1 Generalità sulle comunicazioni

Una **comunicazione** è uno scambio di informazioni fra due sistemi.
Per stabilire una comunicazione fra sistemi diversi sono necessarie almeno quattro cose:
1. almeno due soggetti della comunicazione, che si devono scambiare informazioni
2. le informazioni da comunicare
3. uno o più mezzi fisici attraverso i quali le informazioni possano essere scambiate
4. un insieme di regole che entrambi i soggetti devono seguire per realizzare una comunicazione ordinata.

I soggetti della comunicazione
In ogni **collegamento** ("**link**") fra almeno due soggetti si può distinguere un soggetto che spedisce le informazioni, detto **sorgente**, **trasmettitore** o **emittente** ed almeno un altro che le riceve, detto **destinatario** o **ricevitore**. Questi soggetti vengono anche chiamati "stazioni", "nodi", "punti", senza specificare la direzione della comunicazione. In base a al contesto cui appartengono, vengono usati anche molti altri nomi, alcuni dei quali vedremo in seguito

Le informazioni
L'insieme delle informazioni scambiate in una direzione dai due soggetti durante una comunicazione viene detto "**segnale**". Chiameremo "**messaggio**" un insieme di informazioni che hanno caratteristiche comuni e che viene scambiato in modo unitario in una sola direzione.

Il canale di comunicazione
Un **mezzo di trasmissione** (transmission medium) è un supporto fisico sul quale può avvenire uno scambio di informazioni. Un mezzo di trasmissione è un singolo sistema fisico attraverso il quale può essere effettuata la propagazione di informazione.
Un **canale di comunicazione** è l'insieme dei mezzi, fisici e/o astratti, utilizzati per ottenere il collegamento fra un sorgente ed un destinatario.
Un canale può essere costituito da una successione di diversi mezzi di trasmissione, di dispositivi e di software che insieme realizzano l'intero percorso dal sorgente al destinatario.

Le regole
Perché una comunicazione sia efficace è necessario che entrambi i soggetti usino lo stesso insieme di regole per mettersi d'accordo su come rappresentare le informazioni, come usare il mezzo di trasmissione e in generale su tutti i dettagli necessari perché lo scambio funzioni.
Nel campo delle telecomunicazioni, ed in particolare quando si tratta di software, a queste regole si dà il nome di "**protocolli**".

Le norme
Spesso i protocolli di comunicazione sono stabiliti da "**normative**". Le normative (o "**norme"**) sono specifiche funzionali che sistemi prodotti dall'uomo devono rispettare. Se due diverse aziende producono il loro sistemi in modo che siano entrambi conformi con una norma, essi saranno in grado di lavorare insieme ossia, come si dice in gergo tecnico, di "**interoperare**".
Le norme vengono emesse da organismi di normalizzazione internazionali o nazionali[1], che esprimono il consenso del mondo della ricerca e dell'industria. Nel gergo tecnico le norme sono spesso indicate con il loro nome Inglese di "**standard**".
La realizzazione di sistemi conformi con le norme di solito è volontaria ("norme volontarie") e viene intrapresa dalle aziende per beneficiare dei vantaggi dell'interoperabilità.
Per questioni che riguardano la sicurezza (safety) delle persone, alcune norme possono essere rese obbligatorie dalla legge[2]. In questo caso si parla di "norme cogenti".

Una definizione generale di rete

Una **rete** è un gruppo di entità che possono scambiarsi informazioni, energia o materia.
Questa è una definizione molto generale; seguendola possiamo considerare reti il sistema postale dell'antica Roma, una classe di una scuola, il sistema di distribuzione dell'energia elettrica od il sistema di trasporto merci su gomma.
In una rete di comunicazione avvengono una o più **comunicazioni**, cioè scambi di informazioni.
Anche restringendo il campo alle sole reti di comunicazioni elettriche, esistono molti tipi di reti, con caratteristiche elettriche, protocolli e prestazioni molto diverse.
Diamo ora una serie di definizioni generali che saranno la base per le successive spiegazioni più dettagliate. Le definizioni fanno riferimento alle comunicazioni elettriche, anche se in alcuni casi il loro significato si è esteso alle comunicazioni in senso più generale, come per esempio quelle che si possono stabilire fra esseri umani.

1 p.es. UNI e CEI realizzano le norme italiane, la CEI realizza le norme "elettriche" (e anche informatiche), l'UNI le altre. In campo europeo esistono CEN-CENELEC (CENELEC per le norme elettriche), in campo mondiale ISO-IEC (IEC per le elettriche). Altri organismi importanti sono ITU (ex CCITT), IEEE, EIA, ed altri. Inoltre esistono molti comitati costituiti dalle imprese che stendono specifiche e norme, p.es. Vesa, USB, PCI ed altri.
2 p.es. la legge 626 per la sicurezza sul lavoro contiene riferimenti a norme UNI, che quindi devono essere obbligatoriamente rispettate, per legge.

Direzione del flusso dei collegamenti

In base alla direzione ed al momento in cui i dati fluiscono, un collegamento può essere classificato in diversi modi.

Simplex

Nei collegamenti **simplex** la trasmissione può avvenire sempre e solo dal trasmettitore al ricevitore, senza che essi si possano scambiare i ruoli. Per esempio la radio, se considerata dal punto di vista di un solo utente, è una comunicazione simplex, che funziona sempre in un solo senso, senza possibilità di essere invertita.

Half duplex

Nella comunicazione **half duplex** in ogni istante la trasmissione può avvenire in un solo senso, ma il sorgente ed il destinatario possono scambiarsi i ruoli. Questo tipo di collegamento richiede che i due soggetti coinvolti usino uno stesso insieme di regole (uno stesso "protocollo") per stabilire come "passarsi la parola".
In una comunicazione half duplex il mezzo di trasmissione viene sfruttato completamente per una sola direzione alla volta, per cui si può comunicare alla massima velocità possibile per quel mezzo.
Per esempio un fax è un dispositivo half duplex. Infatti dopo la fase di collegamento iniziale il fax che ha iniziato la chiamata spedisce i dati relativi all'immagine da trasmettere, mentre l'altro sta solo in ascolto. Poi il trasmittente comunica che ha terminato ed il ricevente manda un "rapporto" su eventuali errori di trasmissione occorsi. Pertanto il protocollo prevede un modo per "scambiarsi il ruolo" alla fine della trasmissione delle immagini.

Full duplex

In una comunicazione **full duplex** entrambi i soggetti possono trasmettere e ricevere contemporaneamente.
In linea di massima in una comunicazione full duplex la velocità in ogni direzione sarà la metà di quella di una trasmissione half duplex che usi le stesse modalità di trasmissione.
Il telefono è un sistema full duplex.

Collegamenti asimmetrici

Molti sistemi di comunicazione full duplex danno modo di trasferire le informazioni alla stessa velocità in entrambe le direzioni del collegamento. Questo tipo di sistemi sono detti "simmetrici".
Peraltro alcuni sistemi full duplex, progettati per scopi specifici, sono in grado di realizzare una grande velocità in una direzione ed una molto più piccola nell'altra. Questi sistemi di comunicazione vengono detti **asimmetrici**.
Si ha la necessità di sistemi asimmetrici quando non è necessario avere la stessa velocità di comunicazione in entrambe le direzioni e si vuole, o si deve, sfruttare fino in fondo la "velocità" del mezzo di trasmissione.
Un esempio di collegamenti asimmetrici sono i sistemi di "video on demand".

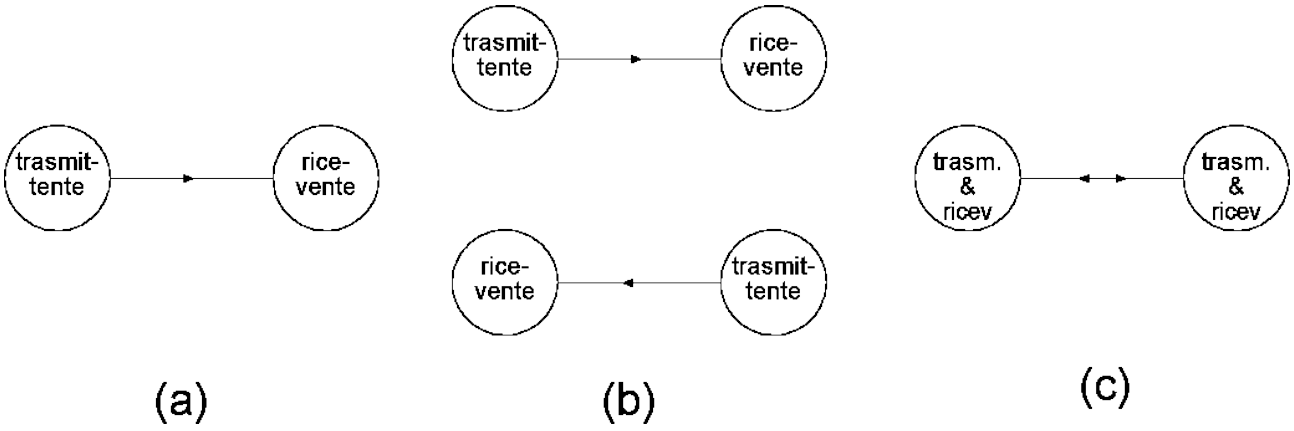

Figura 1: direzione del flusso (a) simplex, (b) half duplex (c) full duplex

Destinazione delle trasmissioni

In un sistema di comunicazione in cui siano presenti più di due soggetti i messaggi possono essere eventualmente diretti a più di un destinatario contemporaneamente.

Comunicazioni da uno a uno (unicast)

Si parla di "**unicast**" quando una comunicazione ha luogo fra un singolo sorgente ed un singolo destinatario. Tutti gli altri soggetti che eventualmente coesistano sulla stessa rete sono esclusi da quella comunicazione. Non è infatti necessario che l'architettura della rete che trasporta comunicazioni unicast debba comprendere solo i due soggetti coinvolti; comunicazioni unicast possono essere stabilite anche in reti complesse, in cui sono presenti milioni di altri soggetti. Per esempio, sono unicast le comunicazioni telefoniche.

Multicast

Si dicono "**multicast**" le comunicazioni da un sorgente a molti destinatari in cui solo un insieme ristretto di chi ha accesso alla rete può ricevere l'informazione. In pratica si tratta di trasmissioni in cui le informazioni vanno da un trasmettitore ad un gruppo ristretto di ricevitori (p.es. le trasmissioni radio della polizia).

