a canale comune. L'autocommutatore, dopo avere instradato una chiamata su un fascio e perciò impegnato un altro autocommutatore, ricevendo da quest'ultimo un segnale di congestione con "crankback", dovuto allo stato di occupazione delle giunzioni uscenti, abbatte la connessione verso valle e cerca di reinstradare la chiamata sulla scelta successiva. In tal modo si passa dal concetto di trabocco per congestione su un fascio a quello di trabocco per congestione su un percorso origine-destinazione [2].

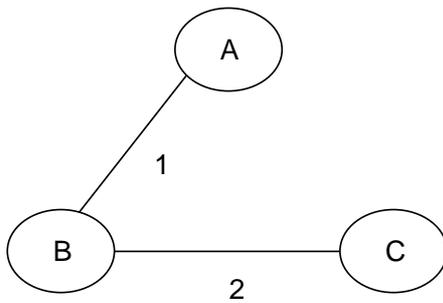


Figura 5 Rete utilizzata nell'esempio di applicazione del metodo della pseudo-inversa per il calcolo dei traffici offerti

Come esempio di quanto detto, si considera la rete di 3 nodi in fig. 5. In questo esempio $F=2, R=6$ ed $M=1$ (un solo percorso è possibile per ogni relazione di traffico).

Sia i l'indice delle relazioni di traffico fra nodo origine e nodo destinazione, detti 1 e 2 i due fasci rispettivamente tra A e B e tra B e C, l'instradamento di ognuna delle 6 relazioni di traffico è illustrato nella successiva Tabella 1.

La matrice \mathbf{Z} è data da:

$$\mathbf{Z} = \begin{bmatrix} s_{11} & 0 & s_{31} & s_{41} & 0 & s_{61} \\ 0 & s_{21} & s_{31} & 0 & s_{51} & s_{61} \end{bmatrix}$$

Se facciamo l'ipotesi che tutti i fasci siano dimensionati con il grado di perdita del 10%, allora: $s_{11} = 0.9, s_{21} = 0.9, s_{31} = 0.81, s_{41} = 0.9, s_{51} = 0.9, s_{61} = 0.81$.

Nel seguito si adottano come nuove incognite le somme dei traffici offerti dalle relazioni di traffico reciproche, ove due relazioni di traffico si definiscono reciproche se sono relative ai due versi opposti del traffico per una medesima coppia di nodi di rete. Chiaramente le nuove incognite sono in numero pari a $R/2$ (3 nel caso in esame). Con le nuove incognite, la matrice dei coefficienti del sistema si può riscrivere come segue:

Relazione di traffico	Percorso	$f(ij1)$	$f(ij2)$
1 (AB)	1	1	0
2 (BC)	2	2	0
3 (CA)	2,1	2	1
4 (BA)	1	1	0
5 (CB)	2	2	0
6 (AC)	1,2	1	2

Tabella 1 Relazioni di traffico e corrispondenti percorsi con riferimento alla rete in fig. 5

$$\mathbf{z}_s = \begin{bmatrix} 0.9 & 0 & 0.81 \\ 0 & 0.9 & 0.81 \end{bmatrix}$$

Se per esempio $b_1=5$ Erl e $b_2=7$ Erl, per quanto detto risulta:

$$\hat{\mathbf{a}}_s = \mathbf{z}_s^T (\mathbf{z}_s \mathbf{z}_s^T)^{-1} \mathbf{b}, \tag{A.8}$$

ove $\hat{\mathbf{a}}_s$ è il vettore delle stime dei traffici offerti.

Risulta perciò:

$$\begin{aligned} \mathbf{z}_s \mathbf{z}_s^T &= \begin{bmatrix} 0.9 & 0 & 0.81 \\ 0 & 0.9 & 0.81 \end{bmatrix} \begin{bmatrix} 0.9 & 0 \\ 0 & 0.9 \\ 0.81 & 0.81 \end{bmatrix} = \\ &= \begin{bmatrix} 1.4661 & 0.6561 \\ 0.6561 & 1.4661 \end{bmatrix} \end{aligned}$$

quindi

$$\begin{aligned} \mathbf{z}_s^T (\mathbf{z}_s \mathbf{z}_s^T)^{-1} &= \begin{bmatrix} 0.9 & 0 \\ 0 & 0.9 \\ 0.81 & 0.81 \end{bmatrix} \begin{bmatrix} \frac{1.4661}{1.719} & -\frac{0.6561}{1.719} \\ -\frac{0.6561}{1.719} & \frac{1.4661}{1.719} \end{bmatrix} = \\ &= \frac{1}{1.719} \begin{bmatrix} 1.3195 & -0.5905 \\ -0.5905 & 1.3195 \\ 0.6561 & 0.6561 \end{bmatrix} \tag{A.9} \end{aligned}$$

In base alle (A.8) ed (A.9) si ottiene in definitiva

$$\hat{\mathbf{a}}_s = \frac{1}{1.719} \begin{bmatrix} 1.3195 & -0.5905 \\ -0.5905 & 1.3195 \\ 0.6561 & 0.6561 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} \cong \begin{bmatrix} 1.43 \\ 3.65 \\ 4.58 \end{bmatrix} \text{Erl.}$$

Si può facilmente verificare che, offrendo alla rete questi traffici e considerando il grado di perdita del 10% su ogni fascio, vengono effettivamente smaltiti 5 erl sul fascio 1 e 7 erl sul fascio 2.

La rete precollegata fonicamente: l'esperimento del quartiere Murattiano di Bari

P. Fedele, R. Imperio (*)

Tra le prestazioni di rete da migliorare grazie all'evoluzione tecnologica, è da considerare con particolare attenzione il tempo necessario per rendere operativo il collegamento. Questa prestazione è infatti tra quelle che più qualificano i servizi offerti.

Per ridurre l'attesa del cliente al minimo, non basta ottimizzare i processi consolidati, ma occorre introdurre nuove filosofie di gestione della rete. Le soluzioni di rete cablate si basano sulla tecnica di realizzare il collegamento o larga parte di esso, in anticipo, prevenendo la domanda del cliente.

Tra queste reti, la rete precollegata fonicamente è un esempio di come l'integrazione delle tecniche, sviluppate singolarmente, possa creare una sinergia che permette un grosso salto di qualità nelle prestazioni della rete.

Nel seguito sono illustrate le procedure di costruzione e gestione di una rete precollegata fonicamente ed è descritta una effettiva realizzazione, sperimentata nel quartiere Murattiano di Bari.

1. Premessa

In Italia, nella grande maggioranza dei casi, la rete di raccolta di utenza è costituita da una rete di distribuzione in rame di tipo elastico. In tal caso sono precostituiti i collegamenti tra centrale ed armadio stradale ripartilinea (rete primaria) e tra armadio ripartilinea ed elemento distributore (rete secondaria).

Tale tipo di soluzione rinvia la costruzione della parte di rete in proporzione più onerosa ad un momento successivo alla stipula del contratto. Inoltre rende possibile modulare e riutilizzare nel tempo le risorse di rete primaria e di commutazione.

Di fatto, in una rete elastica, lo stato di occupazione, o di disponibilità delle risorse, di rete e di commutazione, è rappresentato dalla presenza o meno delle "permutazioni".

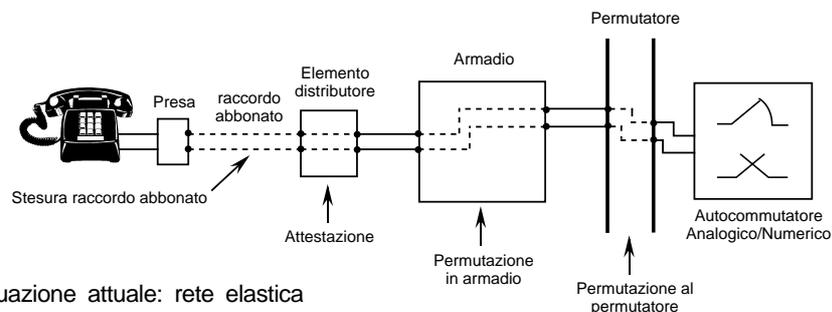


Figura 1 Situazione attuale: rete elastica

Al manifestarsi della richiesta di collegamento da parte di un cliente, si rende necessario costruire il raccordo tra la sua sede e la coppia di rete secondaria dell'elemento distributore. Completati in tal modo i vari tronchi di rete, occorre collegarli fisicamente, tramite permutazioni in armadio ed in centrale, compresa la risorsa di commutazione (come indicato nella fig. 1). Tale operazione avviene dopo la stipula del contratto da parte del cliente.

Pertanto si è costretti, alla rescissione del contratto, a disconnettere i collegamenti all'armadio ripartilinea ed al permutatore di rete per garantire la corretta gestione delle risorse di rete. La modalità di gestione della rete elastica è particolarmente indicata nelle aree con espansione della domanda. Nelle aree stabili e ad alta densità, con elevata mobilità interna, è estremamente probabile che un nuovo impianto sia in realtà il ripristino di un collegamento precedentemente cessato. In definitiva in tali aree territoriali la rete elastica assomiglia ad un puzzle, che viene continuamente disfatto e ricomposto.

(*) ing. Pasquale Fedele, ing. Raffaele Imperio - SIP DR Puglia - Bari

D'altronde proprio in tali realtà risiede solitamente la clientela a più alta intensità di utilizzazione dei servizi. Tale clientela è sempre più esigente e richiede alla rete prestazioni sempre migliori anche per quanto riguarda il tempo di costituzione del collegamento.

