

Convegno Nazionale
Matematica senza Frontiere
Lecce, 5-8 marzo 2003

Matematica e televisione

Marcella Cucurachi
RAI-Roma

Sommario

Si cercherà di evidenziare il ruolo della formazione e delle competenze matematiche nella realizzazione e nella gestione tecnica di un portale di informazione, quale può essere il portale della televisione di stato www.rai.it. Le considerazioni relative a tale esempio possono essere estrapolate a qualunque altro portale.

Analogamente a tutte le architetture di comunicazione, Internet è strutturata su più livelli funzionali: ogni livello offre al livello superiore un insieme di servizi deputati alla risoluzione di una particolare classe di problemi di comunicazione.

La stratificazione dell'architettura Internet è organizzata su quattro livelli:

- strato di interfaccia di rete
- strato dei protocolli di internetworking
- strato dei protocolli di trasporto
- strato applicativo

Soffermandosi su i vari strati di questo schema architetturale si evidenzieranno le conoscenze matematiche che ne sono alla base, si porrà particolare attenzione allo strato dei protocolli di trasporto e lo strato dei protocolli applicativi.

La trasformazione di un'onda elettrica sinusoidale in onda quadra, schematizzazione del bit che viaggia in rete, è frutto di studi e teorie matematiche il cui maggiore artifice è stato Fourier con le sue trasformate. Nella progettazione di una rete intervengono studi statistici per il calcolo d'errore nella trasmissione dei dati e per garantire l'integrità dei dati trasmessi.

Studi matematici sono alla base della codifica dei dati trasmessi; codifiche che comprimono rilevanti moli di dati senza deteriorarli.

(Strato di interfaccia di rete)

Lo strato di Internetworking offre i servizi di instradamento dei dati tra gli host delle varie reti utilizzando il protocollo standard Internet Protocol (IP). Ad ogni "host" collegato in rete viene assegnato un numero intero su 32 bit, notazione "decimale puntata" protocollo IPV4 (Si farà la migrazione al protocollo IPV6, 64 bit), alla base di questa numerazione vi sono delle adeguate regole logico-matematiche che permettono di impostare regole di routing tra host ed apparati di

rete (router, gateway IP)

Uno degli aspetti che più preoccupano nell'affrontare un'interconnessione ad Internet è quello dei potenziali rischi che una tale interconnessione potrebbe portare.

Essi possono essere fatti risalire a due classi principali di problemi:

- la possibilità di accessi esterni non autorizzati alla rete privata
- la possibilità che l'informazione che viaggia sulla rete venga utilizzata non conformemente a quanto desiderato dalle entità comunicanti.

La gestione della sicurezza viene quindi concentrata in pochi punti chiave, in particolare nel punto di interconnessione della rete privata ad Internet, utilizzando dei sistemi chiamati "firewall". Su tali sistemi si impostano delle regole di filtraggio che seguono logiche matematiche di insiemistica.

Per quel che riguarda lo strato applicativo, è lo strato in cui si ha la visibilità del servizio che una tale architettura fornisce.

I servizi Internet sono classificabili come applicazioni distribuite strutturate secondo il modello client/server. Per applicazione distribuita si intende un insieme di processi cooperanti in esecuzione su elaboratori distinti interconnessi tramite una rete.

La realizzazione di tali applicazioni tramite software engineering, implica l'utilizzo delle teorie dei grafi, fondamento di qualunque tecnica di programmazione.

Un affermato standard nella progettazione di Base Dati è il modello Entità-Relazionale, esso è un linguaggio grafico semi formale per la rappresentazione di schemi concettuali.

Un ulteriore utilizzo delle teorie statistiche è quello del calcolo delle page views (in gergo traffico del portale) su tali dati si può valutare il successo o il fallimento della realizzazione di un progetto di questo tipo, in particolare un portale di informazione quale www.rai.it.

1 Introduzione

Aprire un browser e digitare www.rai.it è un gesto che può capitare di fare giornalmente, il risultato è la visualizzazione di quello che è il sito ufficiale della televisione di stato.

Navigando sul sito si possono reperire informazioni di tutti i tipi, la funzionalità maggiore è quella di fornire e garantire l'informazione via web.

In realtà il servizio fornito da questa struttura internet è duplice ovvero:

- È stato fornito ai giornalisti uno strumento per poter pubblicare le notizie sul web
- È stato fornito all'utente la possibilità di avere a disposizione l'informazione via web, telegiornali sia nazionali che regionali, oltre ad altro tipo di informazione.

Fornire questo servizio però è stato il risultato di una progettazione sia hardware che software.

Partendo dal portale RAI, che sarà preso come esempio, si cercherà di evidenziare il ruolo della formazione e delle competenze matematiche nella realizzazione e gestione di progetti di questo tipo.