Broadcast
Le comunicazioni in "**broadcast**" sono quelle che vanno da un sorgente a molti destinatari ed in cui chiunque abbia accesso alla rete può ricevere l'informazione trasmessa da un singolo soggetto. Si tratta di trasmissioni in cui le informazioni vanno da un trasmettitore a tutti i ricevitori possibili (p.es. la televisione "in chiaro").

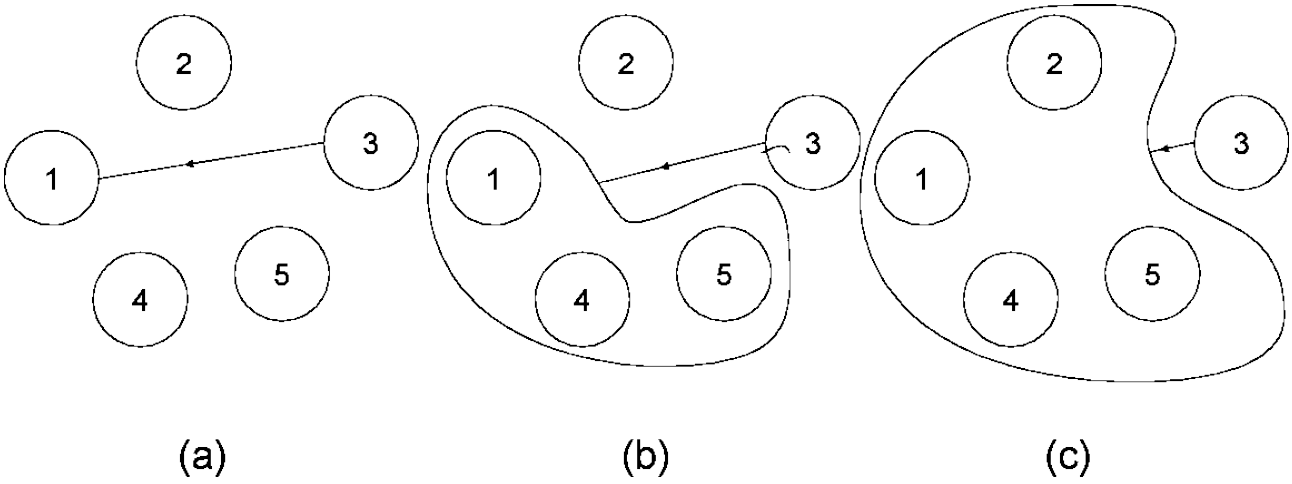

Figura 2: destinazione delle informazioni (a) unicast (b) multicast (c) broadcast

Collegamenti point to point o multidrop

Point to point
Si parla di collegamenti point to point (da punto a punto) quando il mezzo di trasmissione collega solo due punti della rete.
Questo è il modello dei mainframe centralizzati, collegati al resto del mondo con terminali "stupidi" cioè non in grado di funzionare autonomamente.
Nella successiva Figura 3 viene mostrato il collegamento point to point di ogni nodo della rete ad un "concentratore" (nodo 5). Questo tipo di collegamento viene anche usato nelle moderne reti di computer, in quelli che vengono detti "cablaggi a stella" (vedi oltre).

Multidrop
I collegamenti multidrop (o multipunto) sono quelli in cui tratte dello stesso mezzo di trasmissione collegano una alla volta tutte le stazioni della rete.
Le reti reali possono essere costituite di tratti point to point uniti a tratti multidrop.
Una rete fatta con collegamenti multidrop è meno flessibile ed affidabile di una con collegamenti da punto a punto, ma usa una minore quantità di mezzo di trasmissione ed è quindi conveniente se il mezzo di trasmissione è molto costoso.

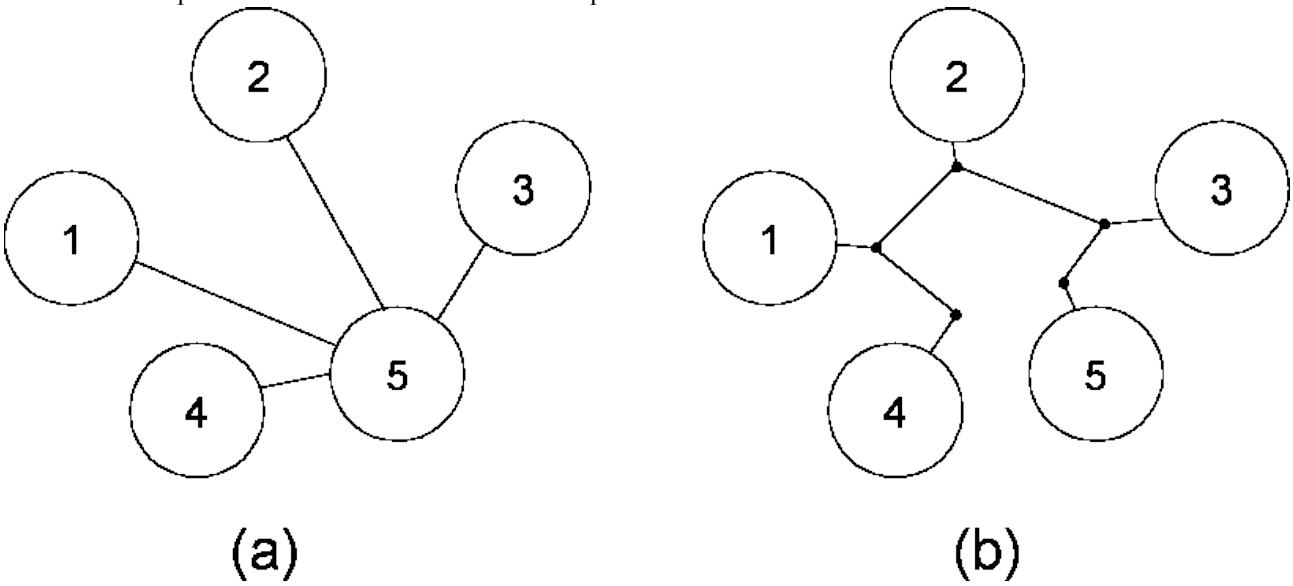

Figura 3: confronto fra collegamento point to point (a) e multidrop (b)

In figura i segmenti rappresentano collegamenti fisici. Si vedono quattro collegamenti point to point con il nodo 5. Passando tutte per il nodo 5 le informazioni possono essere trasferite fra tutti i nodi che le distribuisce agli altri.

Il punto (b) della figura illustra un collegamento multidrop nel quale i 5 nodi sono uniti da un unico cavo che fa "il giro" fra di essi.
La figura spiega in modo esauriente perché è necessario "meno cavo" per collegare tutti i nodi in una rete multidrop: nel caso della rete con collegamenti point to point è necessario collegare ogni volta il nodo concentratore a ciascuno degli altri, mentre nel caso del collegamento multidrop una volta "arrivati" ad un nodo si parte da quello per raggiungere il più vicino.
Per quel che riguarda l'affidabilità, proviamo a supporre che la rete abbia un guasto nel collegamento fra 2 e 3.
Se il collegamento è multidrop la rete non funziona più, oppure, nel caso più fortunato, è divisa in due sottoreti comprendenti le stazioni 4, 1 e 2 e le stazioni 3 e 5.
Se invece il collegamento è punto – punto, è impedita la sola comunicazione fra 5 ed il nodo che ha il collegamento rotto, mentre tutte le altre stazioni possono comunicare regolarmente. Naturalmente se in quest'ultimo caso è il nodo 5 a rompersi tutta la rete non può più funzionare.

Segnali, analogici o numerici (digitali)
TO DO

Parallelismo dei collegamenti
Le comunicazioni cui siamo maggiormente interessati sono quelle che coinvolgono i computer, ragion per cui considereremo molto spesso che i segnali trasmessi siano di tipo numerico, ovverosia che si trasmettano segnali "digitali".
Se la comunicazione prevede la trasmissione di segnali di tipo numerico è possibile realizzare collegamenti nei quali vengono trasmessi un solo bit alla volta oppure più bit contemporaneamente, ciò individua una distinzione fra collegamenti seriali e paralleli.

Collegamenti paralleli
Quando la trasmissione dello stesso messaggio avviene su più linee elettriche contemporaneamente si parla di comunicazione **parallela**. Un parallelismo molto usato è 8 bit, per quanto esistano diverse interfacce che scambiano in parallelo fino a 256 bit.
I cavi paralleli hanno almeno tanti fili quanto è il parallelismo di trasmissione, più alcuni altri per il controllo del flusso (vedi oltre). Per questo motivo e per i maggiori problemi di tipo elettrico rispetto alle interfacce seriali, i collegamenti paralleli sono limitati a distanze piccole.
La velocità di trasmissione nei collegamenti seriali si misura tipicamente in Byte al secondo (indicati dall'Autore come Byte/s e dal resto del mondo, molto impropriamente, come Bps (Byte per second)).

Collegamenti seriali
Quando il collegamento è su distanze superiori alla decina di metri, si preferisce evitare di portare molti fili da un capo all'altro. A tal fine si trasmettono le informazioni su un unico filo (o su una unica coppia di fili), un bit di seguito all'altro. Questo tipo di collegamento viene detto "**seriale**".
Dato che l'elaborazione dei dati all'interno del computer avviene sempre con un certo grado di parallelismo, per questo tipo di collegamento saranno necessari dispositivi in grado di effettuare una conversione da parallelo a seriale e viceversa.
La velocità di trasmissione nei collegamenti seriali si misura in bit al secondo (bit/s o bps[3] (bit per second)).

Controllo del flusso dei dati
Quando si parla di "controllo del flusso" si intende il metodo con cui i soggetti della comunicazione stabiliscono se e quando essa può aver luogo. Per poterlo fare essi si danno vicendevolmente l'autorizzazione a spedire e ricevere i messaggi.
In Inglese e nel gergo tecnico questo meccanismo viene detto "**handshake"** (stretta di mano, termine piuttosto autoesplicativo). Se si è certi che il destinatario non avrà mai la necessità di rallentare la trasmissione, si può effettuare una comunicazione anche senza handshake.
Ci possono essere due modalità di handshake: hardware e software. Si sceglierà fra di esse a seconda della velocità di risposta e del numero di linee di trasmissione che si vorrà avere.