In una rete elastica il tempo per collegare la sede del cliente alla rete di telecomunicazioni è dato dalla somma dei tempi necessari per l'espletamento delle seguenti operazioni elementari:

- 1) acquisizione della domanda presso lo sportello fisico o telefonico;
- 2) emissione dell'ordinativo di lavoro;
- 3) costruzione del raccordo di abbonato;
- 4) realizzazione delle permutazioni presso l'armadio ripartilinea stradale ed il permutatore di rete in centrale.

Mentre alcune operatività possono avere tempi brevi (ad esempio l'acquisizione della domanda), altre richiedono anche tempi logistici di spostamento del personale (permutazioni) o utilizzano procedure "batch" presso i centri di elaborazione dati (emissione ordinativo di lavoro).

Il tempo limite a cui può tendere tale soluzione tecnica è pertanto di alcuni giorni.

2. Le reti cablate

2.1 L'evoluzione verso le reti cablate

Se si vuole ottenere che il tempo di realizzazione del collegamento sia inferiore al limite minimo occorrente nel caso di rete elastica, occorre evolvere verso diverse soluzioni di gestione della rete di raccolta. Infatti eventuali ottimizzazioni del processo descritto nella rete elastica non possono portare a sostanziali riduzioni del tempo occorrente. È necessario operare in modo che il collegamento risulti già preconstituito nel momento in cui il cliente esprime la sua richiesta.

Nelle reti cablate risultano già posate le prese presso le sedi dei clienti ed i relativi raccordi di abbonato. Questi sono rigidamente collegati al permutatore di rete mediante permutazioni precostituite negli armadi ripartilinea.

Cablando la rete, però, occorre tenere memoria delle risorse già impegnate dal cliente e di quelle in attesa di essere utilizzate al concretizzarsi della domanda. Pertanto è indispensabile un opportuno "software" capace di assicurare una corretta gestione della rete in tal senso.

2.2 Problematiche realizzative: l'"indirizzo"

Il problema cardine di una rete rigida fino ai clienti è costituito dalla definizione esatta ed univoca della sede del cliente, mediante un codice denominato "indirizzo".

Tale "indirizzo" non va confuso con il civico che individua non la singola sede del cliente, ma lo stabile. Può

invece essere una valida definizione di "indirizzo": "via, civico, complesso, scala, piano, interno". L'univocità e certezza dell'"indirizzo" è condizione essenziale, in quanto una volta che si è associato il cliente all'"indirizzo" è presso quest'ultimo che si attiva il collegamento preassegnato e pertanto un "indirizzo" errato comporta l'attivazione di un diverso cliente.

È necessario che le regole che identificano univocamente l'"indirizzo" non siano note solo al gestore, che realizzerà la rete cablata, ma anche ai suoi clienti, che dovranno esplicitarlo all'atto della domanda, governando in tal modo l'utilizzazione della rete stessa.

2.3 Applicazioni in Italia

Nel '93 Sip ha realizzato due tipologie di gestione della rete rispondenti alle esigenze prima descritte:

- Gestione della mobilità di utenza.
- Rete cablata.

La gestione della mobilità di utenza si basa sulla modifica della procedura commerciale di gestione utenza (TGU) e della banca dati descrittiva della rete di distribuzione (BDRR). In particolare tali procedure sono state modificate in modo da tenere memoria dei collegamenti cessati e della loro esatta ubicazione, identificata dall'"indirizzo". Tali collegamenti, quindi, sono visti come assegnati alle unità locative corrispondenti agli "indirizzi". Pertanto non vengono cessati realmente, ma restano costituiti sul piano "hardware". Alla prima richiesta verranno identificati e riassegnati mediante gestione virtuale in TGU ed ACL. Il collegamento cessato verrà comunque interrotto al permutatore urbano, per non vincolare la risorsa di commutazione.

La rete cablata, invece, prevede di realizzare, in occasione dei lavori di rete, il collegamento delle unità locative sulle quali si possa prevedere nel breve-medio termine il concretizzarsi di una richiesta di collegamento (fig. 2).

Il lavoro di impresa per la stesura del raccordo di abbonato risulta, a differenza di quanto avviene operando secondo i modi tradizionali, concentrato temporalmente

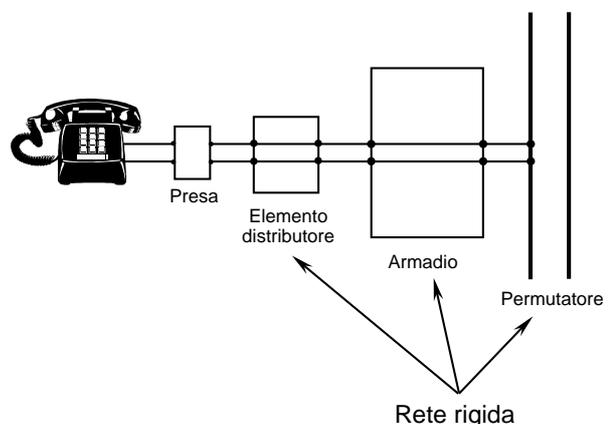


Figura 2 Rete cablata

e di consistenza rilevante nella località di intervento. Pertanto per tali allacciamenti si riesce ad ottenere un sostanziale risparmio di costo, grazie alle economie di scala e ad una ottimale programmazione di questi interventi, svincolati dalle scadenze temporali imposte dalle richieste della clientela.

I collegamenti da realizzare sono individuati mediante una previsione mirata, scaturita nel corso della indagine ecografica, eseguita dal progettista della rete. Sono successivamente gestiti in analogia a quanto descritto nella mobilità di utenza.

3. La rete precollegata fonicamente

3.1 Evoluzione della rete cablata

Una generica rete cablata prevede l'irrigidimento della distribuzione dalle sedi del cliente fino al permutatore urbano, situato nell'edificio sede dell'autocommutatore. Dove tale autocommutatore è completamente numerico, la rete cablata può evolvere verso una forma di rete che unisca, ottimizzandole, le prestazioni dell'irrigidimento della rete di distribuzione con tutte le prestazioni offerte dalla tecnica di commutazione numerica.

La rete precollegata fonicamente è l'evoluzione delle reti cablate verso una soluzione di precostituzione totale del collegamento, autocommutatore compreso.

In tale soluzione impiantistica, oltre a riportare rigidamente ogni sede di potenziale cliente fino al permutatore di competenza, viene realizzata anche la permutazione verso una porta di centrale libera.

3.2 Prerequisiti

La rete precollegata fonicamente permette di utilizzare in modo integrato tutte le prestazioni della tecnica numerica di commutazione. In particolare, le prestazioni

dell'autocommutatore numerico utilizzate nella rete precollegata fonicamente sono:

- Numerazione libera al permutatore.
- Prestazione Hot Linee.
- Gestione della categoria di utente.

3.2.1 Numerazione libera al permutatore

In una centrale analogica c'è una sostanziale coincidenza tra le porte di accesso (attacchi di utente) ed il numero di centrale.

Nelle centrali numeriche, invece, questo vincolo non è più esistente, ed è quindi possibile numerare gli attacchi di utente secondo qualsiasi ordinamento. Tale ordinamento è esplicitato tramite una tabella di corrispondenza biunivoca tra le porte della centrale e la loro numerazione.

La prestazione di "numerazione libera" al permutatore consiste nella possibilità di variare in tempo reale tale tabella di corrispondenze, tramite transazione software. In tal modo l'operatore, tramite terminale, può impostare qualsiasi associazione biunivoca tra le porte fisiche ed i numeri assegnati all'autocommutatore.

Quella appena descritta è una prestazione analoga a quella svolta dal permutatore di rete, che permette di modificare la corrispondenza tra le coppie distribuite ed i numeri di centrale, mediante interventi fisici sulle permutazioni.

Utilizzando tale analogia è possibile interpretare la prestazione di numerazione libera come un immaginario permutatore software, interposto tra i numeri e le porte di centrale.

Pertanto, la prestazione di numerazione libera inserisce a monte del permutatore di rete un permutatore software, duplicando la possibilità di variare la corrispondenza tra numeri e coppie distribuite. È pertanto possibile, se non opportuno, rendere inattivo uno dei due permutatori ed in particolare quello di rete, per gli indubbi vantaggi che derivano dal fatto di sostituire una operatività fisica con una transazione software (fig. 3).

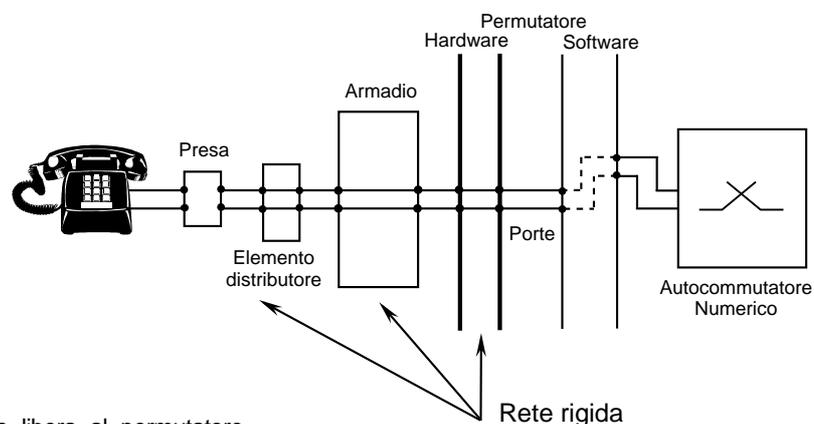


Figura 3 Numerazione libera al permutatore

3.2.2 La prestazione Hot-line

La prestazione di Hot line prevede che lo sgancio del microtelefono comandi direttamente la connessione del chiamante con un numero predisposto, senza la necessità di selezionare la corrispondente numerazione.