Analogamente a tutte le architetture di comunicazione, Internet è strutturata su più livelli funzionali: ogni livello offre al livello superiore un insieme di servizi deputati alla risoluzione di una particolare classe di problemi di comunicazione.

La stratificazione dell'architettura Internet è organizzata su quattro livelli:

Strato dei Protocolli Applicativi
Strato dei Protocolli di Trasporto
Strato dei Protocolli di Internet Working
Strato di interfaccia di rete

Stratificazione architettura Internet

Nel livello più alto risiedono le applicazioni dell'utente ovvero i processi che richiedono lo scambio di dati. L'applicazione interagisce con lo strato di trasporto per il trasferimento delle informazioni. Tali dati attraversano tutti i livelli dell'architettura di comunicazione, ognuno dei quali, per fornire i propri servizi, vi aggiunge in testa o in coda informazioni: i dati così "imbustati" vengono quindi passati al livello inferiore. In ricezione, i dati passano dai livelli inferiori a quelli superiori; ciascun livello, prima di consegnare i dati al livello superiore, preleva ed analizza l'intestazione di sua competenza.

2 Strato di interfaccia di rete (Network Interface)

Questo strato si occupa di gestire la comunicazione sulla rete fisica cui è connesso l'elaboratore e di fornire allo strato superiore un insieme di servizi indipendenti da essa. L'architettura Internet è indipendente dalle caratteristiche della rete sottostante: ciò significa che è possibile interconnettere reti diverse tra di loro. Infatti per ogni tipologia di rete, viene fornito un opportuno strato software (Network Interface) che ne maschera le caratteristiche peculiari. Il compito dello strato di Network Interface è quello di inviare fisicamente il pacchetto sulla rete fisica verso la quale il protocollo IP ha deciso l'instradamento.

Questo significa:

- determinare l'indirizzo fisico (ad esempio MAC address) corrispondente all'indirizzo virtuale di destinazione (IP address)

- inviare il pacchetto, interfacciandosi con il driver della scheda, verso la destinazione di cui ormai conosce l'indirizzo fisico.

Dunque questo strato è responsabile dell'associazione tra indirizzo virtuale di destinazione e corrispondente indirizzo fisico.

Il bit che viaggia in rete viene gestito in questo strato. Il bit viene schematizzato elettricamente come un treno di segnali rettangolari, ovvero dei valori di tensione che oscillano tra $+5V$ a $-5V$ oppure da $+5V$ a $0V$, con la corrispondenza tra:

$$\begin{aligned} +5V &= 1 \\ -5V &= 0 \end{aligned}$$

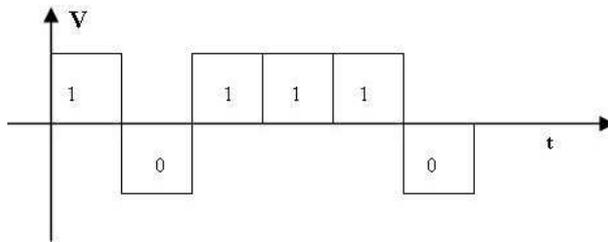


Figura 1: Schematizzazione di una stringa di bit nel tempo

Per stabilire la frequenza di trasmissione e quindi velocità di trasmissione si ricorre ad un operatore funzionale che applicato ad un segnale definito nel dominio del tempo ne individua un altro nel variabile continua della frequenza.

Tale operatore funzionale è noto come Trasformata di Fourier.

La trasformata di Fourier è quindi utile per rappresentare quei segnali per i quali non sussiste una struttura periodica. L'operazione di trasformazione è a volte indicata con la simbologia $X(f) = F\{x(t)\}$. La sua definizione formale dal punto di vista analitico è:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$$

la cui esistenza è garantita per segnali $x(t)$ impulsivi. Un segnale impulsivo è anche di energia, mentre non è sempre vero il viceversa. Spesso però, $X(f)$ esiste anche per segnali di energia; vedremo inoltre che può essere definita (grazie ad operazioni di passaggio al limite) anche per segnali di potenza periodici.

L'antitrasformata di Fourier $F^{-1}\{\}$ è l'operatore analitico che svolge l'associazione inversa a $F\{\}$, e che consente di ottenere, a partire da un segnale definito nel dominio della frequenza, quel segnale nel dominio del tempo

la cui trasformata è il primo segnale. L'operazione di antitrasformazione è definita come

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{j2\pi ft} dt$$

e vale ovunque $x(t)$ sia continuo, mentre nelle discontinuità di prima specie fornisce il valor medio di $x(t)$. Il risultato della trasformata

$X(f) = M(f)e^{j\phi(f)}$ è anche detto spettro di ampiezza complessa, mentre $M(f)$ e $\phi(f)$ sono detti spettri di modulo e fase.