Handshake hardware
Quando il sorgente è pronto a spedire dati lo comunica al destinatario attraverso linee elettriche approntate allo scopo. Il destinatario, con altre linee apposite, comunica il momento in cui la trasmissione può avere effettivamente luogo.
Un esempio potrebbe essere questo: il trasmittente alza una linea di "richiesta di trasmissione" (request o strobe), poi attende fino a che il ricevente risponde alzando un'altra linea, che ha il significato di "autorizzazione alla spedizione" (acknowledge), solo allora spedisce il messaggio. Quest'ultimo esempio è solo un'indicazione di massima, si tenga presente infatti che esistono molti schemi per realizzare l'handshake hardware, sostanzialmente ogni interfaccia hardware ha il suo. Alcuni schemi privilegiano l'affidabilità, altri la velocità di trasmissione.
Qualora uno dei due soggetti che comunicano abbia la necessità di non ricevere dati per un certo tempo basta che non che tenga bassa la linea di acknowledge e l'altro non spedirà altri dati.

3 Nella letteratura tecnica, che è di origine americana, l'unica differenza fra i Byte/s (che viene scritto "Bps") ed i bit/s (che viene scritto "bps") è la lettera maiuscola o minuscola. Ciò, oltre a mostrare ancora una volta la scarsa attitudine degli Americani con le unità di misura, può far sbagliare di un fattore di 8 solo scrivendo una lettera maiuscola invece che minuscola.

Si noti che il controllo dello stato delle linee di handshake può essere modificato via software (p.es con delle OUT, con un 8086) oppure può essere lasciato completamente ad un circuito elettronico che si occupa autonomamente della sua gestione (p.es con un 8255 programmato in modo 1). In entrambi i casi comunque la modalità di controllo del flusso viene definita "handshake hardware".

Handshake software

Invece che con linee fisiche il controllo del flusso avviene attraverso particolari simboli o messaggi, riconosciuti da entrambi i soggetti, che vengono trasmessi attraverso lo stesso canale e nello stesso modo del normale scambio di informazioni. In pratica chi riceve i dati, se si trova in difficoltà ad elaborarli alla velocità a cui arrivano, può spedire a chi trasmette un simbolo speciale che significa "fermati!". Il trasmittente, non appena riceverà questo simbolo interromperà la trasmissione e la riprenderà solo quando il ricevente, dopo che si è messo in condizione di accettare nuovi dati, spedirà un altro simbolo speciale, che significa "puoi riprendere a trasmettere".

È chiaro che per poter effettuare un handshake software è indispensabile che la comunicazione sia full duplex, dato che chi deve far sospendere la comunicazione deve poter spedire il codice mentre sta contemporaneamente ricevendo. Peraltro in questo caso non sarà richiesta la presenza di linee elettriche specifiche, da dedicare all'handshake.

Caratteristiche del traffico di rete

Trattiamo in questo paragrafo dei diversi tipi di traffico che possono essere scambiati in una rete. Anche questo caso è tipico delle reti che scambiano informazioni attraverso segnali di tipo numerico.

A seconda del sistema o dell'applicazione che genera il flusso d'informazioni le caratteristiche dei messaggi da spedire possono essere molto diverse.

L'aspetto principale che distingue i vari tipi di flusso di informazioni (di "traffico") è la tolleranza alla variazione dei ritardi di trasmissione. Alcuni messaggi possono essere trasmessi con ritardi variabili, anche considerevoli; altri tipi di traffico non ammettono grosse variazioni nei ritardi, perché altrimenti la loro validità viene inficiata.

Traffico burst

Un "**burst**" (raffica) è un blocco di dati che deve essere trasmesso ad alta velocità.

Si dicono "burst" quelle trasmissioni che alternano momenti in cui c'è la necessità di un'alta velocità ad altri in cui la trasmissione potrebbe essere del tutto inattiva. Il traffico dati generato da una trasmissione burst è perciò concentrato in istanti specifici, mentre per la gran parte del tempo non esiste.

Le trasmissioni di dati burst più tipiche sono quelle nelle reti di computer. Le esigenze di questo tipo di trasmissioni sono infatti concentrate nel tempo e durano per un tempo relativamente breve. Si pensi per esempio alla richiesta di un file da un server di rete. Quando la stazione non ha bisogno dei dati potrebbe anche fare a meno della rete, ma nel momento in cui richiede il file, allora ha bisogno di velocità ed affidabilità (mancanza di errori), ma può sopportare lievi ritardi, anche variabili.

Il traffico burst, a volte detto "traffico asincrono", è imprevedibile per sua natura e si presta difficilmente ad essere "pianificato".

Traffico real time

I dati real time hanno vincoli stringenti di tempo. Richiedono di essere trasmessi con ritardi massimi fissi, se possibile anche "garantiti". Se i dati giungono più tardi del tempo massimo previsto, essi non sono più utilizzabili. Questo tipo di traffico richiede un flusso di dati continuo ad una velocità costante o con piccole variazioni. Non può tollerare ritardi della rete lunghi o troppo variabili.

Un esempio tipico è la trasmissione della voce non compressa su reti numeriche. In questo caso la voce viene campionata ed i dati divisi in blocchi, che vengono successivamente trasmessi. Se i dati non arrivano a destinazione in tempo, essi non potranno essere riprodotti alla ricezione, traducendosi in un disturbo nel segnale audio. Se ciò accade frequentemente la qualità della riproduzione può divenire inaccettabile. Per contro la presenza di un errore non sarà decisiva, come succede per esempio in un trasferimento di file. Per il traffico real time si possono quindi accettare alcuni errori ma non troppo ritardo nella trasmissione, né variazioni troppo accentuate di questo ritardo.

1.1 Mezzi di trasmissione

Mezzi guidati

Si dicono "guidati" quei mezzi di trasmissione in cui l'onda elettromagnetica che trasporta l'informazione si propaga nella stessa direzione del mezzo. In questi casi l'onda viene "incanalata" nel mezzo e costretta a seguire il suo percorso. In parole povere, i mezzi guidati sono i cavi e consimili. Sono mezzi guidati: il doppino telefonico, il cavo coassiale, la fibra ottica. Ne tratteremo in seguito.

Mezzi non guidati

I mezzi non guidati sono quelli in cui l'onda elettromagnetica che trasporta l'informazione si può propagare in tutte le direzioni. Perché ciò accada sono necessari un dispositivo per la trasmissione, che sia in grado di far propagare l'onda, ed uno per la ricezione, che la raccolga; essi di solito si chiamano antenne. I mezzi non guidati sono l'aria ed il vuoto (in rari casi l'acqua!).

In appendice si trova un capitolo sui mezzi di trasmissione.

1.2 Sincronizzazione elettrica

Per effettuare una comunicazione elettrica è necessario che tutti i soggetti interessati conoscano esattamente il momento in cui i dati sono da considerare "buoni" per la lettura da perte del ricevente. I due sistemi si devono dunque "sincronizzare".

Interfacce asincrone

Nelle interfacce asincrone gli istanti in cui incomincia la trasmissione dell'informazione non sono stabiliti "a priori"; è possibile che fra la trasmissione di un "simbolo" di informazione ed il successivo passi un tempo qualunque.

Il trasmittente ed il ricevente possiedono due "orologi" indipendenti per misurare il tempo e

Ogni volta che viene spedito un "simbolo" od un pacchetto di informazioni, il trasmittente ed il ricevente si sincronizzano, poi possono perdere il contatto per un tempo indefinito.

Vantaggio: richiede una limitata accuratezza relativa dei due orologi. E 'un sistema più semplice

Svantaggio: la risincronizzazione frequente causa overhead.

Interfacce sincrone

Nelle interfacce sincrone l'orologio del trasmittente e quello del ricevente si mantengono sincronizzati in modo permanente, attraverso uno di due metodi illustrati in seguito.

1. un segnale di sincronizzazione (clock) viene trasportato contemporaneamente all'informazione, su di una linea separata. Questo tipo di collegamento sincrono è tipico dei bus del computer e non è agevole per le comunicazioni fra computer, che avvengono su distanze di solito tanto grandi da non poter garantire la coerenza delle temporizzazioni di sincronizzazione.
2. la sincronizzazione fra ricevitore e trasmettitore viene stabilita tramite lo stesso segnale che trasporta l'informazione, ma il segnale è fatto in modo da garantire il mantenimento della sincronizzazione, per esempio assicurando la presenza di una transizione per ogni bit trasmesso.

In questo caso la coerenza fra i due orologi va mantenuta per tempi più lunghi. I due orologi devono essere perciò più stabili. Inoltre è necessario che venga sempre spedito un segnale, che mantiene la sincronizzazione, anche nei momenti in cui non serve spedire alcuna informazione.

1.2.1 Uso del mezzo di trasmissione

Abbiamo detto che in presenza di dati real time c'è la necessità di collegamenti a velocità di trasferimento costante. Dovendo fare la scelta di un mezzo di trasmissione, verosimilmente useremo quello che offre una velocità solo di poco superiore a quella, costante, richiesta dal traffico dati.

Il discorso non regge quando, invece, siamo in presenza di traffico burst. In questo caso potremmo scegliere un mezzo, veloce, in grado di poter sopportare la massima velocità dei momenti di picco, durante i burst. Fatta questa scelta ci accorgeremmo che quel mezzo è inutilizzato per la maggior parte del tempo, perché il traffico burst spesso non c'è!

Dato che la maggiore velocità del canale di comunicazione significa anche un suo maggior costo, questo sarebbe un grande spreco di risorse.

Queste due diverse esigenze individuano diverse modalità di uso del canale di trasmissione, che provvediamo ad introdurre.

Reti a mezzo di trasmissione condiviso ("shared medium" networks)

Ogni stazione si collega ad un unico mezzo di trasmissione, comune. In queste reti una stazione può trasmettere contemporaneamente a tutte le altre. Per questo vengono anche dette "broadcast network". In questo caso tutte le stazioni collegate si dividono la velocità del mezzo. Se il mezzo può trasportare 10 Mbit/s e 20 stazioni necessitano di trasmettere a 1 Mbit/s le stazioni dovranno attendere per accedere al mezzo di trasmissione.

