

**STAGE FORMATIVO AI LNF
PER GLI STUDENTI
DELL'E.FERMI EDIZIONE 2001**

PERIODO: 11 GIUGNO - 6 LUGLIO

TOTALE ORE: 160

RELAZIONE DI SISTEMI E PROGRAMMAZIONE



**STUDENTI:
CAPRI FRANCESCA
CAPRI SARA**

**TUTOR:
SFILIGOI IGOR
FORTUGNO FABIO**

INTRODUZIONE

I Laboratori Nazionali di Frascati (LNF) sono il più grande Laboratorio dell'Istituto Nazionale di Fisica Nucleare (INFN), l'ente che promuove, coordina e finanzia, la ricerca nel campo della fisica subnucleare e nucleare.

L'INFN affida ai LNF il compito di costruire macchine acceleratrici di particelle in grado di sondare i costituenti del nucleo e di produrre in laboratorio nuove forme di materia e da allora i LNF sono stati protagonisti in tutti i settori di ricerca propri dell'INFN.

Questo stage rappresenta il lavoro portato avanti durante tutta la durata dello stage.

Questo stage è diviso in due parti

- *una parte riguarda lo studio riferito all'architettura di reti e sistemi;*
- *l'altra parte riguarda la programmazione e quindi la realizzazione di codici che consentono la gestione di una tabella(database).*

ARCHITETTURA DI RETI E SISTEMI

1 L'EVOLUZIONE DELLA RETE

Sin dalla meta' degli anni '60 l'evoluzione tecnologica consentiva, in un centro di calcolo, di collegare ad un grosso calcolatore centrale (mainframe-terminali) delle postazioni di lavoro che permettevano anche ad utenti lontani di accedere alle risorse del mainframe. Inizialmente, queste erano le reti in uso.

Le postazioni di lavoro erano costituite da terminali privi di risorse proprie, senza grosse capacita' di elaborazione e controllati direttamente dal calcolatore centrale. Queste reti vengono chiamate reti da terminali.

L'impossibilita' di eseguire lavoro autonomo e il controllo centralizzato della comunicazione rendeva il tutto piu' lento, questo limite e' stato risolto introducendo un calcolatore in cui vi e' un grande numero di elaboratori autonomi e interconnessi fra loro, cioe' non deve esserci una relazione tipo master/slave e inoltre devono essere in grado di scambiare informazioni con un opportuno mezzo fisico.

Un sistema di calcolo di questo tipo e' detto rete di elaboratori (computer network) che consente ad ogni utente che ne fa parte di avere accesso a servizi, siti, sistemi e persone.

Rete di elaboratori non e' sinonimo di sistema distribuito, perche' in quest'ultimo l'esistenza di piu' elaboratori e' invisibile all'utente, che ha l'impressione di avere a che fare con un unico sistema di calcolo. Mentre in una rete di elaboratori, l'utente e' a conoscenza dell'esistenza di piu' elaboratori.

Gli scopi principali di una rete sono :

-CONDIVISIONE DI RISORSE: rendere tutti i programmi, i dati e le apparecchiature disponibili a chiunque nella rete, indipendentemente dai punti in cui sono situati la risorsa e l'utente;

-ALTA AFFIDABILITA': mettere in rete sorgenti alternative delle risorse potendole utilizzare in qualsiasi momento;

-FLESSIBILITA' DEL SISTEMA: possibilita' di aumentare le prestazioni del sistema aggiungendo, entro certi limiti, altri processori con l'aumento del carico di lavoro;

-DIMINUIZIONE DEI COSTI: una rete di pc costa meno di un mainframe;

- COMUNICAZIONE FRA PERSONE*:possibilita' di tenere un rapporto,scambiare informazioni e inviarsi messaggi con utenti che vivono a grandi distanze;
- ACCESSO AD INFORMAZIONI REMOTE*:accesso ad informazioni che si trovano su un' altra macchina ;
- DIVERTIMENTO*.

2 TIPI DI RETE

Una rete e' formata di nodi e collegamenti tra I nodi.

Le reti possono essere suddivise in tre grandi famiglie con architettura,caratteristiche e risorse fisiche diverse:

- LAN(LOCAL AREA NETWORK)*:reti locali composte anche solo da due personal computer;
- WAN(WIDE AREA NETWORK)*:reti geografiche,molto complesse con migliaia di computer,distribuiti su un vasto territorio;
- MAN(metropolitan area network)*:rete metropolitana;

2.1 RETI LAN

Le LAN vengono utilizzate su distanze brevi tra luoghi in uno stesso edificio(ufficio)o edifici adiacenti.Il loro impiego richiede che esse vengano installate all'interno di una superficie privata senza alcun attraversamento di suolo pubblico, utilizzando apparecchiature e canali privati.

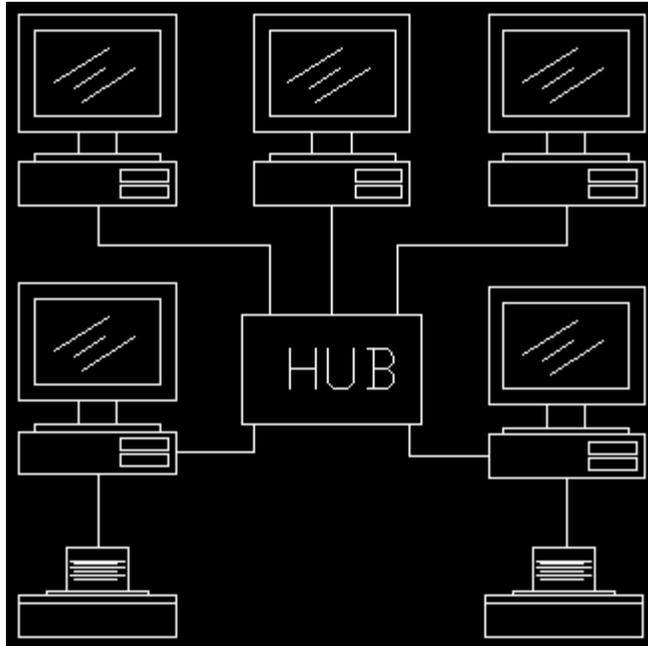
Una caratteristica principale delle LAN e' che utilizzano un unico mezzo trasmissivo ad alta velocita' condiviso da tutti i sistemi collegati.