Tale selezione può partire immediatamente o dopo alcuni secondi per permettere la selezione di una numerazione diversa, disattivando in tal modo la prestazione hot line.

3.2.3 Gestione delle categorie di utente

In una centrale numerica è possibile assegnare fino a otto diverse "categorie d'utente". Tali categorie abilitano ai diversi tipi di traffico (urbano, teleselettivo, servizi di emergenza, etc.) e vanno dalla possibilità di effettuare qualsiasi selezione al funzionamento come terminale solo ricevente (nessun tipo di traffico uscente consentito).

3.3 L'"indirizzo" nella rete precollegata fonicamente

Il fatto di aver connesso l'autocommutatore alla rete cablata ha di fatto creato una corrispondenza biunivoca tra le unità locative servite e le porte di centrale. Queste, quindi, identificano in maniera univoca la sede da collegare. Sono pertanto un valido "indirizzo". Per tale motivo è superfluo, nella rete precollegata fonicamente, rilevare l'"indirizzo" inteso come "via, civico, complesso, scala, piano, interno".

3.4 Operatività

Nella rete precollegata fonicamente il collegamento è già interamente costituito dalla sede del potenziale cliente fino alla porta dell'autocommutatore.

L'evasione di una richiesta di collegamento consiste nella assegnazione di un numero commercialmente disponibile; nell'abbinamento porta-numero; nella gestione della categoria della porta di centrale. Una volta individuata la porta corrispondente alla sede del cliente basterà, infatti, abilitare tale porta a tutti i tipi di traffico.

Per rilevare l'"indirizzo" (porta di centrale) si possono seguire due possibili metodologie.

La prima consiste nel "leggere" il numero del chiamante (dalla segnalazione CCS7) da parte dell'operatore commerciale interessato. In questo caso, naturalmente, la richiesta deve essere effettuata dalla sede da attivare. In modo alternativo, si può assegnare alle porte da attivare la prestazione di hot line, in modo che, una volta che il cliente si collega ad essa, venga indirizzato verso una specifica casella fonica, ove una opportuna fonìa lo invita a prendere nota di un codice identificativo ed a chiamare lo sportello telefonico (187) (fig. 4).

La porta di centrale, configurata in hot line, con attivazione della prestazione dopo alcuni secondi, è abilitata anche alla selezione del 187, per permettere di rivolgersi allo sportello telefonico dallo stesso apparecchio.

Il codice identificativo comunicato al cliente è costituito da un identificativo della porta dell'autocommutatore preceduto da una cifra descrittiva della sede connessa (abitazione, ufficio, ente, moroso etc.).

4. Sperimentazione di Bari nel quartiere Murattiano

4.1 Il quartiere Murattiano

Il quartiere Murattiano costituisce il cuore della città di Bari, con una elevata concentrazione della clientela ad alta intensità di utilizzazione dei servizi. La forte espansione degli investimenti di rete, soprattutto nelle aree a maggior velocità di ritorno economico, ha prodotto in tale zona la presenza di buone risorse di rete. In particolare gli autocommutatori sono completamente numerici, con possibilità di numerazione libera e prestazione hot line, e la rete di distribuzione è quasi interamente a completa saturazione, con presenza pressoché totale di predisposizioni interne agli stabili.

Un territorio ideale per sperimentare processi di ottimizzazione ed integrazione nell'uso di tali risorse (fig. 5).

Disabilitazione traffico entrante
 Disabilitazione traffico uscente eccetto 187
 A 5 secondi dallo sgancio instradamento automatico sul fascio di giunzioni "caselle vocali" e commutazione all'utente del codice porta (xyz)

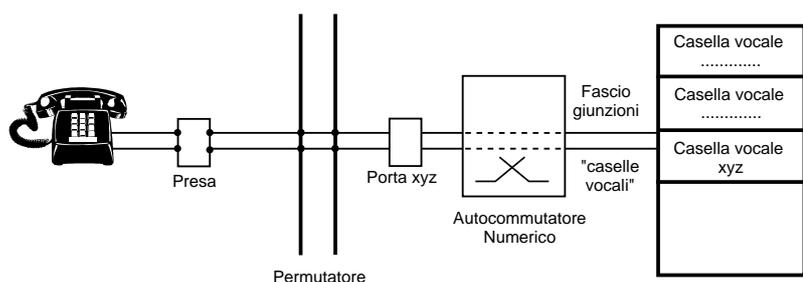


Figura 4 Impianto precollegato

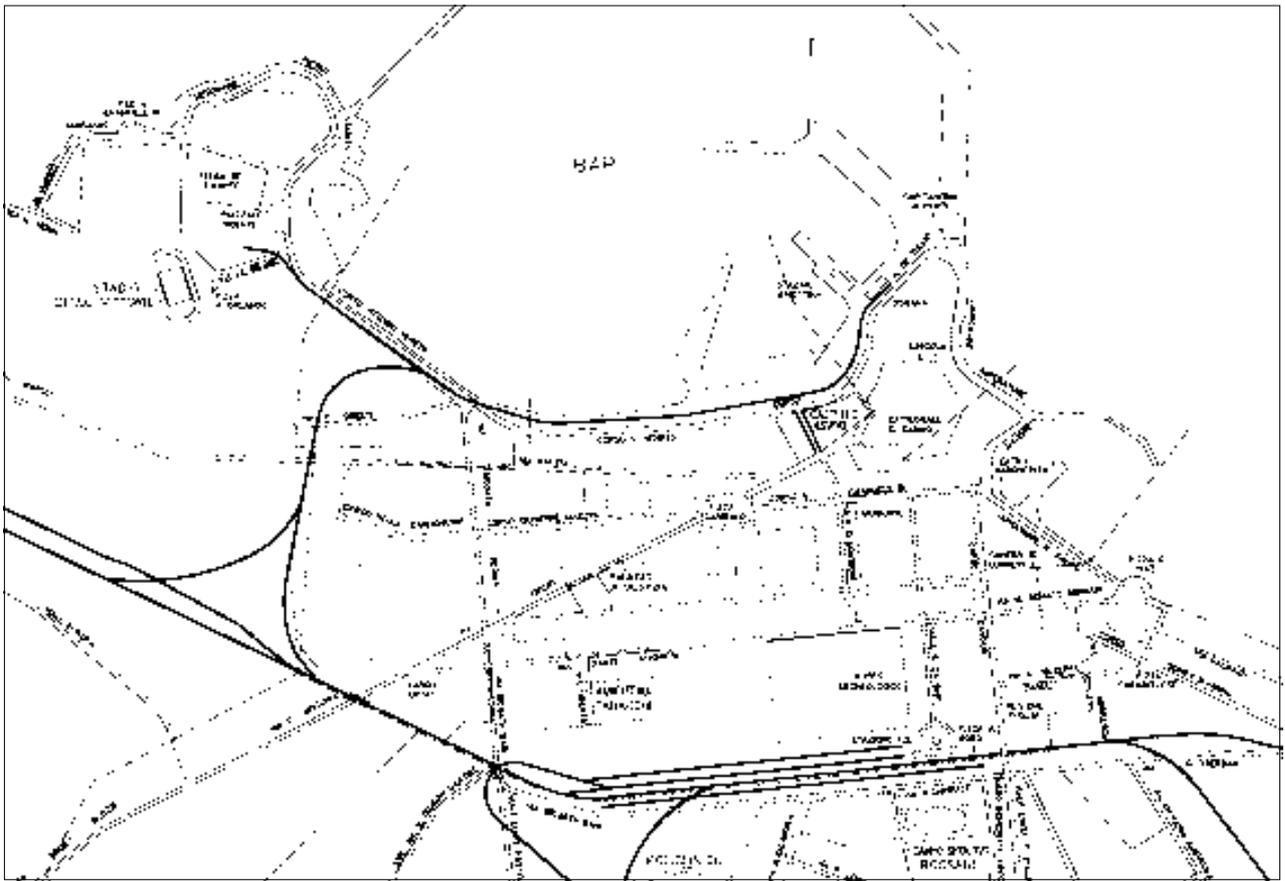


Figura 5

4.2 Problematiche gestionali

Volendo raggiungere entro l'anno '92 la realizzazione di una rete cablata, era condizione necessaria evitare ogni modifica delle procedure gestionali o consistenti variazioni di operatività. Pertanto, indipendentemente dalla tipologia di rete cablata, occorre definire come riconoscere il potenziale cliente della sperimentazione. Tale attività deve essere necessariamente attuata dall'operatore commerciale al quale il cliente si rivolge. È quindi necessario che l'operatore sia in grado di individuare se il cliente che effettua la domanda ricada o meno in una zona di rete cablata. La zona interessata dalla sperimentazione, inizialmente non coincidente con l'intera area di centrale, deve pertanto essere caratterizzata in TGU. Ciò è stato reso possibile classificando in TGU le zone di armadio interessate alla sperimentazione mediante l'inserimento della sigla "CAB" in un campo non utilizzato. In tale modo l'operatore, inserito l'indirizzo in procedura, è in grado di verificare l'appartenenza a rete cablata esaminando la pagina relativa all'area di armadio di competenza. L'aver utilizzato come entità minima discriminabile l'area di armadio ha comportato la necessità di un intervento totale all'interno dei relativi confini territoriali di competenza.