La condivisione del mezzo di trasmissione significa che esso sarà sfruttato in modo più efficiente in presenza di traffico burst. Infatti, quando un nodo della rete non avrà necessità di trasmettere, ce ne potrà essere un alto pronto a farlo. Così il mezzo di trasmissione sarà impegnato per più tempo alla sua massima velocità.

Reti a switch condivisi ("switched" networks)

I nodi si collegano ad un nodo speciale (switch), che a sua volta li collega fra di loro, a due a due od a gruppi, con un circuito fisico. In questo modo nella rete sono possibili diverse comunicazioni contemporanee alla velocità massima dei mezzi di trasmissione.

Reti a mezzo di trasmissione non condiviso ("crossbar" networks)

Ogni stazione si collega direttamente, con link da punto a punto, a tutte le altre stazioni. In questo caso ciascuno di questi collegamenti può funzionare alla massima velocità possibile nel mezzo di trasmissione prescelto; non c'è rallentamento della velocità di comunicazione dovuto alla presenza di più stazioni in rete.

Spedizione dei messaggi sulla rete

Reti a commutazione di circuito (circuit switching)

Si parla di reti a commutazione di circuito quando, prima della comunicazione, vengono effettuati dei collegamenti temporanei (fisici o virtuali) fra i soggetti della comunicazione, che così acquisiscono il mezzo di trasmissione in modo esclusivo, almeno per il tempo in cui il circuito è attivo. Perché una comunicazione di questo genere possa avvenire è prima necessario che sia stabilito a priori, prima dell'inizio della comunicazione, il percorso che il messaggio deve pren-

dere. Ciò non avviene nei due casi successivi, in cui il percorso non deve necessariamente essere stabilito prima dell'inizio della comunicazione.
In reti a commutazione di circuito per cambiare il destinatario di una comunicazione è necessario cambiare il circuito.

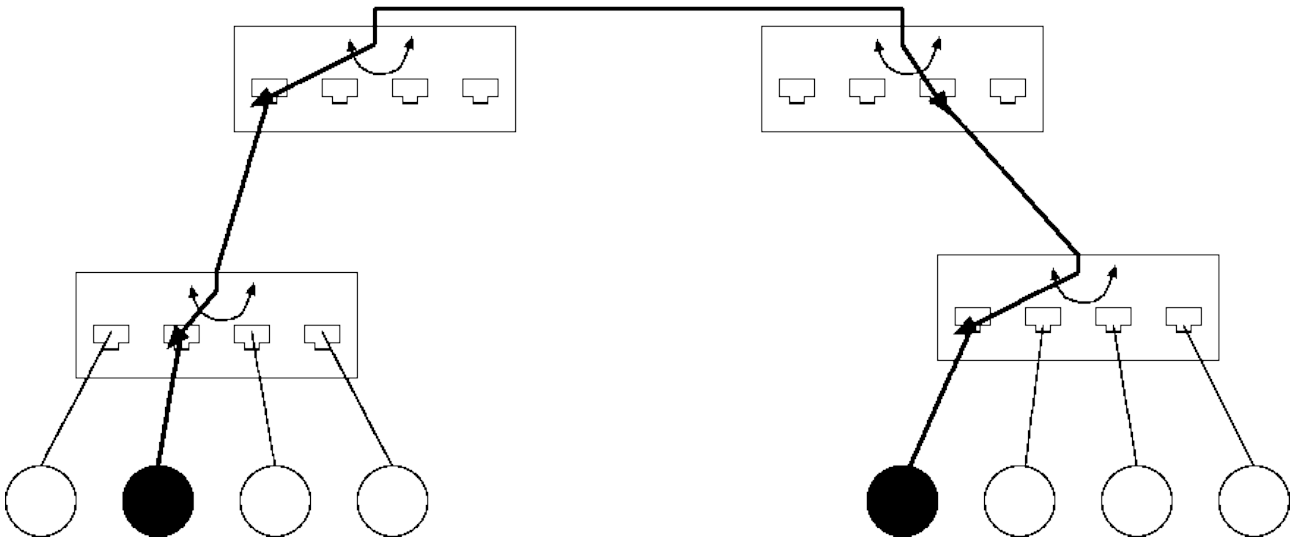

Figura 4: formazione di un canale di trasmissione con commutazione di circuito

Reti a commutazione di messaggio (message switching)

Nelle comunicazioni non sempre è necessario utilizzare un canale di comunicazione fino al massimo delle sua possibilità. Per esempio, ci possono essere momenti in cui non si deve trasferire alcun dato, altri in cui serve avere la massima velocità possibile. In una rete a commutazione di circuito in queste situazioni la capacità del canale (larghezza di banda) viene "sprecata".
Per rendere più efficiente le reti ed utilizzare al meglio la larghezza di banda dei canali si può usare lo stesso canale per trasportare informazioni di soggetti diversi. Nel canale verranno convogliati messaggi destinati a soggetti diversi, così quando un soggetto non ha la necessità di trasmettere, ce ne potrebbe essere un altro che sfrutta il canale.
Invece di essere il circuito che viene commutato per raggiungere la destinazione è il messaggio che, con tecniche diverse sulle quali torneremo poco più avanti, viene fatto giungere a destinazione condividendo i mezzi di trasmissione.
Naturalmente perché ciò accada è necessario che il messaggio di una stazione possa in qualche modo raggiungere la sua destinazione e bisogna anche poter distinguere fra le diverse stazioni che stanno in rete.
Per poter giungere alla stazione giusta il messaggio dovrà contenere una qualche forma di identificazione. Di solito la stazione di destinazione viene identificata con un numero, che prende il nome di **indirizzo** di destinazione. Il messaggio spedito contiene anche l'indirizzo di identificazione del mittente, in modo che la stazione ricevente possa sapere chi le spedisce il messaggio, per potergli rispondere.
In reti a commutazione di messaggio i messaggi lunghi sono spediti tutti di seguito, fino a che non sono terminati.

Reti a commutazione di pacchetto (packet switching)

Se un nodo in una rete a commutazione di messaggio deve trasmettere un messaggio molto lungo, blocca tutte le trasmissioni in quella rete per molto tempo. Per evitare questo problema si usano le reti a commutazione di pacchetto.
A questo scopo si dividono i messaggi in unità più piccole, detti "pacchetti" (packets), e si spediscono, uno di seguito all'altro, pacchetti appartenenti a molte comunicazioni lungo lo stesso canale trasmissivo. Naturalmente questi pacchetti dovranno portare, oltre all'indicazione del destinatario e del mittente, anche qualche forma di **numerazione** progressiva del **pacchetto**, perché chi riceve possa verificare l'arrivo di tutti i pacchetti in cui il messaggio originario era stato suddiviso. Nelle reti a pacchetto il canale di trasmissione è occupato solo per la durata della trasmissione di un singolo pacchetto. Poi viene liberato e può essere "preso" da un'altra stazione. Così tutte le stazioni avranno la possibilità, ogni tanto, di trasmettere, anche quando una di esse stia trasmettendo un messaggio molto lungo.

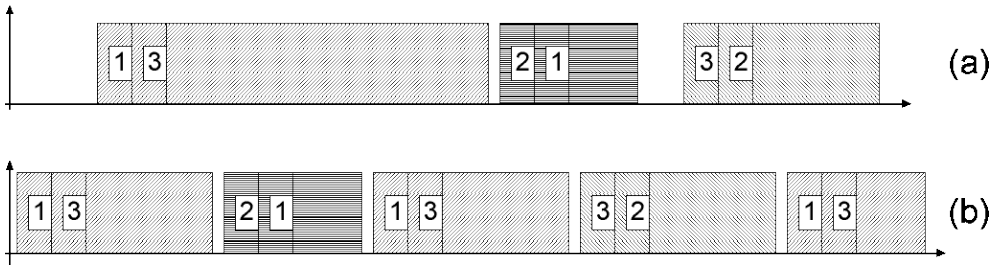

Figura 5: confronto fra commutazione di messaggio (a) e commutazione di pacchetto (b)

Accesso al mezzo di trasmissione
Quando il mezzo di trasmissione è condiviso fra più nodi della rete è necessario che uno solo alla volta trasmetta, anche quando diversi ne avrebbero la necessità. Per individuare quale, fra i nodi in attesa di trasmettere, lo potrà fare, si usano diverse tecniche. Vediamo quelle più comuni:

Accesso master - slave
In questo caso esiste nella rete un soggetto dotato di maggiore autorità sugli altri, detto master (padrone). Questo soggetto, che è unico in un certo periodo di tempo, può stabilire in ogni istante chi, fra gli slave (schiavi) che ne hanno bisogno, deve avere accesso in trasmissione al mezzo. Il master perciò controlla tutto il flusso delle informazioni e senza di esso non è possibile che avvenga alcuna comunicazione.
L'accesso master - slave è tipico della maggior parte dei bus interni dei computer e si usa in quelle interfacce che seguono la stessa filosofia del bus interno, estendendola anche alle periferiche. Un esempio è l'interfaccia IEEE 488, che si usa nella comunicazione con strumenti di misura e di laboratorio, od in alcune reti di campo per l'automazione industriale.
In alcuni tipi di rete master – slave il master può cedere il controllo ad un altro nodo della rete, che può fungere da master a sua volta, a patto che in ogni istante ci sia un solo master "in carica".

Accesso a contesa (contention access)
Nel caso dell'accesso a contesa non esiste un master che decide per tutti, ma tutti devono ottenere il mezzo di trasmissione in concorrenza fra loro, rispettando regole specifiche che assicurano che un solo soggetto alla volta possa trasmettere. Queste regole vanno sotto il nome di MAC (Medium Access Control). Nelle reti locali i MAC più tipici sono CSMA e "Token pass". Nelle reti di telecomunicazioni e in quelle di automazione industriale è usato il MAC TDMA (Time Division Multiple Access).

1.3 Tipi di rete
Abbiamo già dato una definizione di rete, vediamone ora altre, più specifiche e pertinenti al campo che ci interessa.

Reti per audio

Reti telefoniche
Reti complesse, per audio con banda limitata, adatte a trasmettere a bassa qualità una voce umana intelligibile. Reti in cui, prima della comunicazione, viene stabilito, per mezzo di centralini di commutazione un collegamento punto-punto fra i due utenti. Usano una grande varietà di mezzi di trasmissione, ma sono generalmente cablate in rame per la parte finale di connessione agli utenti finali ("ultimo miglio").