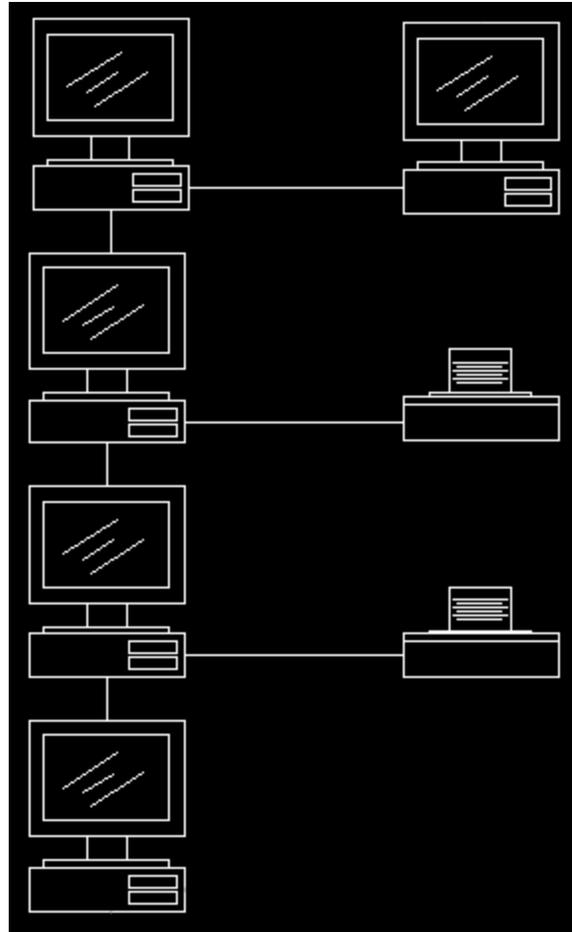
Le LAN sono in generale reti broadcast:un utente trasmette e tutti gli altri ricevono.Sono possibili diverse topologie, le piu' diffuse sono il bus ed il ring:

- *A TOPOLOGIA BUS*: in ogni istante solo un elaboratore puo' trasmettere,gl altri devono aspettare;e' necessario, un meccanismo di arbitraggio per risolvere i conflitti quando due o piu' elaboratori vogliono trasmettere contemporaneamente;lo standard IEEE chiamato Ethernet, che e' una rete basata sul bus dove gli elaboratori trasmettono quando vogliono,se c'e' una collisione aspettano un certo tempo e poi riprovano.
- *A TOPOLOGIA RING*:in un ring ogni bit percorre l'anello in un tempo inferiore a quello di trasmissione di un pacchetto;anche qui e' necessario un meccanismo di arbitraggio basato sul processo di un gettone(token)che abilita' alla trasmissione;lo standard IEEE,derivante dalla rete IBM token

ring. Nelle LAN e' necessaria la presenza di indirizzi per stabilire il reale destinatario e il mittente della trasmissione.

ESEMPI DI RETI LAN:





2.2 RETI WAN

Le WAN vengono utilizzate su aree geografiche molto estese si parla di comunicazioni anche di migliaia di chilometri. Il tipo di canale di trasmissione utilizzato, per motivi di costo, diventa rilevante per cui si dovranno effettuare delle scelte che garantiscono una trasmissione affidabile.

Al problema delle lunghe distanze e' collegata la presenza dei dispositivi che rigenerano e ampliano il segnale che sta viaggiando lungo la rete:

-un insieme di elaboratori sui quali girano i programmi usati dagli utenti;

-un subnet che connette gli elaboratori fra loro. Il suo compito e' di trasportare messaggi da un elaboratore ad un altro.

La subnet consiste di due componenti:

~ linee di trasmissione;

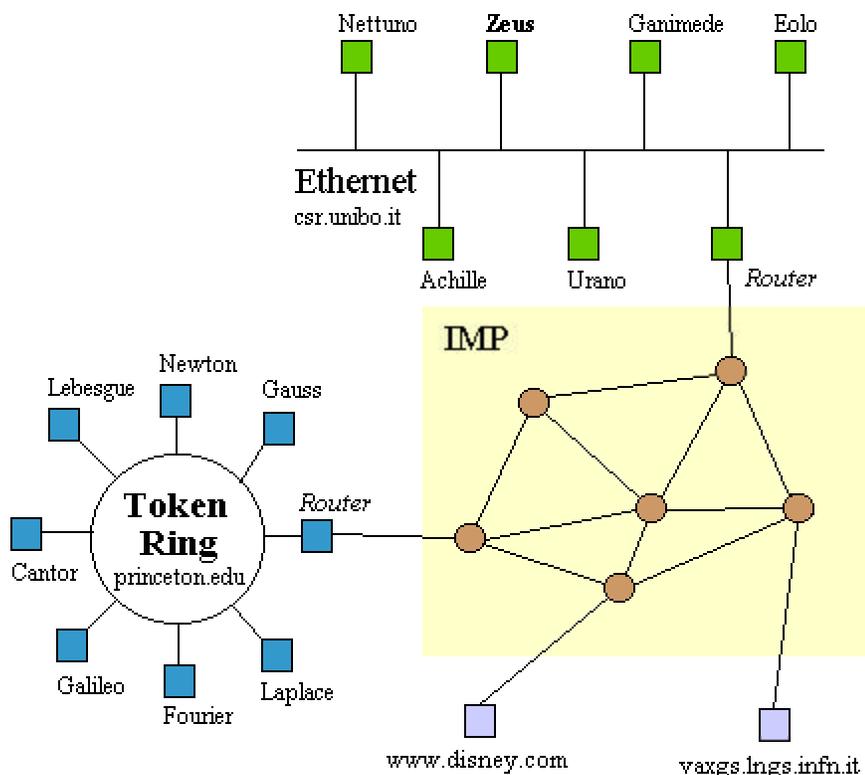
~elementi di commutazione (switching element).