4.3 L' "indirizzo" nella sperimentazione

Una definizione di "indirizzo" di una sede è valido solo se tutti gli elementi che lo compongono sono definiti inequivocabilmente in modo da poter essere rilevati sia dal gestore che dal cliente. La prima definizione di indirizzo analizzata nella sperimentazione è stata "via, civico, complesso, scala, piano, interno" secondo l'impostazione consolidata. Nei centri storici delle nostre città non è però sempre possibile riscontrare tutti gli elementi di un tale "indirizzo". Le scale non sono sempre univocamente rubricate, spesso manca una numerazione degli interni, la presenza di ammezzati e sopraelevazioni rende ambigua la definizione dei piani. Altri enti erogatori di servizi per superare tale difficoltà numerano progressivamente i singoli accessi (esempio numero di posizione del contatore dell'energia elettrica).

Si è così giunti a definire l'"indirizzo" della sede del cliente come "indirizzo" della porta ad essa collegata.

Come metodologia di rilevazione di tale indirizzo, si è preferito utilizzare le caselle vocali. Il marcaggio del chiamante non è stato utilizzato perché non trasparente alle apparecchiature in uso presso lo sportello telefonico. Di contro l'utilizzo delle caselle

vocali ha permesso una maggiore flessibilità di informazione ed ha reso possibile aggiungere al codice identificativo una cifra di informazione aggiuntiva, ad esempio, per riconoscere facilmente la richiesta di attivazione di una sede (e non un cliente) precedentemente cessata per morosità, riducendo in tal modo le possibilità di frode. Un ulteriore vantaggio nella scelta della casella vocale è consistito nella possibilità, rilevato il codice identificativo, di effettuare la domanda di nuovo impianto da qualsiasi sede. Ciò può essere utile, ad esempio, ai clienti che trovino scomodo recarsi presso la sede da attivare negli orari di servizio dello sportello telefonico.

4.4 L'intervento in rete di distribuzione

Nella zona interessata alla sperimentazione è stato necessario attuare un intervento impiantistico di irrigidimento della rete di distribuzione e prolungamento delle coppie fino alle unità locative da servire. Il progettista, quindi, ha dovuto ripercorrere la prassi di rilevamento del fabbisogno coppie, proiettato nel breve periodo, secondo le nuove norme regolate dalla banca dati della rete cablata. In particolare si è affrontato il maggior onere di mantenere memoria della esatta assegnazione di ogni risorsa.

La differenza tra quanto rilevato e la consistenza costituisce l'insieme dei collegamenti da effettuare per portare a compimento l'intervento impiantistico di rete precollegata fonicamente relativamente alla rete di distribuzione.

Per un corretto utilizzo della procedura ACL/BDRR, che gestisce l'assegnazione delle coppie di rete, è stato necessario creare un numero adeguato di domande fittizie da associare alle coppie precollegate in modo da renderle "prenotate", cioè non assegnabili ad altri. Le stesse, all'atto della attivazione, saranno trasformate in "occupate", sostituendo la domanda fittizia con il reale numero assegnato. Il codice della domanda fittizia è funzione dell'"indirizzo" per permettere una rapida ricerca della risorsa assegnata.

4.5 L'intervento in commutazione

Per rendere possibile la gestione della numerazione libera, è stato necessario rendere accessibile la lettura del permutatore software al personale addetto alla manutenzione delle coppie di rete. In tal modo tale personale può ricostruire la relazione tra i numeri e la loro posizione al permutatore di rete, informazione indispensabile per la manutenzione.

L'accesso al permutatore software è stato garantito mediante un personal computer murale dotato di adeguato software. Nella memoria del computer è

registrata l'associazione porta-posizione al permutatore fisico, non soggetta a variazione. Il personal computer è collegato alla banca dati dell'autocommutatore dalla quale è capace di rilevare l'associazione numero-porta. Il programma, attraverso la combinazione delle associazioni numero-porta e porta-posizione al permutatore, è in grado di associare ad ogni numero la sua posizione fisica. Quest'ultima, allorché l'operatore richieda l'individuazione di un numero telefonico, viene restituita in forma grafica, mediante disegno schematico del permutatore con lampeggio della posizione corrispondente al numero inserito.

Inoltre è stato necessario installare un numero adeguato di caselle foniche connesse in hot line con le corrispondenti porte di centrale.

4.6 Analisi dei costi benefici

La valutazione economica può offrire dei riscontri positivi circa l'obiettivo che la rete precollegata fonicamente si prefigge: offrire al cliente un servizio in real time attraverso la minimizzazione delle operazioni necessarie all'attivazione.

Sulla scorta della valutazione economica dell'investimento è possibile individuare indici di territorio che evidenzino la minore o maggiore convenienza alla estensione della rete precollegata fonicamente.

Per il calcolo dei ritorni di investimento occorre innanzitutto definire la struttura dei costi:

- C_t : costi della indagine territoriale atta al censimento dell'area;
- C_a : costi di adeguamento della procedura gestionale commerciale;
- C_r : costo della costruzione della rete rigida;
- C_f : costi delle caselle foniche;
- C_l : costi dell'hardware per la numerazione libera.

Accanto a tali costi si evidenziano i benefici:

- B_r : tempi di allacciamento dell'ordine dei minuti;
- B_o : percezione di migliore qualità offerta al cliente;
- B_c : minori costi per l'attivazione;
- B_s : mancato onere della rimozione delle permutate per la cessazione dell'impianto.

In funzione di tale struttura di costo e benefici si definisce classicamente il "valore attuale netto", VAN, come:

$$VAN = \frac{\sum_k C_k - \sum_j B_j}{(1+i)^n},$$

dove i rappresenta il tasso di attualizzazione ed n il numero degli anni.

Il VAN, così definito, è un indicatore che, in caso di positività di segno, rappresenta la convenienza dell'investimento.

La positività dell'indicatore per la sperimentazione effettuata in Bari è essenzialmente legata ai seguenti due indici interni territoriali:

- Disponibilità di risorse, reti e centrali, ottimizzate sul territorio.
- Buona domanda interna con un battente che consente di sfruttare le risorse di rete.

Più in generale si prospetta una buona redditività nelle aree di centrale con un gradiente di domanda accentuato, tale da ammortizzare convenientemente le risorse impegnate.

Di natura non trascurabile appaiono i benefici dovuti ad una diversa organizzazione dell'attività gestionale della struttura commerciale ed ad una diversa struttura dei costi impiantistici orientata alla costituzione delle infrastrutture di rete.

4.7 I primi risultati della sperimentazione

La prima area di sperimentazione, approntata nel '92, è stata costituita dall'area cavo "Principe Amedeo" della centrale Centro. Tale zona ha costituito un primo campione di rete precollegata fonicamente, anche se non partecipata alla clientela, ed è stata utilizzata per verificare in campo la correttezza delle scelte operate e per permettere un primo addestramento degli operatori. Superata questa fase di avvio si è proceduto ad un ampliamento della zona territoriale della sperimentazione, che consentisse una valutazione degli aspetti legati all'utilizzo da parte dei clienti ed ai ritorni aziendali.

A partire dal mese di aprile '93 si è proceduto secondo le modalità di rete precollegata fonicamente relativamente a tutte le cessazioni dell'intera area di centrale. Contemporaneamente si è provveduto a realizzare l'intervento impiantistico di rete di distribuzione nelle altre aree cavo della centrale Centro. A maggio un quarto dell'area di centrale risulta coperto e tale percentuale è stata ampliata al 50% nel mese di giugno per giungere alla copertura dell'intera area di centrale alla fine di settembre.

Sulla parte sottoposta ad intervento impiantistico l'operatore commerciale riscontra in TGU la rete precollegata fonicamente e pertanto invita il cliente a "leggere" l'identificativo inserendo l'apparecchio telefonico nella presa della sua sede. Anche sul resto della centrale esistono impianti sparsi di rete precollegata, che possono essere comunque utilizzati dal cliente che proceda secondo le modalità di rete precollegata fonicamente anche senza esplicito invito dell'operatore commerciale. Tutti gli impianti di rete precollegata fonicamente sono comunque individuati e quindi utilizzati da parte del personale del centro lavori impianti di abbonato che li riconosce mediante la banca dati ACL/BDRR.

5. Conclusioni

La sperimentazione avviata sul territorio di Bari si propone di perseguire i seguenti obiettivi:

- Tempi di accesso al servizio di fonia praticamente nulli.
- Migliore immagine verso il cliente.
- Migliore efficienza di gestione da ottenere sviluppando le integrazioni tra le diverse procedure.
- Verifica in campo delle interoperatività e della integrazione per consentire una valutazione delle scelte effettuate, contribuendo all'evoluzione dei sistemi di gestione con l'esperienza maturata.
- Redditività degli investimenti con più pronta velocità di ritorno.

L'integrazione delle tecniche, già sviluppate per diverse esigenze ed esperienze, rappresenta l'originalità della sperimentazione in atto e dimostra che i singoli benefici ottenibili da ciascuna delle applicazioni portano nell'integrazione ulteriori e significativi sviluppi e benefici.