Reti radiofoniche
Reti di tipo broadcast o multicast, su mezzo di trasmissione non guidato (onde elettromagnetiche propagate in atmosfera), che trasmettono segnali analogici modulati in ampiezza od in frequenza. Trasmettono per scopo di intrattenimento o per ordine pubblico, emergenza, .. .

Reti per video
Reti di tipo broadcast o multicast che trasmettono segnali audio + video per scopo di intrattenimento, di informazione o per segnalazione - sicurezza.

Reti televisive
Reti che trasmettono il loro segnale via "etere", da antenne a terra o da satellite. Attualmente basate per la gran parte su tecniche di trasmissione analogiche, si va sempre più diffondendo la tecnica digitale (DVB Digital Video Broadcast), soprattutto nelle reti più avanzate che trasmettono via satellite o ad alta definizione (HDTV).

Reti per TV via cavo
Reti per video su mezzo di trasmissione guidato. Nelle reti più diffuse di questo tipo il supporto è un cavo coassiale a larga banda di alluminio, che, interrato nelle città, raggiunge la maggior parte delle abitazioni negli USA ed in alcune nazioni europee. La trasmissione può avvenire in broadcast od in multicast ed è asimmetrica. Di solito c'è la possibilità tecnica per l'utente di interagire con il sistema. I cavi trasportano molti canali video con tecniche di multiplazione FDM. In Italia alcune grandi aziende, come la Società Autostrade o le FS hanno reti interne di sorveglianza video, alcune cablate in fibra ottica.

Reti per "streaming" o "video on demand"
Al contrario delle altre reti per video, non funzionano in broadcast. Quando la trasmissione è multicast di solito si parla di "streaming video", mentre si si trasmette in unicast, realizzando una sorta di collegamento punto - punto "virtuale" fra la centrale e l'utente, si dice che la trasmissione è in "video on demand". In questo caso ogni utente della rete può vedere una trasmissione diversa.
Le tecnologie odierne per la trasmissione e l'elaborazione dei segnali rendono possibile la realizzazione di queste reti, che sono già vendute commercialmente in modo massivo. I mezzi di trasmissione per il video on demand possono essere sia guidati che non guidati (rete telefonica o rete wireless metropolitana). I protocolli di più altro livello ormai affermati sono quelli di Internet (IP = Internet Protocol).

Reti per dati e reti di computer
Si parla di rete di computer quando un computer in grado di funzionare autonomamente è collegato ad almeno un altro computer, anch'esso in grado di funzionare autonomamente, ed essi sono in grado di scambiarsi messaggi e/o condividere alcune delle loro risorse (memorie di massa, stampanti ..).

Reti locali (LAN)
La sigla LAN significa Local Area Network. Una LAN è una rete di computer ubicati a distanza relativamente piccola (dell'ordine del chilometro). Data la distanza relativamente piccola, i computer possono essere collegati con mezzi di trasmissione e dispositivi che permettono alte velocità, con un tasso di errori di trasmissione molto basso.
I computer collegati alla LAN possono essere del tutto diversi, nell'architettura (e quindi con diverse famiglie di CPU) e nel software (con diversi sistemi operativi). Essi rispondono peraltro allo stesso standard di rete, che stabilisce, fra le altre cose, anche come devono essere rappresentate le informazioni, e sono quindi in grado di "**interoperare**" cioè di funzionare insieme.
Dal punto di vista fisico tutti i dispositivi che sono collegati ad una LAN sono uguali. Per esempio, i dati che vengono trasmessi ad una stampante con scheda di rete, che non ha neppure un "monitor" vengono trasmessi nello stesso modo di quelli che vanno ad un mainframe con centinaia di terminali. Ciò vuol dire che la differenziazione delle funzioni dei dispositivi presenti in una LAN avviene a livelli di astrazioni più alti del livello fisico. A livello fisico ogni computer di una LAN è uguale agli altri.
I computer che sono collegati ad una rete vengono chiamati in molti modi, come vedremo in seguito. Per indicare un generico computer collegato alla rete, senza alludere ad alcuna delle sue specifiche funzioni, useremo il termine stazione ("workstation" o "station") o anche nodo.

LAN da pari a pari (**peer to peer** o paritetiche)
Nelle reti peer to peer nessuna delle stazioni ha più importanza rispetto alle altre. Ciascuna delle stazioni può mettere a disposizione delle altre alcune delle sue risorse, come cartelle di file, stampanti ad essa collegate, altri dispositivi di memorizzazione come nastri di backup e molte altre "risorse". Nelle reti peer to peer tutte le stazioni sono in grado di realizzare tutte le funzioni di rete, cioè sono in grado sia di mettere a disposizione risorse, sia di utilizzare quelle messe a disposizione da altre stazioni.
La protezione nell'accesso alle risorse nelle reti peer to peer è regolata a livello della singola stazione. In ogni singola stazione che mette a disposizione risorse di rete, ci sarà un elenco di utenti diverso dalle altre, ed un elenco di permessi per l'accesso alle risorse diverso da tutte le altre stazioni.

LAN basate su server (client - server o "server based")
Nelle reti client - server una o più stazioni hanno la funzione di mettere a disposizione e gestire le risorse da condividere. Le stazioni che hanno questa funzione vengono dette "**server**", mentre quelle che usano solo le risorse di altri sono dette "**client**".

I server di una rete basata su server possono essere **dedicati** o **non dedicati**. I server dedicati possono svolgere solo la funzione di server di rete, quelli non dedicati possono essere usati anche come computer "general purpose", cioè per l'utilizzo normale (p.es. per elaborazione testi o calcolo con tabelloni elettronici).
Le reti basate su server permettono un accesso molto più controllato e garantiscono perciò una sicurezza molto maggiore. Inoltre, se il server è dedicato, danno anche migliori prestazioni.
Alcuni S.O. che forniscono funzionalità peer to peer sono in grado anche di collegarsi a server di rete, funzionando anche come client. In questo caso essi hanno un accesso controllato alle risorse, attraverso le funzioni del server. Peraltro essi sono anche in grado di mettere a disposizione della rete alcune delle loro risorse. Questo approccio è perciò di tipo "peer" per quanto riguarda le risorse, mentre ha le caratteristiche delle reti a server per l'accesso controllato e la sicurezza. Questo è il caso p.es. di Windows XP (o di una altro S.O. Windows di "tipo client"), quando funziona come client di reti Netware, Unix o con S.O. Windows di tipo "server" (NT, 2000 Server o 2003 Server).

È importante notare che la differenza fra reti peer to peer e client-server è data esclusivamente dal diverso software che vi gira, mentre l'infrastruttura hardware necessaria per i due tipi di rete potrebbe anche essere la stessa e, come già detto, è intrinsecamente paritetica.

Reti di fabbrica e di campo

Una categoria particolare di reti, assimilabile alle LAN, è quella delle reti di fabbrica. Sono le reti che servono a trasportare le informazioni, rilevate "sul campo" dai sensori, ai sistemi di controllo e supervisione che regolano il funzionamento di ogni ambiente di produzione automatizzato. Oltre che la regolazione automatica dei processi di produzione, lo scopo di queste reti è anche la gestione della produzione e della qualità.

MAP

TODO

Field bus (reti di campo)

Una necessità molto sentita nel campo dell'automazione industriale è quella della minimizzazione del numero e della lunghezza dei cavi richiesti. Si immagini una linea di produzione industriale con centinaia o migliaia di sensori e attuatori. Fino a poco tempo fa ciascuno di essi veniva collegato ad un armadio di controllo con almeno due fili. Essi servivano per la trasmissione del segnale al sistema centrale di regolazione e controllo. Questa trasmissione avveniva di solito per mezzo di un segnale analogico, in corrente, da 4 a 20 mA. Con l'affermarsi dei microcontrollori a bassissimo costo i sensori di oggi possono essere "intelligenti", contendo una CPU, e si possono collegare in reti "a bus", cablate linearmente (vedi il paragrafo "Disposizione delle stazioni in una rete"). I cavi collegano i sensori uno dopo l'altro, portando, in linea di principio, solo due di fili all'armadio, in luogo delle centinaia necessarie usando l'approccio più vecchio. Questo permette di ottenere grossi risparmi sia in cavo che in lavoro per il cablaggio. A ciascun sensore dovrà essere assegnato un numero univoco, detto indirizzo, che servirà per identificarlo quando dovrà spedire e ricevere informazioni.
La stessa necessità è sentita nelle automobili, ove i sistemi elettronici sono sempre più numerosi, aumentando il numero di cavi necessari. Con una rete a bus tutti i dispositivi possono essere collegati con due soli fili per le comunicazioni e due, più spessi, per la potenza.

Reti geografiche (WAN)

La sigla WAN significa Wide Area Network e, come dice il nome, implica la copertura di aree più estese di quelle coperte dalle LAN. I mezzi di trasmissione ed i dispositivi di comunicazione utilizzati nelle WAN sono di solito in grado, almeno allo stato attuale dell'arte, di consentire velocità di trasmissione molto più basse di quelle ottenibili con una LAN.

Internet

Oggi è la rete geografica per eccellenza, la rete delle reti, che connette milioni di computer che usano lo stesso insieme di protocolli di rete (TCP/IP).

WAN aziendali

Prima dell'esplosione del fenomeno Internet, gli esempi più importanti di reti geografiche erano le reti private realizzate dalle grandi aziende multinazionali per collegare le loro divisioni sparse per il mondo. Fino a non molto tempo fa esse erano di solito basate su software scritto ad hoc e non standard. Dopo l'esplosione di Internet e delle sue tecnologie le stesse tecniche usate per Internet sono state "riciclate" anche nelle reti aziendali e sono stati coniati allo scopo due nuovi termini:

Intranet

Una Intranet è una rete privata che ha una infrastruttura hardware del tutto separata dalla grande rete Internet, ma che usa gli stessi strumenti software e gli stessi protocolli. La separazione totale dalla Internet garantisce maggiore sicurezza, dato che nella grande rete non sono rari i casi di persone che, incoraggiate anche dall'illusione della anonimato, possono provare a collegarsi ed eventualmente a far danni nei computer altrui. Dato che la rete fisica di una Intranet è del tutto separata da quella della Internet essa offre perciò maggiori garanzie di sicurezza. Naturalmente l'infrastruttura separata della Intranet renderà necessario acquistare tutte le apparecchiature e le linee di trasmissione che la realizzano, senza condividere la spesa fra molti utenti, come succede in Internet. Ciò rende i costi spesso tanto alti da poter essere affrontati solo da aziende molto grandi.