Un esempio di rete WAN e' l'architettura SNA ,dove la rete ha la forma di un quadrilatero e il pacchetto prima di raggiungere il destinatario,si puo' dividere lungo le linee che portano al nodo desiderato per poi ricompattarsi quando esso viene raggiunto.Questo lavoro assicura che se una linea dovesse rompersi la connessione non andra' in errore.

Una WAN e' utilizzata per connettere piu' LAN fra loro, molto spesso LAN distinte vengono interconnesse con dispositivi detti BRIDGE (che spiegheremo in seguito) per molte ragioni:

- due LAN differenti,ad esempio Ethernet ed una token ring,vogliono dialogare fra di loro;
- si vuole una LAN la cui lunghezza superi i limiti massimi consentiti,ad esempio 12,5 km per Etherne;
- si desidera suddividere una LAN in molteplici LAN.

ESEMPIO DI RETE WAN:



2.3 RETI MAN

Le MAN si estendono su distanze medio-piccole e utilizzano principi e tecnologie caratteristiche delle reti geografiche pure estendendosi in un'area limitata.

3 MEZZI DI COLLEGAMENTO

Si è detto che una rete consente di tenere collegati tra loro molti nodi. La scelta dei mezzi di collegamento è una delle più importanti nel costruire una rete, in quanto deve prevedere la possibilità di future espansioni.

Al giorno d'oggi i più comuni mezzi di comunicazione sono:

3.1 DOPPIO INTRECCIATO

Questo cavo è il più anziano e tuttora il più usato nelle reti locali vista la sua economicità e la sua elevata velocità di trasmissione (fino a qualche km mantiene velocità di diversi Mb/s).

Consiste di una coppia di conduttori in rame intrecciati l'uno con l'altro, in forma elicoidale, per minimizzare le interferenze tra i fili.

È usato in particolare per collegare terminali del sistema telefonico, per trasmissioni analogiche e digitali con una larghezza di banda che dipende dalla sezione del filo e dalla distanza da attraversare.

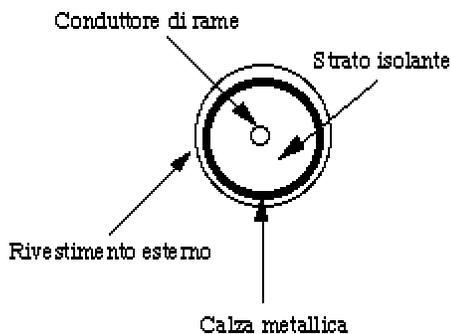
Esistono due diversi tipi di doppi:

- *Quello di Categoria 3, non schermato, usato nelle linee telefoniche tradizionali;*
- *Quello di Categoria 5, più stabile sulle lunghe distanze (perché più schermato) e più idoneo alla trasmissione di dati ad alta velocità.*



3.2 CAVO COASSIALE

Questo cavo e' molto simile a quella delle comuni antenne televisive, ed e' costituito da un conduttore centrale in rame circondato da uno strato isolante all'esterno detto calza metallica.



Offre un miglior isolamento rispetto al doppino e consente di trasmettere dati con buona velocita' su lunghe distanze (massimo 180 km).

Al giorno d'oggi questa soluzione e' pero' in disuso sia nelle reti locali (dove si preferisce il doppino, piu' economico e veloce), sia sulle lunghe distanze dove sta prendendo piede la fibra ottica .



3.3 FIBRA OTTICA

Le fibre ottiche sono uno dei mezzi piu' recenti che stanno rivoluzionando le telecomunicazioni.

Un sistema di trasmissione ottica ha 3 componenti:

~mezzo di trasmissione:

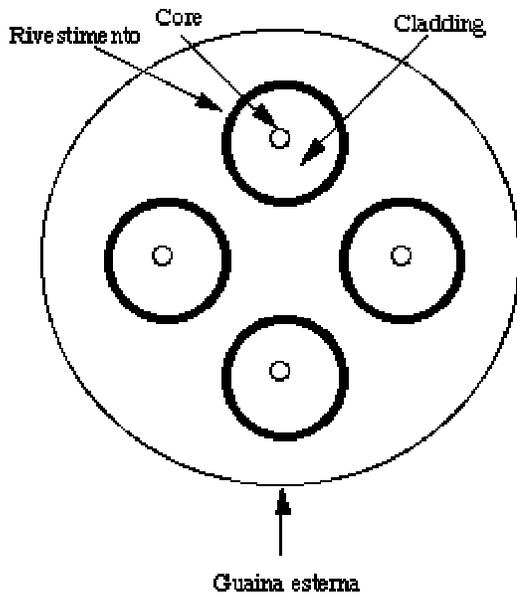
E' costituito da un sottilissimo cilindro centrale in vetro (core) con indice di rifrazione n_1 , circondato da uno strato esterno di vetro, chiamato mantello (cladding), avente un diverso indice di rifrazione n_2 e da una guaina protettiva, il tutto e' raggruppato in una guaina contenitrice esterna;

~sorgente luminosa;

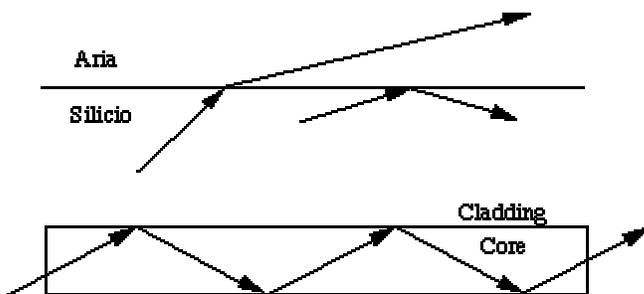
Puo' essere un semplice LED o comunque un dispositivo in grado di generare impulsi luminosi quando viene attraversato da corrente elettrica;