Il risultato di tale integrazione è una rete di telecomunicazioni, la precollegata fonicamente, dalle prestazioni particolarmente competitive, che si distingue dalle altre soluzioni in particolare per le due seguenti peculiarità:

- La rete precollegata fonicamente utilizza come banca dati la rete stessa, interrogata direttamente dal cliente senza alcun intermediario. Ciò comporta una totale assenza di errori dovuti al mancato aggiornamento della banca dati o ad errori di accesso alla stessa.
- La rete precollegata fonicamente prevede che ogni operazione hardware sulla rete sia sostituita da una transazione software. Il tempo di realizzazione del collegamento è quindi costituito dalla somma dei tempi di alcune transazioni e dei tempi di colloquio tra gli operatori. È quindi possibile giungere a tempi di attivazione dell'ordine di alcuni minuti.

In particolare, grazie al fatto di aver eliminato ogni operazione hardware nella costituzione del collegamento, le transazioni possono essere remotizzate e comandate da un unico operatore che, se abilitato a svolgere anche le attività commerciali in diretto colloquio con il cliente, può eseguire l'attivazione in linea.

Inoltre il fatto di operare esclusivamente con transazioni software rende la realizzazione della rete precollegata fonicamente indipendente dalla soluzione hardware di rete adottata (coppie in rame, Fiber To The Office, Fiber To The Curb, Fiber To The Home, etc.).

Centri di esercizio e manutenzione delle centrali numeriche

F. Segre, A. Trifiletti (*)

In questo articolo vengono descritte le principali caratteristiche e funzioni del CEM (Centro di Esercizio e Manutenzione delle centrali numeriche).

Dopo una prima parte descrittiva generale, sono state analizzate più nel dettaglio la configurazione del CEM della Liguria e le attività svolte dal personale del Nucleo Operativo di Supervisione Regionale (NOSR), evidenziando i principali vantaggi ottenuti nell'esercizio delle centrali numeriche con l'utilizzo di detto sistema.

1. I centri di controllo e manutenzione della rete

L'evoluzione delle strutture dei sistemi di telecomunicazione impone la necessità di operare su impianti di complessità crescente nel tempo, aventi dimensioni che abbracciano porzioni di rete sempre maggiori.

Nei riguardi delle funzioni di controllo e manutenzione, questo profondo cambiamento tecnologico e strutturale dà luogo a due importanti conseguenze:

- le conoscenze di dettaglio sulla costituzione e sul funzionamento degli apparati non possono più essere diffuse come una volta, a livello dei tecnici che operano in periferia;
- le funzioni diagnostiche devono analizzare e correlare contemporaneamente una molteplicità di dati provenienti in tempo reale da impianti e sedi diverse.

A queste nuove esigenze si aggiunge poi la necessità di operare da pochi punti di comando verso più sedi decentrate per le funzioni di configurazione e di gestione della rete.

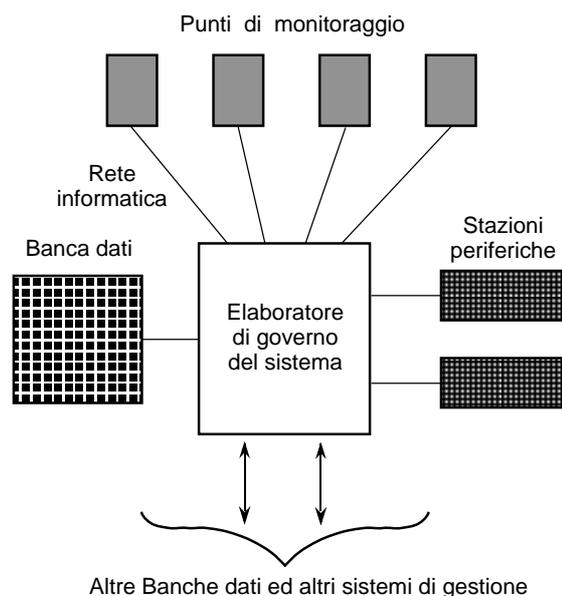
Questi sono i principali motivi che hanno indotto ad organizzare le funzioni manutentive e gestionali in forma centralizzata, mediante la costituzione di pochi nodi di esercizio e manutenzione posizionati al centro di reti stellari di trasporto dei dati.

I principi generali ispiratori delle architetture di controllo sono fondamentalmente due:

- massima centralizzazione dei punti di raccolta ed elaborazione dei dati e delle funzioni decisionali di rilievo sul funzionamento della rete;
- massima diffusione verso la periferia delle informazioni raccolte e dei risultati delle analisi svolte nei centri suddetti.

Tenendo conto di questi criteri generali, un sistema di controllo per l'esercizio e la manutenzione è, nella sua forma più semplice, costituito dalle seguenti parti funzionali (vedi fig. 1):

- punti di monitoraggio, o di esecuzione comandi, posti in prossimità o all'interno degli oggetti da controllare;
- uno o più elaboratori interconnessi per il governo dell'intero sistema e per l'elaborazione dei dati, posto in corrispondenza del centro di supervisione;
- una rete informatica, a struttura stellare, per la trasmissione dei dati di controllo;
- una o più banche dati centralizzate;
- un numero limitato di stazioni periferiche;



(*) ing. Franco Segre, ing. Antonino Trifiletti - SIP DR Liguria - Genova

Figura 1 Parti funzionali di un sistema di controllo della manutenzione

- interconnessioni con altre banche dati e altri sistemi di gestione, di visione o di stampa dei dati.

Questa struttura di base può poi essere arricchita dalla presenza di stazioni di pre-elaborazione, atte a ridurre la quantità di informazione che deve essere direttamente trattata dal centro.

2. Funzioni dei CEM

Focalizzando l'attenzione sulle centrali numeriche di commutazione della rete italiana, l'organizzazione sopra descritta fa capo ai CEM (Centri di Esercizio e Manutenzione) situati presso i capoluoghi di Regione nei Centri di Supervisione Regionali (CSR).

Nella rete SIP sono stati introdotti CEM di tipo differente per le quattro marche di autocommutatori esistenti (ITALTEL - FACE ALCATEL - FATME - SIEMENS).

Essi svolgono una serie di funzioni, che possono essere così classificate:

- funzioni di esercizio
- funzioni di manutenzione
- funzioni amministrative
- funzioni di supporto
- funzioni di interfacciamento con altri sistemi

Le funzioni di esercizio si prefiggono principalmente di consentire agli operatori posti nel CSR di presiedere alle configurazioni degli autocommutatori (fasce - instradamenti - tipi di segnalazione), di caricare il software delle centrali e controllarne l'efficienza, di gestire il processo delle misure di traffico e di qualità e di raccoglierne ed elaborare i risultati.

Le funzioni di manutenzione consentono la raccolta degli allarmi, provvedono alle relative segnalazioni agli operatori, gestiscono le schede delle irregolarità ed alimentano la banca dei relativi dati statistici.

Le funzioni amministrative gestiscono le categorie d'utente e consentono la raccolta dei dati dei contatori elettronici e della documentazione degli addebiti.

Le funzioni di supporto gestiscono il dialogo uomo-macchina e l'archivio delle operazioni svolte.

L'interfacciamento con altri sistemi deve essere assicurato soprattutto nei riguardi di quei sistemi esterni che debbono elaborare i dati rilevati nelle centrali e raccolti presso i CEM: in particolare i dati di tassazione (da trasferire al CED) e di misure di traffico (da trasmettere ai sistemi specifici di trattamento di questi dati).

3. Struttura e organizzazione di un CEM

Per esaminare concretamente la struttura e l'attività di un CEM, viene preso a riferimento quello della LIGURIA per la gestione delle centrali ITALTEL.

3.1 Interconnessione del CEM con il mondo esterno.

Nella fig. 2 sono rappresentate le linee di interconnessione con gli oggetti controllati, indicati con simboli ovali; le altre linee collegano il CEM con i terminali di supervisione e con gli altri sistemi di elaborazione (rappresentati con rettangoli).

Per il CEM di Genova gli oggetti controllati sono i seguenti:

- 2 Stadi di Gruppo di Transito (SGT) di tipo UT 100 (GENOVA Lagaccio e SAVONA Letimbro);
- 1 Stadio di Gruppo di Transito (SGT) provvisorio di tipo TN 16 (SAVONA Letimbro);
- 19 Stadi di Gruppo Urbani (SGU) di tipo UT 100;
- 2 Elaboratori di Supporto di Centrali CT2/T1 (2 centrali);
- 6 Elaboratori di Supporto di Centrali CT2/T2 (9 centrali);
- 3 Elaboratori di Supporto (ES) di centrali UT 20 (4 centrali).

Gli altri sistemi di elaborazione verso i quali il CEM deve trasferire dati sono:

- il CSN (Centro di Supervisione Nazionale) per i dati di intensità di traffico e di congestione dei fasci di ultima scelta;
- l'UCRD (Unità di Concentrazione Raccolta Dati sul traffico) per i dati di misura del traffico e dell'accessibilità;
- l'SSTR (Sistema di Supporto Tecnico Regionale) per i dati di tassazione e di documentazione degli addebiti;
- il CED (Centro di Elaborazione Dati) per l'attuazione della procedura relativa agli utenti morosi.