Extranet
Una Extranet è una rete privata che condivide la infrastruttura hardware e software di Internet. La sicurezza deve quindi essere garantita dal solo software. Al giorno d'oggi questa soluzione è più economica ma molto meno sicura della Intranet. È comunque in corso un grosso sforzo nello sviluppo di protocolli standard per i trasferimenti "sicuri" in Internet, la crittografia e l'autenticazione dell'utente. Si parla in questo caso di **VPN**: Virtual Private Network, cioè di reti che sono "virtualmente private" perché garantiscono la riservatezza delle comunicazioni come farebbe una rete ad infrastruttura privata anche se i dati vengono effettivamente trasmessi su una rete pubblica.

Collegamenti su rete telefonica

Per collegare due computer attraverso il sistema telefonico esistono due possibilità: collegamenti su rete commutata e collegamenti su linea dedicata (o "affittata" (leased)).

Collegamenti su linea telefonica commutata
Usano la rete telefonica "normale". Per effettuare un collegamento si chiamerà il numero del destinatario come se si volesse chiamare "a voce". Una volta che il destinatario risponde, i due sistemi possono iniziare la comunicazione, con le tecniche che verranno spiegate in seguito.

Collegamenti su linea telefonica dedicata
Usano particolari linee telefoniche, affittate dal gestore della telefonia. Queste linee sono sempre attive fra i due punti e per collegarsi non c'è bisogno di selezionare il numero dell'altro punto.

Le caratteristiche delle linee dedicate sono: buona qualità elettrica del collegamento (che può significare una maggiore velocità di trasmissione), equalizzazione fissa, costo fisso, non dipendente dal tempo di collegamento, né dal traffico.
I vantaggi dei collegamenti su linea commutata sono: costo a tempo, quando non si usa non si paga.
Per scegliere bisognerà tener conto della densità di traffico prevista, della sua distribuzione nell'arco della giornata, della necessità di trasmissione in tempo reale. Se la trasmissione deve avvenire in tempo reale per tutto l'arco del giorno la linea dedicata è inevitabile, la linea commutata è invece molto vantaggiosa se il traffico è concentrato tutto in alcuni momenti della giornata.

Un'altra classificazione dei collegamenti telefonici fra due computer distingue i collegamenti su linea analogica da quelli su linea digitale.

Collegamenti su linea telefonica analogica
E' ancora il collegamento telefonico più usato, che fa uso delle normali linee telefoniche che sono in tutte le case da circa un secolo e che costituiscono quello che viene definito "POTS" (Plain Old Telephone System, il "buon vecchio" sistema telefonico). Per far comunicare due computer attraverso linee analogiche è necessario utilizzare un modem per ogni estremo del collegamento (vedi oltre).

Collegamenti su linea telefonica numerica (ISDN)
E' il modo di collegare due computer per telefono che verrà usato in un futuro prossimo. La rete telefonica del futuro sarà basata, a tutti i livelli, su tecniche numeriche (digitali). Anche nel telefono di ogni casa arriveranno flussi di bit, invece di un segnale analogico da riprodurre nella cornetta. L'apparecchio telefonico avrà il compito di effettuare la conversione A/D e D/A e di far partire le relative sequenze di bit verso la centrale telefonica. Il collegamento fra due computer utilizzerà la stessa tecnica di trasmissione e, naturalmente, non avrà bisogno di conversioni A/D e D/A.

1.3.1 La "convergenza" e le "reti integrate"

E' ormai chiara da tempo la tendenza verso il digitale di tutte le tecniche di trasmissione. La telefonia sta lentamente ma senza sosta convertendosi al digitale; ormai tutte le comunicazioni fra le centrali telefoniche avvengono tramite segnali digitali, mentre i collegamenti degli utenti sono in grandissima parte analogici. Peraltro è possibile, su richiesta, per la maggioranza degli utenti italiani estendere il collegamento con tecniche digitali fino al proprio apparecchio telefonico (vedi ISDN, nel seguito). La sostituzione completa delle linee analogiche è comunque piuttosto lontana nel tempo. Le reti telefoniche cellulari tendono verso il digitale, oggi la rete più diffusa è ancora analogica, ma sarà presto superata dal GSM digitale, che garantisce più servizi ed una occupazione della banda molto più efficiente.
Anche la televisione segue la tendenza: alcuni tipi di trasmissioni pay TV sono già digitali, mentre lo sarà, sin dall'inizio, anche la TV ad alta definizione.
E' perciò in atto una "convergenza", in tutti i settori delle telecomunicazioni, verso le tecniche digitali. A convergenza avvenuta, qualunque informazione verrà trasmessa con tecniche digitali, non avrà perciò molta importanza chi e come la trasmetterà, ma solo con quale servizio ed a qual costo. Le aziende che muoveranno dati fra i continenti saranno perciò molte, con estrazioni diverse. Saranno senz'altro in gioco le aziende che oggi gestiscono la telefonia pubblica, ma anche quelle che fanno broadcast televisivo o possiedono reti interne, come le Ferrovie, le Autostrade o certe banche, oltre a nuovi soggetti che nasceranno con l'obbiettivo dei nuovi mercati. E' quindi prevedibile un rimescolamento dei settori nei quali operano le aziende di telecomunicazioni, con spostamenti ed "invasioni" di campo che porteranno concorrenza, cioè progresso tecnico più veloce e benefici economici agli utenti.

1.4 Disposizione delle stazioni in una rete

Parlando di "disposizione" delle stazioni bisogna distinguere due livelli. Un livello "fisico", che riguarda a che cosa è collegata fisicamente ognuna delle stazioni, ed un livello "logico", che si riferisce a con chi "parla"direttamente ciascuna delle stazioni. Nella terminologia al riguardo c'è molta confusione, ogni autore dà la sua interpretazione, molti non fanno distinzioni fra livello logico e fisico. Infatti nel gergo corrente, viene spesso usato il solo termine "topologia", sia per intendere il collegamento fisico fra le stazioni di una rete, sia per intendere l'ordine ed il modo con il quale le stazioni comunicano fra di loro. Alcuni, per distinguere fra i due casi, parlano di "topologia logica" e "topologia fisica".
Siccome fra i due concetti c'è un'effettiva differenza, riscontrabile anche nelle architetture di alcuni tipi di reti locali esistenti, in questo testo si considera la topologia una caratteristica astratta della rete, mentre si dà il nome di "cablaggio" o "layout" ai collegamenti fisici effettivamente realizzati fra le stazioni.

1.4.1 Schemi di cablaggio (wiring layout)

Per noi il layout di una rete è dunque l'insieme dei cammini percorsi dai cavi di collegamento delle stazioni che la costituiscono. Esso dipende fortemente dalle caratteristiche dell'edificio da cablare. I costi del cablaggio possono essere minimizzati con una saggia scelta degli schemi utilizzati

Cablaggio lineare
Il cavo collega, in un unico percorso, a due a due tutte le stazioni, una di seguito all'altra. Una linea fisica unisce una dopo l'altra tutte le stazioni.

Cablaggio a stella
I cavi iniziano tutti da un unico punto, e terminano in ciascuna delle stazioni. In pratica, ciascuna stazione è collegata da punto a punto con un "concentratore", ove convergono tutti i cavi della rete.
In molti tipi di rete da ogni nodo di una rete cablata a stella può partire una nuova stella. In questo caso il cablaggio viene definito "ad albero" o a "stella connessa".

Cablaggio ad albero o a stella connessa
In genere si intende fare un piccola differenza fra cablaggio ad albero o a stella connessa. Per cablaggio ad albero si intende un cablaggio che stabilisce una gerarchia o una direzione fra i vari rami della rete, mentre in una stessa connessa i vari rami sono "uguali".

Cablaggio a maglia
Si parla di cablaggio a maglia quando per raggiungere la destinazione è possibile percorrere più di un percorso.

Con il cablaggio lineare si può minimizzare la quantità di cavo utilizzata, non essendo necessario concentrare tutti i cavi in un solo punto. Per contro con tale cablaggio un'interruzione del cavo porta alla caduta di tutta la rete, mentre lo stesso evento in un cablaggio a stella porta alla caduta del solo collegamento da punto a punto fra il concentratore e la stazione interessata (ciò, naturalmente, solo nel caso un cui il concentratore sia in grado di rilevare automaticamente il guasto ed escludere la stazione rotta). Con un cablaggio a stella anche la diagnostica e la gestione della rete sono semplificate.
Nelle applicazioni reali di solito si usa un approccio misto; per esempio si può scegliere di collegare con cablaggio lineare alcuni punti di concentrazione (p.es. uno per ogni piano di un edificio, in una LAN) dai quali poi si diparte un cablaggio a stella.

Cablaggio strutturato
Il cablaggio strutturato è l'idea del cablaggio a stella portata all'estremo. In questo tipo di cablaggio si concentrano in un unico punto *tutti* i cavi (di solito UTP o STP (vedi oltre)) che forniscono comunicazioni all'azienda, quindi non solo i cavi di rete, ma anche i cavi telefonici. Tutti i cavi termineranno, nell'armadio del concentratore, su di un pannello con connettori simili a quelli delle prese telefoniche americane, ma un po' più grandi. Vicino a questo pannello ci saranno gli ingressi del centralino telefonico e quelli dei dispositivi della rete locale (hub). Con corti cavi (patch) si possono collegare i connettori dell'armadio concentratore al dispositivo che si vuole collegare a quel cavo. In questo modo si collegano le prese a muro "del telefono", presenti in ogni ufficio, alle apparecchiature relative alla telefonia od a quelle della rete locale. Se si vorranno fare delle modifiche al cablaggio della rete locale o della rete telefonica, nella maggior parte dei casi bisognerà solo cambiare la disposizione delle "patch". Ciò aggiunge molta flessibilità al progetto della rete. Un altro vantaggio del far coincidere l'armadio del centralino telefonico con quello della rete locale sta nel fatto che spesso nelle canalizzazioni dell'impianto telefonico sono già posati cavi in più, di riserva; sarà perciò possibile utilizzare per la rete dati quei cavi già pronti, senza spese di posa. Per il cablaggio strutturato esiste lo standard ISO/IEC 11801.
Se l'azienda è su più piani od in più edifici vicini, una rete dorsale (backbone), di solito cablata linearmente, collega fra loro i vari armadi di concentrazione.