~rilevatore del segnale:

fotodiodo che emette un impulso elettrico quando viene attraversato da radiazioni luminose;



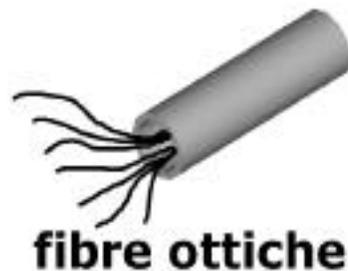
Le fibre ottiche sfruttano il principio della deviazione che un raggio di luce subisce quando attraversa il confine fra due materiali diversi. La deviazione dipende dagli indici di rifrazione dei due materiali. La presenza di un impulso di luce puo' segnalare un bit a 1, mentre la sua assenza indicherebbe un bit a 0. Il raggio di luce quando entra nella fibra genera un'onda riflessa ed un'onda rifratta (segnale perso). Studiando la geometria oculare si verifica che esiste un angolo limite chiamato φ_m dove non esiste il segnale di rifrazione ma solo quello di riflessione e quindi il raggio entrera' con quest'angolo in modo da non perdere parte dell'onda.



Le fibre ottiche si dividono in due categorie:

- *MULTIMODALI*: raggi diversi possono colpire la superficie con diversi angoli (detti *mode*) proseguendo quindi con diversi cammini. Il diametro del core è di 50 micron, come quello di un capello;
- *MONOMODALI*: sono così sottili (il diametro del core è 8-10 micron) che si comportano come una guida d'onda: la luce avanza in modo rettilineo, senza rimbalzare. Sono più costose ma reggono distanze più lunghe (fino a 30 km).

Oggi le fibre ottiche sono usate principalmente per i grandi server visto la loro incredibile velocità di trasporto dati (arrivano a 1.5Gb/s e oltre) e la loro capacità di isolamento ma il costo è ancora molto elevato.



PUNTI DI AGGREGAZIONE

Gli SWITCH sono dei punti di aggregazione (elementi di commutazione) in cui convergono tutti i cavi che collegano i computer in rete. Essi svolgono una funzione importante in quanto regolano e gestiscono il traffico dei dati sulla rete.

Per prima cosa offrono agli utenti una velocità di trasmissione molto alta, se l'efficienza di trasmissione massima è di 10Mb/s, ogni computer ha i propri 10Mb da sfruttare.

Inoltre gli SWITCH inviano i pacchetti di dati solo alla porta specificata dal destinatario in quanto riesce a stabilire una connessione temporanea tra la sorgente e la destinazione chiudendola alla fine della connessione

lasciando sempre libera la linea per gli altri computer. E' come se si creasse, ai nodi (nodo=dispositivo indirizzabile individualmente connesso ad una rete) che devono comunicare, una propria piccola rete evitando cosi' l'eventuali collisioni che si potrebbero creare fra i nodi appartenenti alla rete .

Un termine usato per definire questi elementi di commutazione e' il ROUTER che e' uno smistatore di traffico che riceve un pacchetto da una linea in ingresso, lo memorizza per intero in un buffer interno e appena la linea e' libera, instrada il pacchetto su essa. I router nelle reti si occupano del traffico verso l'esterno.

Questi elementi sono anche molto intelligenti perche' leggono l'indirizzo completo per determinare il punto successivo a cui inviare il pacchetto di dati. Poi grazie ad una tabella di routine, fanno arrivare I dati a destinazione attraverso percorsi piu' veloci, infatti se trovano una strada occupata ne cercano una alternativa senza bloccare il traffico della rete.

Il BRIDGE(cpu+ memoria)quindi consente di collegare due (o piu') reti con la stessa tipologia in modo d' averne una sola consentendo conseguentemente la possibilita' di piu' nodi di colloquiare tra loro come se appartenessero ad un'unica rete.

La caratteristica che differenzia un BRIDGE da un ROUTER e' che i primi agiscono sulle buste di livello due e i secondi su quelli di livello tre.

5 TOPOLOGIE DI RETE

Tutti gli elementi che compongono una rete di calcolatori possono essere connessi fisicamente tra loro in moltissimi modi secondo le necessita'.

Le piu' comuni topologie di rete sono:

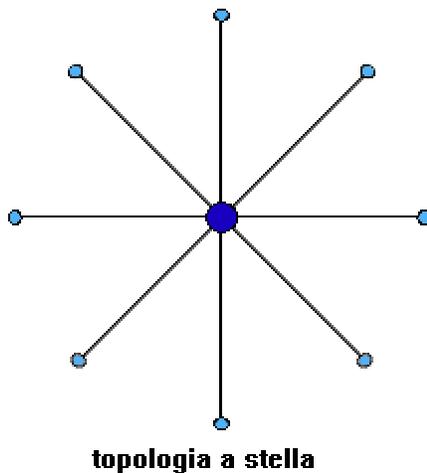
- A STELLA
- AD ANELLO(TOKEN-RING)
- AD UN ALBERO O GERARCHICA
- A BUS

5.1 TOPOLOGIA A STELLA:

La topologia a stella e' la topologia piu' antica usata nella trasmissione dati .

Questa tipologia prevede un nodo centrale al quale sono direttamente connessi tutti gli altri.

Da questo punto comune si diramano tutti i cavi che collegano le stazioni. Considerando il tipo di cavo utilizzato,il costo totale e' all'incirca uguale a quello delle altre topologie.



5.2 TOPOLOGIA AD ANELLO:

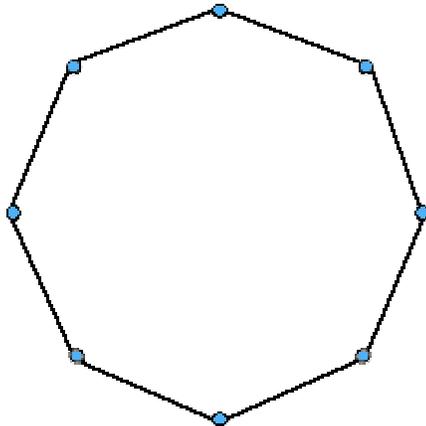
Nella topologia ad anello ogni nodo e' fisicamente connesso con altri due in modo da costruire una struttura circolare.