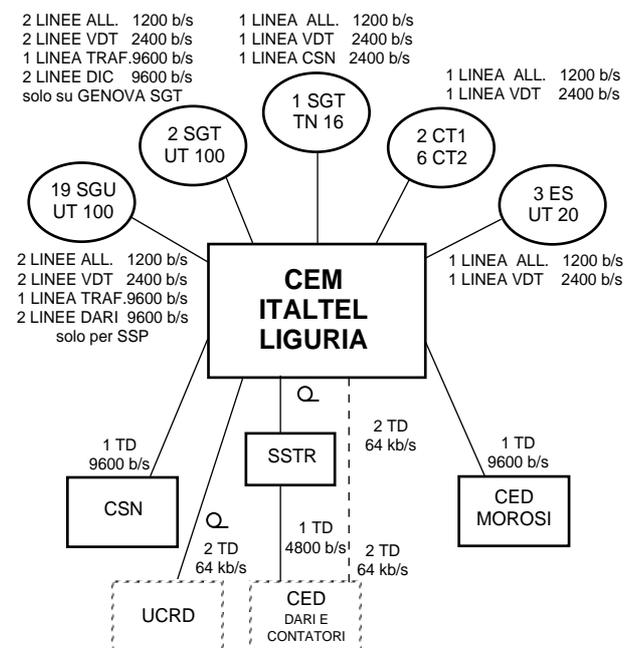


Figura 2 Interconnessioni tra il CEM della Liguria ed il mondo esterno

3.2 Struttura interna del CEM

Nella fig. 3 è rappresentato il sistema di elaborazione centrale.

L'attuale configurazione utilizza un elaboratore VAX 6530 della Digital, che tramite HSC "Hierarchical Storage Controller" (dispositivo di controllo della memoria di massa dischi-nastri) gestisce 20 dischi e tre unità nastri.

Tutte le scritture su disco vengono eseguite in doppio, in modo da garantire la salvaguardia dei dati nei casi di una singola rottura.

Guardando il disegno notiamo che tramite LAN (Local Area Network) l'elaboratore dialoga con il mondo esterno, utilizzando opportune interfacce (DELNI, DECSERVER, DEMPR, DEMSA).

L'utilizzo dell'interfaccia DELNI permette il collegamento con molte periferiche con un'unica connessione sulla LAN: una DELNI ha 8 porte su cui si possono collegare 8 DECSERVER, da 8 porte cadauno, per un totale di $8 \times 8 = 64$ porte utili.

Sui DECSERVER vengono configurati terminali locali o remotati, e tutte le linee verso le centrali elettroniche.

Le porte dei DECSERVER possono essere configurate in modo indipendente a velocità diverse a seconda dell'utilizzo.

Il DEMPR permette, come i DELNI, un'espansione da 1 a 8, ed è connesso ad un apparato modulare che gestisce il collegamento con le periferiche per il trasferimento dei dati di Traffico Globale relativi alle prestazioni di documentazione degli addebiti.

La Work-Station è un PC 386 con video grafico, sul quale vengono visualizzati tutti gli allarmi di UT e del CEM.

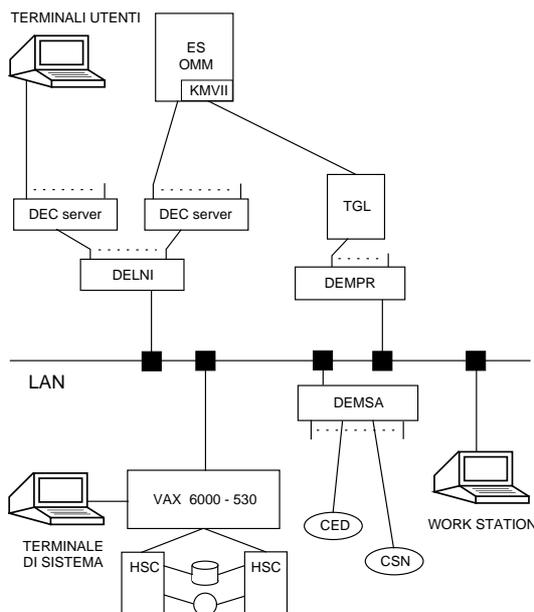


Figura 3 Sistema di elaborazione centrale del CEM

Il lavoro di acquisizione allarmi da parte dell'operatore viene notevolmente facilitato con un sistema grafico a finestre e colori differenti.

3.3 Organizzazione dei centri operativi

I centri operativi controllati dal CEM sono di tre tipi:

- Centri di singole Aree Impianti di Centrale (AIC);
- Centri di Supervisione e Esercizio (CSE) di Unità Territoriali, comprendenti più AIC;
- Centro di Supervisione di Regione (CSR-NOSR).

A ciascuno di questi centri il sistema attribuisce funzioni di gestione diversificate in relazione alle portate territoriali di competenza. I corrispondenti posti d'operatore hanno equipaggiamenti differenti, adeguati alle specifiche esigenze di controllo: gli AIC e i CSE sono dotati solo di terminale d'operatore per l'interattività con il sistema e il prelievo delle statistiche, mentre presso il CSR sono presenti anche due work-stations per la visualizzazione degli allarmi della regione e una consolle per la gestione del sistema.

Gli operatori sono suddivisi in "classi d'utenza" a seconda delle operazioni che devono svolgere e alle quali sono abilitati:

- il *gestore* del sistema può agire su tutte le funzioni; in particolare è l'unico che può modificare le funzionalità del CEM e la gestione degli altri operatori;
- il *supervisore* gestisce la configurazione degli AIC ed il controllo delle procedure automatiche;
- il *tecnico* svolge le funzioni di trattamento e risoluzione degli allarmi di competenza;
- il *system manager* è l'unico operatore abilitato ad intervenire sul Sistema Operativo.

4. Prestazioni del CEM

Passiamo rapidamente in rassegna le principali prestazioni espletate dal sistema :

a) Gestione dei dati di configurazione degli autocommutatori:

Le regole di instradamento e di tassazione del traffico, per le diverse centrali governate da un CEM, sono memorizzate in un data-base centrale e nelle banche dati dei vari autocommutatori. E' previsto che l'allineamento dell'archivio centrale con quelli periferici venga effettuato in via automatica dal CEM, in funzione delle informazioni ricevute dalle centrali. Attualmente questa prestazione non è ancora presente.

b) Supervisione e gestione del traffico:

Si tratta di una funzione accentrata presso il Centro di Supervisione Nazionale (CSN). Il CEM ogni 15 minuti riceve i dati di intensità di traffico e di perdita dei fasci di

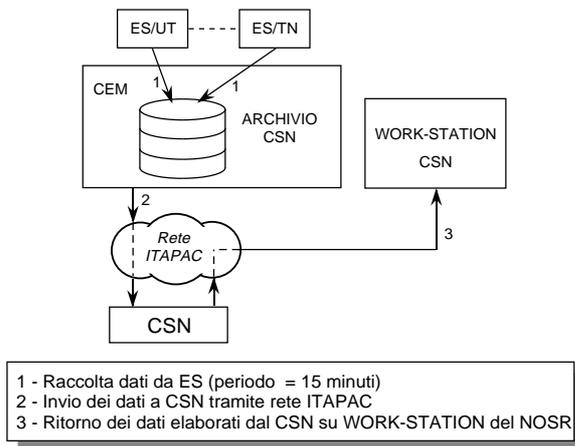


Figura 4 Struttura della supervisione del traffico

ultima scelta da tutti gli autocommutatori controllati e li invia al CSN tramite rete Itapac (vedi fig. 4). Nella stessa sede regionale è situata una work-station del CSN per la visualizzazione dei dati e delle situazioni critiche.

c) *Gestione misure:*

Le misure di traffico e di accessibilità eseguite dagli autocommutatori vengono organizzate ed attivate dal CEM, che ne determina le caratteristiche ed il calendario, e ne raccoglie i risultati, normalmente su base giornaliera, per memorizzarli su nastro e per trasferirli successivamente al sistema di elaborazione specifica di tali dati (UCRD o Elaboratore di Misure di Traffico Regionale, EMTR) (vedi fig. 5).

Le misure di disponibilità e di occupazione dei fasci, che consentono di determinare i valori di intensità di traffico, vengono predisposte per durate di 14 giorni, con una raccolta di dati presso il CEM ogni 5 giorni. Le misure di accessibilità, effettuate per campionatura, hanno una durata di almeno 7 giorni, con raccolta dei dati ogni due giorni. Le misure del traffico su singoli utenti sono permanenti, ed i dati vengono trasferiti al CEM ogni 10 giorni. Oggi si eseguono anche misure sui servizi ausiliari (STS), con cicli di durata di due settimane.

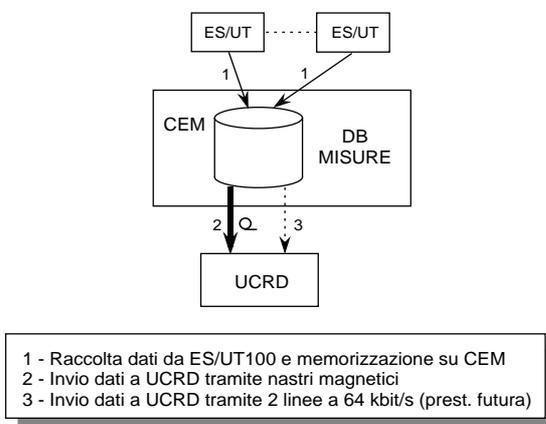


Figura 5 Struttura per la gestione delle misure di traffico

d) *Telecaricamento software:*

Le nuove versioni del software delle centrali (cambi release) e le relative variazioni (patch), che vengono attualmente caricate tramite comandi periferici, saranno inviate direttamente dal CEM (questa prestazione non è ancora in esercizio).

e) *Gestione del sincronismo data-ora:*

E' previsto che il CEM sincronizzi gli autocommutatori di sua competenza, allineandoli una volta al giorno alla data e all'ora correnti.

f) *Supervisione e gestione allarmi:*

Gli allarmi che il CEM riceve dagli autocommutatori possono essere classificati in funzione sia della tipologia degli oggetti controllati (allarmi di utenti e di giunzione), sia delle azioni da compiere in conseguenza delle segnalazioni (verifiche preliminari, invio immediato di un tecnico in loco).