1.4.2 Topologie di rete

La topologia di una rete é il modo con cui sono collegate logicamente le varie stazioni che la compongono. La topologia di una rete ha a che fare con il modo e con l'ordine con il quale le diverse stazioni colloquiano fra loro per trasportare le informazioni ma non con la disposizione fisica delle varie stazioni. In una rete con una certa topologia le varie stazioni possono essere collegate fisicamente con cablaggi diversi.

Bus
In una topologia a bus un messaggio trasmesso da una stazione può essere ricevuto contemporaneamente da tutte le altre stazioni della rete; solo la stazione interessata alla comunicazione utilizzerà il messaggio ricevuto. Il mezzo di trasmissione è condiviso da tutte le stazioni, per cui una sola stazione alla volta può trasmettere.

Anello
In una topologia ad anello ciascuna stazione è collegata logicamente con due sole altre stazioni; da una di esse riceve le informazioni, che elabora, eventualmente modifica, ed inoltra alla seconda delle stazioni a cui è collegata. La fila di stazioni, sistemate una di seguito all'altra, si richiude a formare un anello, per cui, se una stazione trasmette un messaggio alla stazione successiva, dopo un certo periodo di tempo se lo vedrà ritornare, di solito modificato, dalla stazione precedente. In linea di principio, con una topologia ad anello il mezzo di trasmissione può essere sia condiviso che commutato, anche se le realizzazioni pratiche più importanti usano un mezzo di trasmissione condiviso.

Indipendentemente dalla topologia adottata, la decisione su quale debba essere la stazione che può trasmettere è compito del "Medium Access Control" (MAC), che vedremo in seguito.

1.5 Protocolli

Tutte le comunicazioni richiedono una rete, ma essa non basta per assicurare una comunicazione efficace; è necessario che alla rete si aggiunga un insieme di regole che stabiliscano le modalità con cui lo scambio d'informazioni deve avere luogo.

Si dice protocollo un insieme di convenzioni e regole che governano uno scambio di informazioni attraverso una rete.

Protocolli stratificati (layered protocols) o "stack di protocolli"

1.5.1 Protocolli orientati al carattere (Byte oriented)

Protocolli che trasmettono le informazioni come "caratteri", costituiti ciascuno da un certo numero di bit (di solito 8). Per controllare la comunicazione questi protocolli fanno uso di caratteri speciali (codici di controllo). Esistono cioè particolari "caratteri" ai quali il protocollo assegna un significato particolare. Per esempio, nei protocolli che fanno uso del codice ASCII il numero 7 (bell) significa che il ricevente deve fare suonare una campanella. La presenza di caratteri di controllo del protocollo può essere riconosciuta sia da parte di dispositivi hardware che da software.
Con i protocolli Byte oriented è difficile spedire informazioni binarie, come per esempio un programma eseguibile. Infatti fra i dati binari da spedire possono essere compresi i numeri che codificano i codici di controllo. Per spedire questi numeri il protocollo dovrà realizzare "trucchi" piuttosto complicati, che diminuiscono l'efficienza della trasmissione (vedi, in seguito PPP).
Il codice ASCII, usato in molti protocolli Byte oriented, ha come caratteri speciali tutti i codici che vanno da 0 a 20h (32d). I caratteri di tastiera non sono caratteri di controllo e quindi la spedizione di file ASCII può essere effettuata senza problema anche con protocolli Byte oriented (vedi, in seguito, Xon - Xoff).

1.5.2 Protocolli orientati al bit (bit oriented)

Protocolli che trasmettono le informazioni come sequenze di bit. Ad ogni bit della sequenza può essere assegnato un significato particolare. Nei protocolli bit oriented le informazioni sono di solito organizzate in "**frame**" (termine che in Inglese significa "cornice" e che, in questo contesto, viene tradotto in Italiano come "trama").
La posizione del bit all'interno del frame ne determina il significato. Il protocollo specifica il significato di ogni singolo bit, raggruppando in "campi" (field) i bit che trasmettono la stessa informazione. Per spedire un'informazione di controllo basta che il protocollo preveda un campo apposito, opportunamente specificato, nel quale si potrà trasmettere l'informazione voluta. La gestione delle spedizioni e ricezioni è semplificata, dato che si possono spedire tutti i numeri, non limitati dalle informazioni di controllo. In questo caso si complica il software che deve estrarre le informazioni dal frame.
Il contenuto di alcuni campi di un frame può influenzare il formato del resto del frame. Ci possono essere cioè dei campi che indicano la presenza o meno di altri campi nel resto del frame. Il frame di un protocollo bit oriented può perciò essere variabile, sia come lunghezza che come forma e contenuto.
Nel gergo tecnico delle reti locali, il frame viene anche detto **PDU** (Protocol Data Unit) ed è l'unità minima di informazione che può essere trasmessa nella rete.
I protocolli più moderni sono in gran parte di tipo bit oriented.

1.6 Sicurezza delle comunicazioni

La sicurezza delle comunicazioni implica almeno tre aspetti: la certezza che il trasferimento è stato eseguito regolarmente, la certezza che le informazioni vadano solo a coloro che sono autorizzati a conoscerle, la sicurezza che le informazioni ricevute non andranno perdute nel corso del tempo. Possiamo indicare il primo tema come "affidabilità" dei collegamenti, il secondo come "riservatezza", sul terzo aspetto torneremo parlando di reti locali.

1.6.1 Affidabilità dei collegamenti

Protocolli a ricezione garantita
Alcuni protocolli effettuano controlli per assicurarsi che ogni messaggio spedito giunga regolarmente a destinazione, un po' come spedire una raccomandata con ricevuta di ritorno. Questi protocolli vengono detti "a ricezione garantita" (guaranteed delivery); in questi protocolli il ricevente deve aver modo di restituire al mittente la conferma (acknowledge) che il messaggio è stato ricevuto. Per essere a ricezione garantita non è necessario che il protocollo controlli se il messaggio arrivato è stato corrotto durante la trasmissione (anche se è molto improbabile che lo faccia). Esistono protocolli nei quali non è previsto che il destinatario, quando riceve un messaggio, non spedisca alcun messaggio di acknowledge al mittente. Questi protocolli non sono a ricezione garantita.

Protocolli a spedizione affidabile (senza errori)
Esistono protocolli che sono in grado di stabilire se i pacchetti spediti da mittente si sono persi nella rete o se sono giunti a destinazione "corrotti", cioè se contengono errori rispetto a quanto spedito. Se un protocollo è in grado di fare la verifica degli errori di trasmissione e se provvede anche ad assicurarne la automaticamente alla correzione viene detto protocolli "**a spedizione affidabile**" (reliable delivery).
Dei dati trasmessi da un protocollo affidabile "ci si può fidare" perché non sono "sbagliati", almeno nei limiti delle possibilità del protocollo di rilevare gli errori. Il fatto che i dati di un protocollo affidabile non siano sbagliati non significa che sono anche "sicuri", perché per poter essere considerati "sicuri" si deve essere certi:
1. dell'identità del trasmittente
2. del fatto che i dati trasmessi non siano intercettabili, o quantomeno non possano essere leggibili da chi riuscisse ad intercettarli lungo il loro percorso.

Per capire se ci sono stati errori nella trasmissione si usano le tecniche di rilevazione degli errori, trattate brevemente nel seguito.

Rilevazione e correzione degli errori
Se si fa una trasmissione di un file (si pensi ad un programma) è quasi sempre indispensabile che nemmeno un bit di quanto ricevuto sia diverso da ciò che è stato spedito. Poiché è impossibile avere la sicurezza che tutti i dati spediti non vengano corrotti durante la trasmissione è necessario approntare delle tecniche per riconoscere alla ricezione se i dati non sono affidabili. Qualora si accertasse la corruzione di alcuni bit, sarà compito del ricevente effettuare la correzione, se il metodo utilizzato lo permette, oppure chiedere alla sorgente la ritrasmissione della parte di messaggio interessata dall'errore.
Per poter rilevare gli errori, ed eventualmente per correggerli senza ritrasmissione, è necessario che sia spedita più informazione di quella strettamente indispensabile. I codici che si adotteranno avranno cioè delle "ridondanze".
Si parla di rilevazione degli errori quando è possibile identificare la presenza di errori nella trasmissione o nella memorizzazione di un'informazione. Si parla di correzione degli errori quando alcuni degli errori rilevati possono essere anche corretti utilizzando gli stessi bit di ridondanza trasmessi con il segnale, cioè senza chiedere la ritrasmissione della parte errata del messaggio.
Il trasmittente fa uso di un algoritmo che, utilizzando le informazioni da trasmettere, gli permette di calcolare dei dati ridondanti. Successivamente esso trasmette sia le informazioni, sia i dati ridondanti. Il ricevente usa le informazioni ricevute per calcolare indipendentemente il valore dei dati ridondanti, con lo stesso algoritmo che aveva usato chi trasmetteva. Se il calcolo fatto dal ricevente è diverso dai valori ridondanti che ha ricevuto, si ha la certezza che i dati si sono corrotti nella trasmissione. Se invece essi sono uguali, c'è una probabilità alta che i dati siano stati ricevuti correttamente, ma non c'è mai la certezza assoluta. La probabilità di considerare giusti dei dati sbagliati è bassa, utilizzando gli algoritmi migliori, e può essere abbassata a piacimento aumentando la quantità di informazione ridondante. Per valutare la bontà degli algoritmi per la rilevazione degli errori se ne definisce la sensibilità e l'efficienza.

Sensibilità di un codice ridondante
La sensibilità di un codice ridondante è la sua bontà nello scovare gli errori, cioè la percentuale di errori che esso è in grado di rilevare, rispetto alla quantità di informazioni trasmesse.