L'anello puo' essere bidirezionale o unidirezionale.

Per quanto riguarda un anello unidirezionale,i costi iniziali restano bassi,i costi di comunicazione subiscono un incremento e la sicurezza dei dati resta inalterata.In questo caso,il messaggio viaggia dalla sorgente verso il nodo destinatario seguendo l'anello:non esiste un nodo prevalente sugli altri,per cui non si crea un concentrazione del lavoro in un unico nodo.

La rete unidirezionale e' molto sensibile ai guasti, se si verifica un'interruzione lungo un canale il nodo in questione resta isolato e gli altri non potranno effettuare alcune connessioni ad altri nodi.

I nodi di una rete bidirezionale hanno la possibilita' di scegliere il percorso da far eseguire al proprio messaggio per raggiungere il nodo desiderato: la scelta e' limitata a due sole possibilita' e avviene principalmente in base al numero di nodi che vanno attraversati.



topologia ad anello

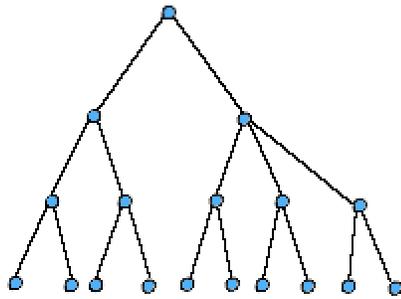
5.3 TOPOLOGIA AD ALBERO GERARCHICO:

Una topologia gerarchica e' rappresentata da una struttura ramificata dei nodi che la compongono: ogni nodo (esclusa la radice) e' figlio di un nodo padre e avra' eventualmente dei nodi fratelli e dei nodi figli.

i collegamenti tra nodi figli e padre sono molto veloci e diretti, mentre restano lenti e laboriosi i collegamenti su lunghe distanze, con un affollamento di richieste man mano che ci si avvicina al nodo radice della gerarchia.

Il vantaggio maggiore e' la semplicita' della gestione.

Lo svantaggio e' che se il sistema principale, e' sovraccarico di lavoro, puo' provocare il rallentamento dei servizi per tutti gli utenti.



topologia ad albero

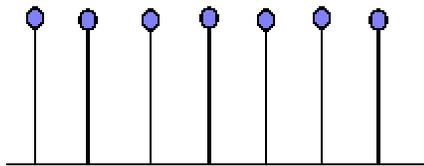
5.4 TOPOLOGIA A BUS:

Questa struttura e' molto semplice utilizza un unico tratto di cavo al quale sono collegate le stazioni rete.

Tutte le stazioni condividono questa cavo e i dati trasmessi da una qualsiasi di esse vengono ricevuti da tutte le stazioni collegate a tale cavo.

Le estremita' del segmento di cavo sono terminate da dei punti di terminazione .

Data la semplicita' di questa topologia il costo dell'implementazione e' basso. Il cavo unico puo' comportare un problema di notevole entita', in quanto nel caso in cui si guasti, nessuna delle stazioni potra' trasmettere. Un altro grande problema e' dato dal fatto che il cavo, essendo di lunghezza fissa, non consente un inserimento illimitato di nodi. Questo problema e' stato risolto inserendo nella rete il dispositivo BRIDGE.



6 PROTOCOLLI DI COMUNICAZIONE

I componenti che costituiscono la rete per poter comunicare tra loro e, se necessario, verso l'esterno hanno bisogno di un linguaggio di comunicazione standardizzato.

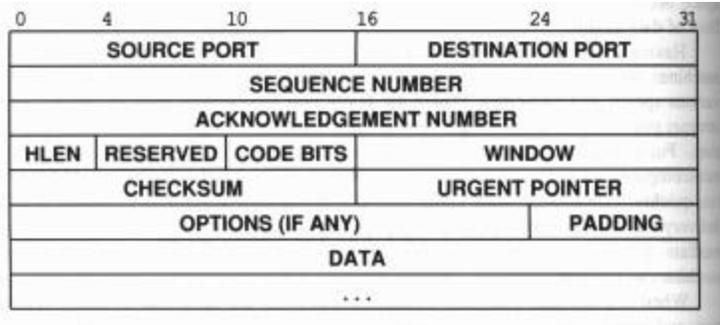
6.1 PROTOCOLLO TCP/IP

Il protocollo piu' usato e' il TCP/IP. questo protocollo ha principalmente la funzione di trasportare i dati attraverso la rete e compie questa operazione in due momenti distinti.

Inizialmente il protocollo TCP (transmission control protocol) suddivide i dati in pacchetti, indica la modalita del trasporto e di conseguenza anche il modo in cui dovra manipolare i dati ricevuti. Successivamente si assegna ad ogni pacchetto l'indirizzo IP del mittente e quello del destinatario in modo da indicargli gli estremi del percorso che dovra compiere.

Prima di trasmettere il pacchetto di dati successivo ci si assicura che il precedente sia arrivato.

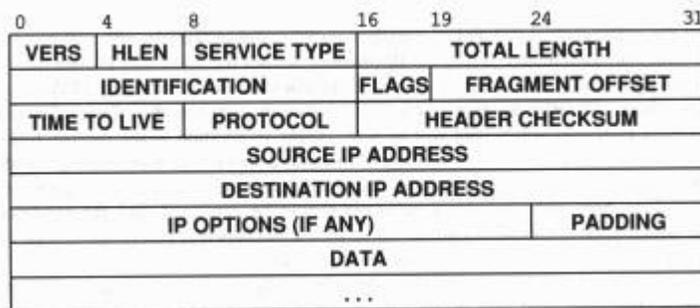
Con il TCP/IP si assegna ad ogni host (computer centrale) un indirizzo binario a 32 bit. Per rendere questi indirizzi piu' comprensibili essi sono suddivisi in quattro gruppi di bits con i rispettivi valori scritti in decimale e separati da punti.