Inoltre, per qualche tipo di reclamo di utente, il CEM viene interessato per attivare i test necessari a discriminare se il guasto è da attribuire all'autocommutatore o alla linea di abbonato.

g) *Gestione statistiche di manutenzione:*

Le procedure previste consentono di raccogliere, di classificare e di sintetizzare i dati significativi sulle operazioni di manutenzione e sulle schede di irregolarità.

h) *Gestione degli utenti morosi:*

Dal CEM è possibile disabilitare e reinserire in servizio gli utenti morosi, in base alle informazioni ricevute dal Centro di Elaborazione Dati, con il quale è collegato tramite una linea dedicata. La procedura si attiva automaticamente due volte al giorno, e può essere inoltre avviata manualmente dall'operatore. Le centrali, a lavoro ultimato, inviano al CEM l'esito dell'operazione. Le fasi del processo sono rappresentate schematicamente in fig. 6. Il sistema consente pure di effettuare controlli

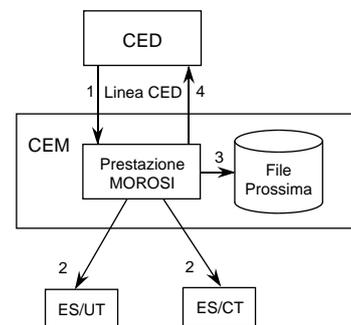


Figura 6 Struttura della prestazione morosi

di allineamento e di congruenza tra i dati di memoria del CED e quelli reali delle centrali. Un pregio di questa procedura è la rapidità di intervento nella fase di riattivazione del servizio.

i) *Gestione dei contatori elettronici:*

Il CEM, con un processo automatico, richiede giornalmente alle centrali governate di trasmettere il contenuto dei contatori d'utente. Il numero acquisito di scatti viene memorizzato in doppio sul data-base locale, e trasferito periodicamente al CED su linea dedicata, o tramite nastro (fig. 7). Tutte queste operazioni si svolgono normalmente secondo precisi orari, ma possono essere anche attivate manualmente da operatori abilitati. Si attuano inoltre dopo un'indisponibilità prolungata dei collegamenti tra centrali e CEM.

l) *Documentazione degli addebiti:*

Il CEM riconosce i numeri d'utente appartenenti alla categoria che richiede la documentazione degli addebiti, e, con procedure automatiche o manuali, registra i dati nell'archivio locale e li trasferisce al CED (vedi fig. 7). Il CEM ha una capacità di memoria sufficientemente grande per contenere tutti i dati di conteggio e di documentazione prodotti negli ultimi cinque giorni dalle centrali controllate.

5. **Attività giornaliera di un CEM**

Prendendo sempre a riferimento il CEM di Genova, esaminiamo ora l'attività giornaliera svolta dai 9 tecnici, che si alternano in tre turni assicurando un presidio permanente del Centro.

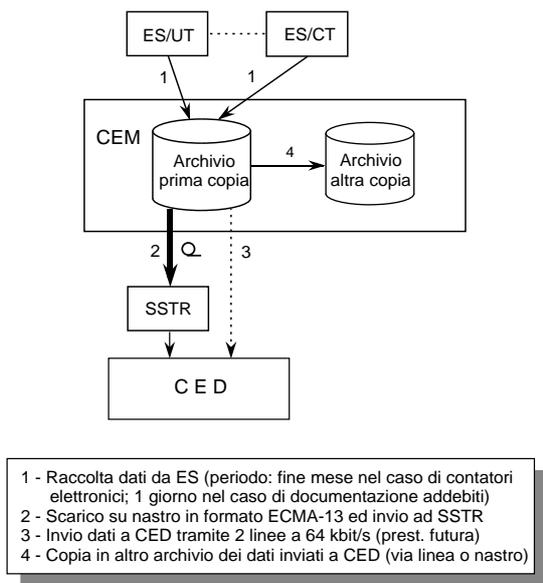


Figura 7 Struttura per la gestione dei contatori elettronici e documentazione degli addebiti

Durante l'orario diurno di base (dalle 8 alle 17), una persona si dedica completamente alla gestione delle misure di traffico, eseguendo operazioni di abilitazione ed interruzione delle misure e individuando eventuali tempi di interruzione, al fine di rispettare la durata prestabilita, mantenendo i contatti con i tecnici che lavorano presso le sedi delle Aree di Impianti di Centrale (AIC).

Il restante personale si dedica in genere alle altre funzioni.

La giornata tipica della vita del CEM è organizzata sulla base di una sequenza di attività sistematiche che vengono eseguite con cadenze periodiche. A queste si sovrappongono le operazioni provocate dall'arrivo di specifici allarmi.

Esaminiamo le principali attività periodiche:

- Uno dei primi interventi mattutini consiste nell'analisi, nella trascrizione e nella contabilizzazione degli allarmi pervenuti durante la notte.
- Nel frattempo l'operatore addetto alle misure predispone il programma giornaliero, in sintonia con le AIC e successivamente provvede all'invio dei dati raccolti all'apposito sistema di elaborazione esterno. Nei casi in cui non è ancora possibile un interfacciamento diretto dei due sistemi, l'operatore dispone per una registrazione dei dati su nastro, da inviare manualmente a destinazione.
- Verso le 9,30 si opera per la registrazione su nastro delle comunicazioni eseguite dagli utenti che hanno richiesto la documentazione degli addebiti.
- Alle 10,30 viene attivata la raccolta dei dati di traffico effettuati dagli utenti collegati alla Rete Intelligente.
- Si effettua pure un controllo sistematico dell'efficienza di tutte le linee di Trasmissione Dati con il CED.
- Nel primo pomeriggio vengono raccolte dalle centrali le informazioni relative alle numerazioni assegnate ai nuovi ampliamenti, al fine di allineare le banche dati del CEM all'effettiva consistenza degli autocommutatori.
- Alle 15 ed alle ore 3 del mattino, sulla base di un file ricevuto dal CED, si impartiscono i comandi relativi all'attivazione dell'utenza morosa. Al termine si controlla l'esito dell'operazione, e, se è necessario, si invia un fac-simile al nucleo operativo interessato per la riattivazione manuale.
- Alle ore 16,40, al termine dell'orario di presidio delle centrali, il turnista del CEM assume le competenze della gestione diretta degli allarmi di tutti gli autocommutatori numerici della Liguria, e li gestisce tramite workstation, prodigandosi per la loro risoluzione.
- Nelle ore serali e durante la notte l'operatore esegue alcune operazioni di routine e controlla che tutti gli autocommutatori inviino regolarmente i loro dati.
- Alla fine del mese occorre attivare la procedura contatori d'utente, che coinvolge, oltre al turnista, una seconda persona. Tramite un'opportuna schedulazione, le centrali inviano al CEM le letture

dei contatori di propria pertinenza. Il CEM raccoglie questi dati, verifica che non ne manchino e li registra su nastro.

Si è detto sopra che, al di fuori dell'orario di presidio delle AIC, l'operatore del CEM controlla gli allarmi di tutti gli autocommutatori numerici, e, in presenza di un Grande Allarme, esegue una prima indagine da terminale, al fine di determinare se è necessario richiedere l'intervento del reperibile specialista.

A tal fine l'allarme viene assegnato ad una delle seguenti classi:

- allarmi che necessitano dell'invio immediato del reperibile di 1° livello (per esempio il "fuori servizio" totale di un Elaboratore di Supporto periferico);
- allarmi che necessitano di comandi di verifica, con successivo tentativo di ripristino telecomandato (per esempio in presenza di allarme di "fuori servizio" di un modulo di centrale, prima si determina il numero ed il tipo del modulo fermo, poi si lancia il comando di "restart", quindi si discrimina se fare o meno intervenire sul posto il reperibile);
- allarmi a soglia sui fasci di giunzione ed allarmi di destinazione sul canale comune di segnalazione: in

entrambi i casi l'operatore, individuato il fascio in allarme, tenta un recupero, e, in caso di esito negativo, ricorre all'intervento del reperibile solo se il fascio in allarme è di ultima scelta;

- allarmi a soglia su utenti bloccati: l'operatore, dopo aver tentato inutilmente il ricupero, ricorre o meno al reperibile, a seconda della classe di utente.

L'operatore, nel trattamento degli allarmi, è aiutato da una guida, contenente precise norme comportamentali per ogni tipo di evento rilevabile.

6. Efficacia del CEM

Nel periodo da aprile a dicembre 1992 sono pervenuti al CEM di Genova 428 allarmi, di cui 213 (pari al 49,8%) sono stati risolti dal personale del Centro, mentre 166 segnalazioni (pari al 38,7%) hanno richiesto l'intervento del reperibile (vedi fig. 8).