Per dare un'idea dell'uso della trasformata di Fourier si riporta la trasformata del segnale che rappresenta il bit ovvero il rettangolo.

2.1 Trasformata di un rettangolo

Disponendo del segnale $x(t) = A \text{rect}\tau(t)$, si calcola lo spettro di ampiezza $X(f)$ che risulta pari a $A\tau \text{sinc}(f\tau)$.

Questo risultato viene graficato in Figura 2:

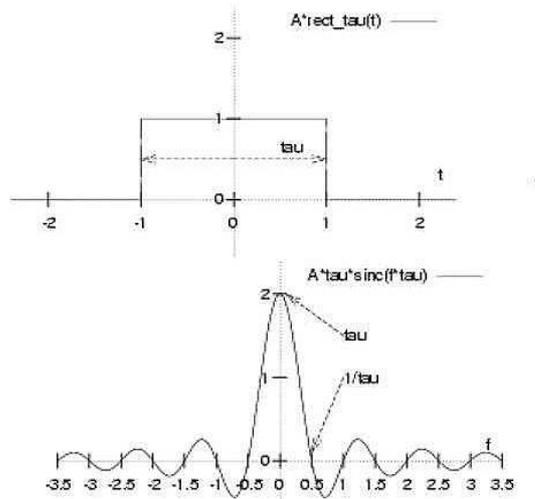


Figura 2: F -trasformata di un rettangolo di base $\tau=2$ ed ampiezza $A = 1$

Il noto andamento $\frac{\sin x}{x}$ rappresenta ora la distribuzione continua in frequenza dello spettro di ampiezza, ed il primo zero della curva si trova presso $f = \frac{1}{\tau}$, in modo del tutto simile al treno di impulsi rettangolari di base τ . Notiamo esplicitamente inoltre che, aumentando la durata del rect, lo spettro si restringe, addensandosi nella regione delle frequenze più basse; mentre al contrario, qualora il rect sia più breve, $X(f)$ si estende a regioni di frequenza più elevata.

2.2 Proprietà della trasformata di Fourier

2.2.1 Simmetria coniugata

Nel caso in cui $x(t)$ sia reale, risulta

$$X(f) = X^*(-f)$$

e quindi la parte reale di $X(f)$ è pari, e quella immaginaria dispari, ossia modulo $|X(f)|$ pari e fase $\arg\{X(f)\}$ dispari.

2.2.2 Dualità

Trasformata ed antitrasformata differiscono solo per il segno. Ciò comporta che se sostituiamo alla variabile f del risultato $X(f)$ di una F -trasformata, la variabile t , si ottiene una funzione del tempo $X(t)$ che, se nuovamente trasformata, fornisce il segnale originario $x(t)$, calcolato nella variabile f , cambiata di segno: $x(-f)$. Il concetto esposto, verificabile analiticamente con facilità, si riassume come

$$\begin{aligned} x(t) &\xrightarrow{F\{\}} X(f) \xrightarrow{t=f} X(t) \xrightarrow{F\{\}} x(-f) \\ X(f) &\xrightarrow{F^{-1}\{\}} x(t) \xrightarrow{f=t} x(f) \xrightarrow{F^{-1}\{\}} X(-t) \end{aligned}$$

e consente l'uso di risultati ottenuti "in un senso" (ad es. da tempo a frequenza) per derivare senza calcoli i risultati nell'altro (o viceversa). Come esempio di questa importante proprietà si riporta la trasformata di un $\text{sinc}(t)$.

Supponiamo di voler F -trasformare il segnale:

$$x(t) = B \frac{\sin \pi t B}{\pi t B} = B \text{sinc}(tB)$$

l'applicazione cieca dell'integrale che definisce la trasformata di Fourier appare un'impresa ardua...

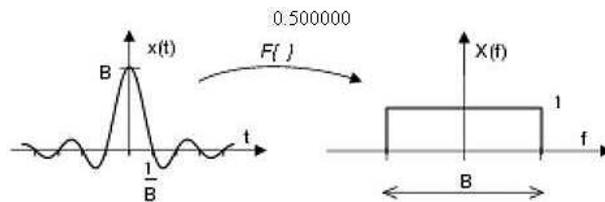


Figura 3: Trasformata di $\text{sinc}(t)$

Allora, ricordando che $F\{rect\tau(t)\} = \tau sinc(f\tau)$, scriviamo direttamente che

$$F\{Bsinc(tB)\} = rect_B(f)$$

Pertanto la trasformata di un $\frac{\sin x}{x}$ nel tempo, è un rettangolo in frequenza.