Efficienza di un codice ridondante
L'efficienza di un codice ridondante è la sua capacità di usare bene i bit di ridondanza.
Essa è legata alla misura della degradazione delle prestazioni che il codice introduce, cioè al numero di bit di ridondanza aggiunto ai dati, all'aumento della memoria richiesta e/o al rallentamento delle operazioni che l'utilizzo di quel codice impone. Tanto minori saranno questi fattori, rispetto al volume di informazioni trasportate, tanto maggiore sarà l'efficienza.

Come in tutte le cose della vita, di solito si deve cercare un compromesso, in questo caso fra sensibilità ed efficienza, anche se alcuni metodi sono migliori di altri in entrambi i campi. Facciamo un esempio, un po' assurdo. Per avere sensibilità al 100% nel rilevamento dell'errore si potrebbe fare il confronto bit per bit dell'informazione ricevuta con una copia "campione" che si sa che non contiene errori. In questo caso però l'efficienza sarà nulla, perché, avendo già ricevuto una copia "certa" dell'informazione, di fatto con la seconda copia non spediamo informazioni nuove.
Una trasmissione così potrebbe essere interessante in sede di test dell'apparato di comunicazione. Infatti in quel caso l'informazione che conta non è "cosa c'è scritto nel messaggio" ma piuttosto se: "quello che ho spedito è arrivato giusto".

Solo allora ha senso spedire una sequenza di bit che sono già noti alla destinazione, perché ogni bit che arriverà diverso da come era previsto sarà senz'altro un errore.

Controllo di parità

Il modo più semplice per rilevare errori di trasmissione è il controllo di parità.
Volendo trasmettere un blocco di bit utilizzando il controllo di parità, si deve aggiungere un bit di ridondanza ai normali bit del blocco. Il valore di questo bit è basato sul numero di bit del blocco che sono a 1. Il bit in più viene detto "bit di parità". La parità viene detta "pari" (even) se si fa in modo che sia pari il numero totale dei bit a uno che vengono trasmessi (compreso il bit di parità). Se quel numero è dispari, la parità viene detta dispari (odd).
Il metodo del controllo di parità non è molto sensibile, p.es. se durante la trasmissione di un blocco vengono corrotti un numero pari di bit l'errore non verrà rilevato. Se la probabilità dell'errore fosse uniforme fra i bit, questo significherebbe una sensibilità del 50% (ci azzecca una volta su due!). A questo proposito si deve considerare che nelle telecomunicazioni normalmente vengono commessi pochissimi errori, ma quando qualche disturbo occasionale ne fa commettere uno, la probabilità che se ne commetta un altro subito dopo è abbastanza alta (si dice che gli errori di trasmissione sono di tipo "burst": non ci sono quasi mai ma quando cominciano sono "importanti" e durano molto, per i tempi dei computer di oggi). L'efficienza del controllo di parità è perciò così bassa che non viene più usato nelle telecomunicazioni dai tempi delle telescriventi. Esso è stato sostituito in tutte le applicazioni "serie" da metodi più sensibili e che utilizzano i bit di ridondanza in modo molto più efficiente.
In altri casi il bit di parità viene usato ancora. Ciò accade per esempio nella memoria RAM di molti PC. Essa è indirizzata a Byte, ma è composta di 9 bit per locazione, con un bit di parità in più. Ciascuno dei bit di ogni locazione della RAM di solito si trova fisicamente in un chip diverso. Dato che la probabilità che due chip di memoria si guastino contemporaneamente è praticamente nulla, l'utilizzazione del bit di parità ha senso ed è praticata.

Checksum

Un checksum ("check-sum", somma di controllo) è la somma dei numeri binari che costituiscono un blocco di dati, fatta buttando via tutti i riporti. Se il numero di bit della somma è limitato a 8 o a 16, si dice che il checksum è a modulo 256 o 65536, cioè oltre a 256 o 65535 si ricomincia da zero. Blocchi di dati diversi possono perciò dare lo stesso checksum, anche se ciò è improbabile per dati corrotti da disturbi "burst". Un checksum non è molto sensibile e lo è sempre meno quanto più lungo è il blocco di dati. Inoltre è completamente cieco alle sequenze di numeri zero, od all'inversione dell'ordine dei numeri del blocco. I checksum si usano ancora per blocchi fino a 255 Byte.

La parola checksum viene anche usata in modo più generale, per intendere un qualsiasi algoritmo che permetta di sapere se un certo insieme di numeri è "cambiato". Per esempio, esistono dei programmi che calcolano un "checksum" dei dati in ogni cartella di un hard disk, o di ogni file eseguibile, in modo da verificare se un virus li ha modificati. In questo caso non si tratta di checksum in senso stretto ma di algoritmi molto più efficienti e bisognerebbe parlare di "hash", piuttosto che di checksum.

"Checksum" MD5

Un esempio di algoritmo di "checksum" molto usato è l'algoritmo MD5.
Si tratta di un algoritmo di pubblico dominio, implementato in tutti i linguaggi ed in tutte le piattaforme, che fornisce la "firma" di un gruppo di dati. Essa è un hash dei bit dei dati calcolato in modo che il cambio anche di un singolo bit dei dati faccia cambiare, con altissima probabilità, anche il valore della "firma".
L'algoritmo MD5 garantisce anche che dalla firma (hash) sia praticamente impossibile calcolare il valore originario.
MD5 può servire per accertarsi che il contenuto di un file non sia stato modificato, per esempio nella sua trasmissione in rete. La cosa può funzionare così:

1. Il mittente calcola la firma MD5 del file da spedire
2. Il mittente trasmette il file ed anche la firma
3. Il ricevente riceve file e firma MD5
4. Il ricevente calcola la firma MD5 con i dati del file che ha ricevuto
5. Se la firma che ha ricevuto è uguale a quella calcolata si può essere praticamente certi che il file non è stato modificato, se è diversa si è senz'altro sicuri che il file è stato modificato.

Un'altra applicazione dell'algoritmo MD5 è la memorizzazione delle password. Quando un utente cambia la sua password essa viene memorizzata nel sistema solo dopo essere passata attraverso un calcolo MD5. In questo modo nessuno, neppure l'amministratore del sistema, è in grado di leggere la password. Infatti dall'hash MD5 non è possibile risalire alla password.
Successivamente, quando l'utente vuole entrare nel sistema, digita la sua password, che viene nuovamente passata in MD5 e confrontata con l'hash MD5 memorizzato. Se i due hash MD5 sono uguali l'utente è autenticato, altrimenti viene respinto.

CRC Codici a ridondanza ciclica (Cyclic Redundancy Code)

Un CRC è un numero ottenuto come risultato di un calcolo, effettuato sopra la sequenza di Byte che si deve trasmettere, con un algoritmo che fa uso di polinomi binari. La sequenza di Byte viene considerata come un grande numero binario, che viene diviso per un polinomio fatto in modo che il resto della divisione sia "il più unico possibile".
Il resto della divisione è il CRC. Gli algoritmi CRC usano la divisione in modulo 2, che necessita solo di operazioni di XOR e di shift, sono perciò veloci da eseguire e facili da realizzare anche in hardware.

Il CRC viene aggiunto alla sequenza di bit del segnale e spedito con essa. Il destinatario riceverà sia i bit di informazione, sia il valore del CRC. Con i bit di informazione e con lo stesso algoritmo usato dal sorgente, il destinatario calcolerà un suo CRC, che dovrà essere identico a quello ricevuto. Se il CRC sarà diverso, si sarà sicuri che i dati si sono corrotti durante la trasmissione. Viceversa, l'identità dei due CRC non darà la certezza che i dati non sono corrotti. Peraltro la probabilità di avere un CRC identico anche nel caso di dati corrotti è molto bassa, e, cambiando algoritmo, può essere bassa quanto si vuole. I CRC sono molto efficienti nell'utilizzare i bit in più (ridondanti) che vengono trasmessi. P.es. il CRC CCITT (ITU), usato nei fax, ha queste sensibilità: è in grado di rilevare il 100% degli errori di un solo bit, il 100% degli errori di due bit separati da meno di 2^{16} bit, il 100% degli errori adiacenti ("burst") fino a 16 bit di seguito, il 99,998% di tutti gli altri errori (da J.Hirst (vedi in "riferimenti")).

Correzione dei dati
La correzione dei dati rovinati dalla trasmissione si ottiene in due modi: con la ritrasmissione o con l'uso di particolari codici ridondanti che permettono anche la correzione. Il primo dei due casi si presenta più di frequente, se che il costo della ritrasmissione non è molto alto. Qualora invece quel costo fosse proibitivo, come nel caso della comunicazione con satelliti lontani, si utilizzano codici a correzione d'errore, che usano gli stessi bit di ridondanza per effettuare la correzione e chiedono la ritrasmissione solo molto raramente (o mai, tenendo i dati "rotti" se non si riescono ad aggiustare).

1.6.2 Riservatezza delle comunicazioni

La riservatezza di una comunicazione è la sicurezza che le informazioni non siano lette da chi non è autorizzato a farlo. Per evitare questa evenienza bisogna tener conto di almeno due aspetti:
1. sicurezza che l'accesso fisico alle informazioni sia riservato solo agli autorizzati (autenticazione dell'utente).
2. sicurezza che le informazioni possano essere comprese solo da chi è autorizzato (crittografia).

Autenticazione dell'utente

Password

"Certificati" e firme elettroniche

Altri metodi

Identificazione biometrica

Crittografia
Qualora le informazioni delle quali si vuole assicurare la riservatezza cadessero comunque in mano a persone non autorizzate, ci si può cautelare con tecniche che fanno in modo che le informazioni che vengono trasmesse non siano "riconoscibili" ad altri che non siano il destinatario. Per far questo esse vengono modificate con un algoritmo. Le tecniche che si usano in questi casi vanno sotto il nome di "crittografia".
Un algoritmo di crittografia trasforma l'informazione che si deve trasmettere in modo che non sia più riconoscibile, a meno di non utilizzare una "chiave" segreta, nota a tutti i destinatari.

Crittografia a chiave privata

Crittografia a chiave pubblica

1.6.3 Compressione dei dati

In un flusso di informazioni ci possono essere ripetizioni ed altre ridondanze. Qualora esse vengano tolte, si avrà la possibilità di trasmettere meno dati e perciò di avere comunicazioni più efficienti.

Algoritmi senza perdita d'informazione (lossless)
RLE
.GIF
.ZIP
Algoritmi con perdita d'informazione (lossy)
.JPG

Crittografia e compressione insieme