Nei primi cinque mesi del 1993 sono stati segnalati 235 allarmi, di cui ben 167 (pari al 71,1%) sono stati risolti direttamente, e solo per 59 (pari al 25,1%) è stato necessario il ricorso al reperibile (fig. 9).

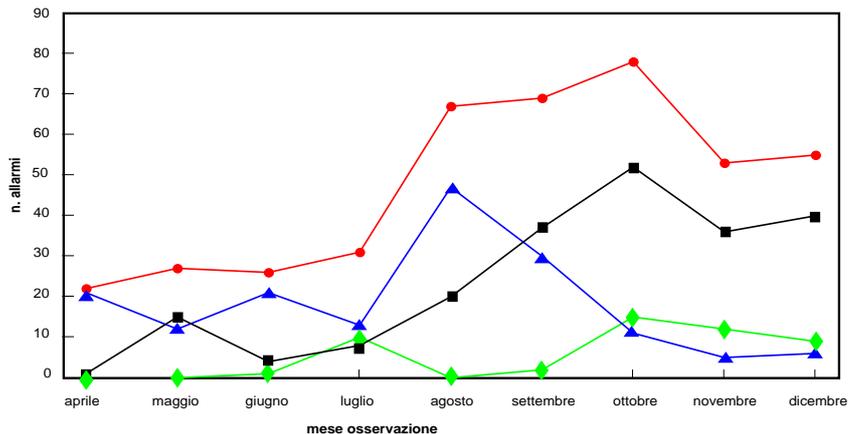
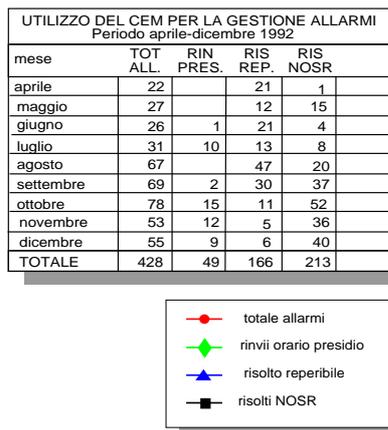


Figura 8 Utilizzo del CEM per la gestione allarmi (periodo aprile-dicembre 1992)

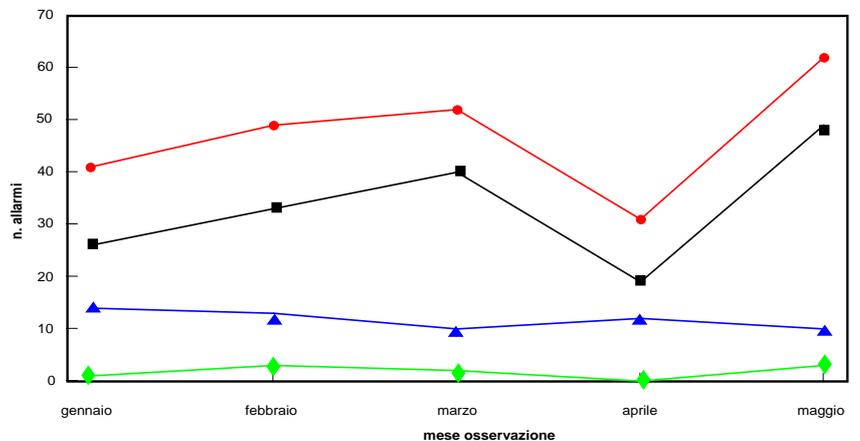
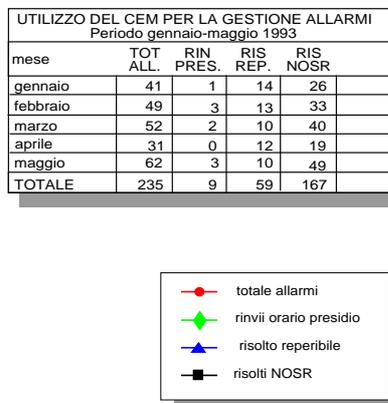


Figura 9 Utilizzo del CEM per la gestione allarmi (periodo gennaio-maggio 1993)

Questo netto miglioramento dell'efficacia delle azioni del CEM è dovuto da un lato alla progressiva esperienza acquisita dal personale, suffragata dalla partecipazione a corsi ripetitivi presso le case costruttrici e dall'addestramento effettuato in Direzione Regionale, dall'altro alla migliore puntualizzazione delle norme operative. Il beneficio è reso evidente dal confronto tra il tempo medio occorrente per la risoluzione diretta dell'inconveniente, pari a 20 minuti, e quello necessario al reperibile per l'intervento esterno, pari a un'ora e 34 minuti.

Gli allarmi pervenuti sono stati ulteriormente suddivisi per tipologia per vedere l'efficacia degli interventi tramite CEM:

- 1) fuori servizio di un doppio blocco di sicurezza di un modulo: creano in genere un grave disservizio per gli utenti ed i collegamenti afferenti al modulo stesso;
- 2) anomalia di un concentratore remoto, per mancanza rete, o per altro grave degrado;
- 3) allarmi per i quali non vi è ancora disservizio, ma esiste il rischio di un imminente fermo di un modulo, per inconvenienti alla sincronizzazione o all'unità di temporizzazione di modulo;
- 4) fuori servizio per un gruppo di utenza, oppure per uno o più fasci di giunzioni;
- 5) modulo fuori servizio;
- 6) fuori servizio di un numero di circuiti di giunzione superiore ad una soglia prefissata, oppure per un guasto all'Elaboratore di Supporto;
- 7) rischio di centrale ferma, per inconvenienti alle parti comuni tra più moduli, quali modulo distribuzione messaggi o generazione e distribuzione temporizzazioni.

Se gli allarmi riscontrati da gennaio a maggio 1993 vengono ripartiti in queste categorie, si ricavano i dati della tabella 1, che consentono di fare le seguenti considerazioni:

- gli allarmi che hanno creato un disservizio immediato totale (categorie 1, 2 e 5) sono stati 48 (20%), con alta percentuale di interventi di ripristino eseguiti direttamente dal CEM (75%);

Periodo 1/1/1993 - 31/5/1993						
NUM. ALL.	TOT. ALL.	RIS. NOSR		RIS. REP.		RINV. PRES.
1	14	10	71,43%	4	28,57%	
2	12	10	83,33%	2	16,67%	
3	17	11	64,71%	6	35,29%	
4	113	84	74,34%	24	21,24%	5 4,42%
5	22	16	72,73%	6	27,27%	
6	34	13	38,24%	17	50,00%	4 11,76%
7	23	23	100,00%			
TOT. AL.	235	167	71,06%	59	25,11%	9 3,83%

Tabella 1 Ripartizione degli allarmi per tipologia

- la categoria prevalente di allarmi è quella che ha posto fuori servizio gruppi di utenti o di giunzioni; anche per questi l'intervento degli operatori del CEM è stato risolutivo nella grande maggioranza dei casi;
- gli inconvenienti alle parti comuni tra più moduli (n. 7) sono sempre stati risolti dal CEM;
- la più bassa efficacia delle azioni del CEM si riscontra negli allarmi (n. 6), per i quali non è spesso possibile agire se non sul posto in cui si verifica il guasto.

I lusinghieri progressi ottenuti saranno oggetto di ulteriori miglioramenti, mediante l'affinamento delle procedure e delle conoscenze, ma già sin d'ora consentono di affermare che l'efficacia del CEM contribuisce decisamente a razionalizzare al meglio l'impiego del personale di manutenzione, a ridurre la durata dei disservizi, e quindi ad innalzare il livello della qualità percepita dai clienti.

7. Prestazioni future dei CEM

Tutte le funzioni descritte saranno soggette a future evoluzioni tendenti a migliorarne le prestazioni, e ad estenderne le applicazioni a categorie particolari di utenti, di servizi e di tipi di traffico. Si cureranno in particolare:

- gli aspetti di presentazione/visualizzazione dei dati;
- l'acquisizione e l'organizzazione dei dati secondo nuovi formati;
- l'estensione agli utenti di servizi ISDN e CLASS;
- la caratterizzazione degli utenti START;
- la gestione dei contatori per i servizi di decade 1.

Inoltre verrà introdotta una nuova prestazione "GRAFO DI RETE" avente lo scopo di fornire un'immediata visualizzazione grafica degli instradamenti (SGU-SGT e SGU-SGU), evidenziando le caratteristiche dei fasci interessati (nome, consistenza, scelte e ultimi dati di misura di traffico). A tal fine, il CEM viene collegato con un elaboratore, che ha il compito di raccogliere i dati degli instradamenti delle centrali.

Con queste ed altre operatività, sarà possibile avere, nell'unica sede regionale, non solo un controllo integrato dell'efficienza degli impianti, ma anche una visione unitaria e completa dell'andamento del traffico e della qualità.

Il passo successivo sarà quello che consentirà, in caso di bisogno, di intervenire attivamente sugli autocommutatori, per far fronte a condizioni particolari di sovraccarico o di guasti esterni che richiedano di modificare instradamenti, creare priorità, parzializzare fasci di giunzione o alterarne momentaneamente le funzioni.

Queste attività devono necessariamente essere gestite in tempo reale, rispetto agli eventi che le determinano, da pochi punti di regia, che abbiano una visione la più ampia possibile su vasti strati della rete.