L'unita' di trasferimento base del TCP/IP e' il datagramma IP.

Questo contiene:

- *L'indirizzo destinazione;*
- *L'indirizzo di partenza;*
- *Un type field che identifica il contenuto del datagramma;*
- *Un checksum field che assicura l'integrita' dei valori contenuti in esso;*
- *Il time to live che specifica quanto tempo puo' sopravvivere un datagramma IP nella rete .*



Il complesso del protocollo TCP/IP puo' supportare un semplice programma applicativo per l'accesso a terminale remoto a carattere TELNET esso permette all'utente di un sito di stabilire una connessione TELNET con un server di un altro sito. TELNET poi passa tutto cio' che e' digitato sulla tastiera dell'utente direttamente al computer remoto, come se fosse stato digitato sulla stessa macchina remota; inoltre TELNET visualizza sullo schermo dell'utente l'output testuale della macchina remota.

INDIRIZZI E NOMI CHE CARATTERIZZANO I CALCOLATORI

Ogni calcolatore e' caratterizzato da un indirizzo detto MAC ADDRESS che viene consegnato all'acquirente dalla casa che fa la macchina nel momento in cui acquista il calcolatore. Infatti quello detto in precedenza e cioe' quando si parlava del datagramma IP che conteneva l'indirizzo sorgente e l'indirizzo destinazione, consiste nel MAC ADDRESS della macchina che invia i dati e di quella che li riceve .

Per ogni MAC ADDRESS e' assegnato un nome simbolico che risulta essere facilmente memorizzabile e identificabile a cui corrispondera' un indirizzo IP, attraverso il protocollo DNS che consente la conversione del numero IP in un indirizzo alfanumerico.

7 TECNICA DI TRASMISSIONE DATI

Adesso ci occuperemo della trasmissione vera e propria di informazioni tra 2 nodi.

Quando un nodo deve comunicare con un altro nodo invia attraverso il protocollo TCP/IP il pacchetto di informazioni al socket (porta) del nodo destinato a ricevere l'informazione.

Questa porta fa un operazione di fork cioe' operazione di duplicazione, una rimane ad ascoltare l'arrivo di altri eventuali segnali su quella porta ,l'altra invia al nodo mittente il nuovo indirizzo a cui risponde e da quel momento potranno colloquiare soltanto su quella porta.

8 ARCHITETTURA ISO-OSI

L'organizzazione internazionale degli standard (ISO International Standard Organization) senti' la necessita' di proporre una serie di standard per le reti di calcolatori e avvio' cosi' il progetto OSI (Open System Interconnection) ossia un modello di riferimento per interconnessione di sistemi aperti.

Gli scopi principali di questo modello sono:

-fornire una base comune su cui sviluppare standard per l'interconnessione di sistemi informatici;

-fornire un modello rispetto a cui confrontare le architetture di rete.

Il modello OSI e' diviso in sette livelli:

8.1 LIVELLO FISICO

La funzione principale di questo livello e' di garantire che i bit spediti vengano ricevuti correttamente dal destinatario.

8.2 LIVELLO DI LINEA (DATA LINK)

I compiti principali di questo livello sono:

-organizza i dati in strutture di bit chiamate frame, che vengono passate a livello di rete che provvedera' a prepararle al viaggio lungo la rete;

-inserisce delle sequenze predefinite di bit di controllo;

-controlla e regola i tempi di comunicazione tra trasmettitore e ricevitore e il flusso di dati.

8.3 LIVELLO DI RETE (NETWORK)

Le funzioni fondamentali di questo livello sono:

-determinazione del modo in cui i frame vengono instradati dalla sorgente verso il nodo destinatario.

La scelta della strada dovra' tenere conto anche del fatto che se ci sono troppi frame lungo la stessa linea si puo' creare traffico;

-risolvere i problemi di interconnessione tra ambienti diversi ossia problemi di interpretazione del messaggio.

8.4 LIVELLO DI TRASPORTO (TRANSPORT)

Questo livello provvede al trasferimento dei messaggi lungo la rete facendo in modo che gli utenti non debbano essere a conoscenza della struttura della rete. Quando il traffico diventa troppo intenso questo livello puo' utilizzare piu' percorsi di rete suddividendo i frame che compongono il messaggio lungo le strade disponibili.

8.5 LIVELLO DI SESSIONE (SESSION)

Una sessione oltre a consentire il trasporto dei dati fornisce anche dei servizi utili in certe applicazioni permettendo ad utente di connettersi ad un sistema remoto in time-sharing o per trasferire un file tra due macchine.

Un importante servizio di sessione e' la sincronizzazione:avendo la possibilita' di inserire nella serie dei dati dei punti di controllo dopo un

guasto dovranno essere ripetuti solo i dati che seguono l'ultimo punto di controllo.

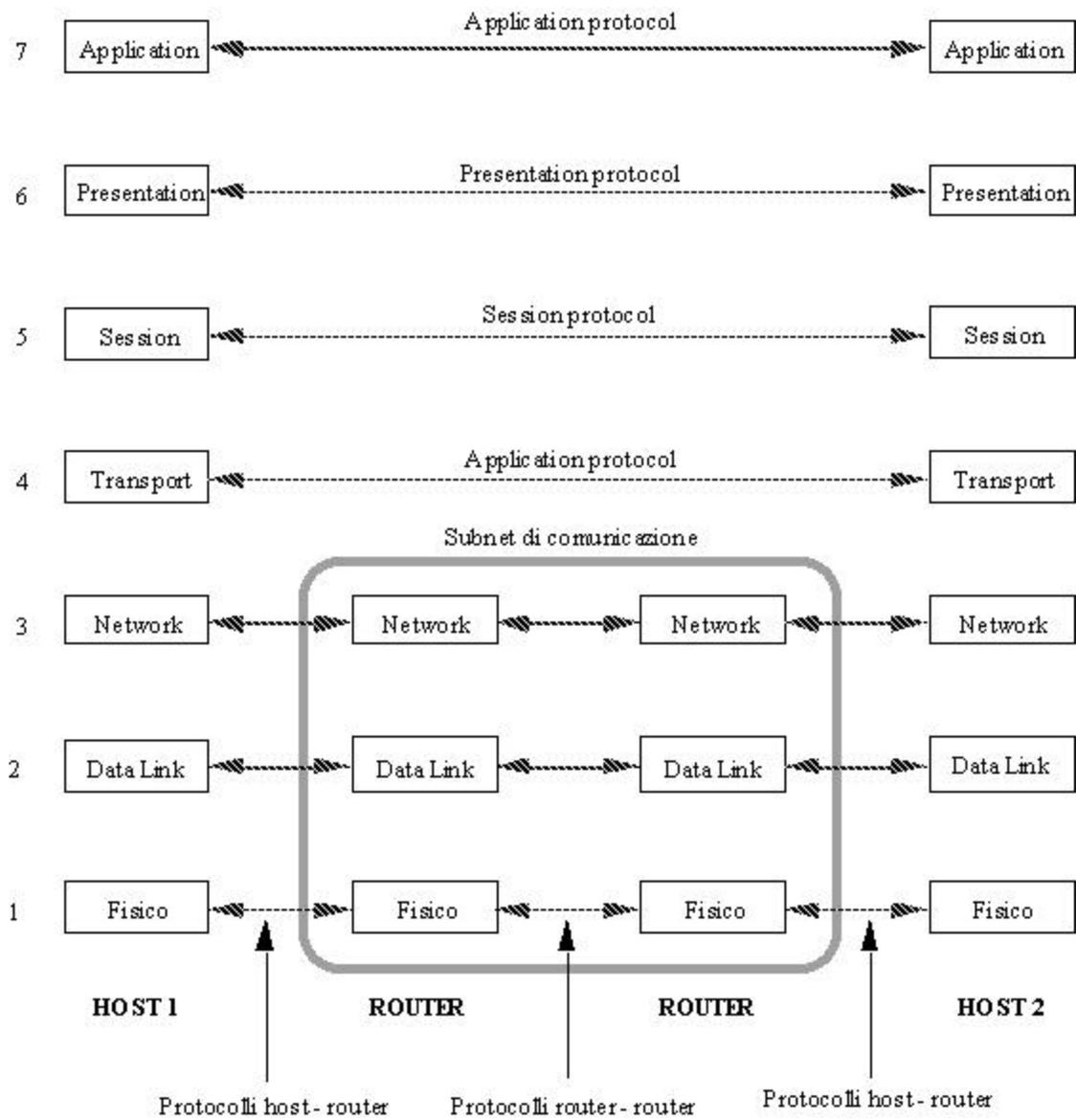
8.6 LIVELLO DI PRESENTAZIONE (PRESENTATION)

Il livello di presentazione si occupa della correttezza sintattica e semantica delle informazioni trasmesse permettendo la comunicazione tra due computer che hanno codici diversi. Questo livello effettua la conversione dalla rappresentazione usata all'interno del computer alla rappresentazione standard della rete.

8.7 LIVELLO DI APPLICAZIONE (APPLICATION)

Questo livello elabora le informazioni che serviranno per una determinata applicazione.

Tramite i servizi di questo livello il processo applicativo accede a quelli dei sottostanti livelli.



9 SISTEMA OPERATIVO: UNIX

Il sistema operativo e' un interfaccia tra l'utente e la macchina installato sul disco rigido(memori di massa).

Il piu' recente ed evoluto dei sistemi operativi e' UNIX.

I dischi rappresentano le memorie di massa piu' utilizzate. Si capisce quindi come l'ottimizzazione dei tempi medi di accesso al disco possa incidere in maniera notevole sui tempi di risposta dei singoli programmi.

Tutti i dischi vengono visti dal sistema come un insieme di settori contigui e numerati in ordine crescente, a partire dalla traccia 0. In particolare il settore 0, detto anche settore di BOOT e' destinato a contenere informazioni necessarie al BIOS per attivare il sistema operativo al momento dell'accensione del computer.

Sul disco rigido c'e' anche una tabella in cui sono registrate tutte le informazioni relative ai file contenuti in certo volume. Queste tabelle vengono chiamate FILE DIRECTORY (ogni file ha un nome).

Ogni riga della tabella contiene tutte le informazioni relative sia ai file sia alle aree disponibili, ovviamente nelle tabelle potrebbero essere presenti anche altre informazioni aggiuntive.

UNIX utilizza per la gestione dell'area dei file su disco, dei puntatori chiamati I-NODE. Ognuno di questi puo' puntare fino a 8 CLUSTER. I CLUSTER sono dei gruppi di settori contigui in cui vengono allocati file e sottodirectory, la loro dimensione varia secondo del tipo di formattazione applicato al supporto magnetico. Se si supera il numero di 8 CLUSTER si aumenta quello dei puntatori, che possono ulteriormente puntare ad altri CLUSTER.

PROGRAMMAZIONE

APPROCCIO ALLA PROGRAMMAZIONE

Per quanto riguarda la parte della programmazione abbiamo avuto un primo approccio con l'ambiente lavorativo.

Abbiamo lavorato con editor emacs ,diverso dall'ambiente in cui lavoriamo a scuola e abbiamo ,con la realizzazione di piccoli e semplici programmi in c++,passo passo, aggiungendo istruzioni e aumentando il grado di difficoltà ,creato il nostro progetto e ci siamo avvicinati ai due linguaggi che noi non conoscevamo l'XML e SQL.

L'XML (EXTENSIBLE MARKUP LANGUAGE) e' un sottoinsieme di SGML.Il suo obiettivo e' quello di consentire ad un genericodocumento di essere fornito, ricevuto ed elaborato su WEB.

XML e' stato progettato nell'ottica di una facile implementazione e per l'interoperabilità sia con SGML che conHTML. L'XML descrive una classe di dati e il comportamento dei programmi che li elaborano.

L'SQL e' l'acronimo di STRUCTURED QUERY LANGUAGE(linguaggio strutturato di interrogazione) e costituisce l'interfaccia standard per la creazione,la manipolazione e il controllo di database.

L'SQL risulta un linguaggio non procedurale, nel senso che l'utilizzatore deve solo specificare quali dati desidera elaborare e non la modalità con cui gli stessi dati devono essere individuati, e richiede la presenza di un linguaggio ospite poiché' al suo interno non prevede nessuna struttura di controllo.

Nella prima settimana abbiamo lavorato insieme. L'esercizio proposto era quello di creare una classe tabella che contenesse cinque colonne in cui i nomi venivano passati al momento dell'istanziamento.

I valori delle cinque colonne erano, per semplificare la realizzazione del codice ,valori interi in quanto lavorare con piu' tipi di dati rimaneva piu' complesso.

Inoltre la classe tabella doveva possedere tre metodi:

- inserimento di una riga(cinque valori);*
- stampa dell'intera tabella ;*
- l'estrazione di una riga, in modo d' avere in output un' altra tabella che contenesse solo le righe che soddisfanno solo le condizioni inserite in input.La condizione veniva aspressa:*

nome_colonne<valore

Prima di arrivare alla risoluzione completa dell'esercizio abbiamo suddiviso il problema realizzando ,inizialmente un codice che consentisse l'inserimento indefinito di colonne e righe attraverso l'estensione VECTOR che e' una classe che ha lo stesso significato di array soltanto che quest'ultima e' statica mentre con VECTOR abbiamo un'estensione a piacere.

Successivamente si e' realizzato un programma che permette di inserire da input il nome delle colonne e stampare solo le righe occupate.

Con l'implementazione di un nuovo codice siamo riusciti ad estendere il programma permettendo di modificare i valori delle colonne avendo cosi'la possibilita' di creare tabelle di ogni genere.

In seguito abbiamo realizzato un codice che oltre al normale inserimento da input consentiva un inserimento acquisito da sistema.

Infine siamo riuscite a terminare l'esercizio aggiungendo dei programmi che ci hanno portato ad inserire una condizione acquisita dal sistema e a stampare la tabella con solo le righe che la soddisfano.

Arrivate a questo punto ci siamo divise il lavoro per la realizzazione del database,una ha svolto la parte riferita all'SQL e l'altra riferita all'XML.

PARTE XML

Per quanto riguarda XML ho utilizzato l'istruzione di scrittura e di lettura sfruttando quanto fatto nei programmi precedenti,ricopiando cioe', la classe tabella.

L'istruzione di scrittura mi consentiva la stampa, come un documento XML. A questo punto sono passata a realizzare un programma che consentiva di leggere un documento XML utilizzando il pacchetto (o libreria) EXPAT. Per semplificare le cose mi sono data un numero di elementi limitato (nome, cognome, via, città) con l'idea di espanderlo ad essere più generale ma il tempo non me lo ha permesso.

Come prima cosa ho preso confidenza con EXPAT e ho cercato di capire come funziona, anche se mi è risultato molto difficile.

Successivamente ho letto il documento XML e lo ho stampato in output (non formattato) e in fine, ho realizzato l'ultima fase, quella che mi è risultata più difficile. Ho inserito i valori partendo dal documento HTML e utilizzando i programmi fatti durante la prima settimana ho messo i valori nella tabella.

Ho incontrato parecchie difficoltà nell'ultima fase in quanto mi è risultato difficile trovare le istruzioni per visualizzare la tabella.

PARTE SQL

Dalla divisione del lavoro, che ci avrebbe portato alla creazione del mini database, io ho svolto la parte che riguardava il linguaggio SQL, riuscendo, nel tempo che era rimasto le istruzioni SELECT e WHERE. La prima clausola ha permesso di costruire una tabella formata solo con le colonne specificate. La seconda ha ridotto la selezione alle righe che soddisfano la condizione.

I programmi che soddisfano le clausole precedenti sono stati svolti in C++, e si basavano sull'interpretazione di stringhe.

Anche qui, come nella parte comune, prima di completare l'esercizio richiesto, e quindi avere in output una tabella che restituisse solo i valori che soddisfano la condizione iniziale, ho lavorato in maniera graduale, dai programmi più facili a quelli complessi.

Per la clausola WHERE il primo codice svolto ha permesso di inserire da input due condizioni legate

dalla parola "AND", implementando poi alcune classi e utilizzando anche il poliformismo offerto dal C++, sono riuscita a far interpretare alla macchina la parola "AND" come prodotto logico, oltre a questo le

ho fatto esaminare le condizioni delle due stringhe legate dall'AND con un metodo specifico.

Nel secondo codice oltre all'AND ho fatto un altro metodo di comparazione che permettesse al processore di comprendere la condizione messa in input con le stringhe stavolta legate dalla parola "OR". Con questo programma ho soltanto visualizzato se la condizione inserita era vera o falsa.

I problemi maggiori che ho trovato nella stesura del codice sono stati l'uso di istruzioni che non conoscevo come "FIND", che e' servita per trovare gli operatori "AND" e "OR" all'interno della condizione, oppure "SUBSTR" per individuare la parte destra e sinistra rispetto all'operatore; e l'applicazione del poliformismo alle classi, perche' e' stato un argomento che a scuola non ho svolto ed ho appreso solo durante lo stage.

Il penultimo programma e' stato l'unione dei due precedenti, perche' consentiva di inserire nella stessa condizione iniziale non solo "AND" o "OR", ma entrambe e ripetuti anche piu' volte sulla stessa stringa. Problemi non ce ne sono stati perche' le istruzioni utilizzate per costruire il codice erano piu' o meno le stesse dei programmi precedenti.

L'ultimo programma richiedeva proprio un metodo che restituisse in output una tabella con solo le colonne che verificavano la condizione inserita.

CONCLUSIONI

Infine l'unione della parte in XML e della parte SQL hanno permesso la realizzazione del vero e proprio database system SQL basato su file descrittivo XML.