2.2.3 Linearità

Discende molto semplicemente dalla proprietà distributiva dell'integrale che definisce la trasformata. Pertanto se:

$$z(t) = ax(t) + by(t) \quad \text{allora} \quad Z(f) = aX(f) + bY(f)$$

2.2.4 Valore medio e valore iniziale

Subito verificabile una volta notato che la F -trasformata, calcolata per $f = 0$, si riduce all'integrale di $x(t)$, e quindi al suo valor medio. Pertanto:

$$m_x = \int_{-\infty}^{\infty} x(t)dt = X(f = 0) \quad e,$$

$$\text{per dualità:} \quad x_0 = x(t = 0) = \int_{-\infty}^{\infty} X(f)df$$

dove l'ultima relazione esprime la proprietà del valore iniziale.

2.2.5 Traslazione nel tempo

Si tratta di una proprietà molto semplice, e che ricorre frequentemente nei calcoli sui segnali. Manifesta la relazione esistente tra la trasformata dei segnali e quella degli stessi traslati, e si esprime con il predicato: se

$$z(t) = x(t - T) \quad \text{allora} \quad Z(f) = X(f)e^{-j2\pi f T}$$

Nella figura 4 si esemplifica il risultato ottenuto per $x(t) = rect\tau(t - T)$, in cui si è posto $\tau = 2$ e $T = .5$, ottenendo in definitiva $Z(f) = X(f)e^{-j2\pi f T} = 2sinc(2f)e^{-j2\pi f}$.

Poniamo ora l'attenzione sul fatto che l'espressione $x(t - T)$ indica un ritardo del segnale $x(t)$ di una quantità pari a T .

La circostanza che questo ritardo si traduca in un andamento lineare della fase di $X(f)$ ha una conseguenza notevole anche nell'altra direzione, ossia:

Una alterazione della fase di $X(f)$ deve essere lineare in f , se si desidera che il segnale mantenga inalterata la sua forma d'onda.

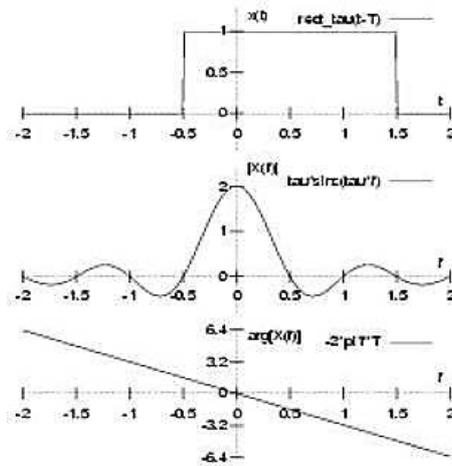


Figura 4: Trasformata di un segnale rettangolare traslato

2.2.6 Traslazione in frequenza (modulazione)

È la proprietà duale della precedente, e stabilisce che se

$$Z(f) = X(f - f_0) \quad \text{allora} \quad z(t) = x(t)e^{j2\pi f_0 t}$$

la proprietà è detta anche di modulazione.

2.2.7 Treno di impulsi

Prima di esaurire l'elencazione delle proprietà della trasformata di Fourier, presentiamo un diverso modo di ottenere lo spettro di un segnale periodico, facendo uso della particolare forma d'onda (ideale)

$$\prod_T(t) = \sum_{m=-\infty}^{\infty} \delta(t - mT)$$

chiamata treno di impulsi, e che si rivelerà di utilizzo frequente nei contesti del campionamento e delle trasmissioni numeriche.

2.2.8 Segnale periodico

Consideriamo un segnale periodico di periodo T espresso come

$$x(t) = \sum_{m=-\infty}^{\infty} g(t - mT)$$

di cui $g(t)$ costituisce un periodo: la riproduzione di infinite repliche di $g(t)$, spaziate di un periodo T l'una dall'altra, ricompongono il segnale periodico

originario. Sfruttando la proprietà di convoluzione con l'impulso traslato, la stessa somma può essere scritta come

$$x(t) = \sum_{m=-\infty}^{\infty} g(t) * \delta(t - mT) = g(t) * \sum_{m=-\infty}^{\infty} \delta(t - mT) = g(t) * \pi_T(t)$$

dove nel secondo passaggio si è sfruttata la linearità della convoluzione. Ricordando ora la proprietà della moltiplicazione in frequenza, troviamo $X(f) = G(f) \cdot F\{\pi_T(t)\}$; ci accingiamo allora a determinare $F\{\pi_T(t)\}$, ossia la trasformata del treno di impulsi.

2.2.9 Trasformata del treno di impulsi

La trasformata di un treno di impulsi è a sua volta un treno di impulsi, di periodo inverso a quello originario.

$$F\{\pi_T(t)\} = F\left\{ \sum_{n=-\infty}^{\infty} \Pi_n e^{j2\pi n F t} \right\} = \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T}\right) = \frac{1}{T} \pi_{\frac{1}{T}}(f)$$

ottenendo il risultato:

$$F\{\pi_T(t)\} = \frac{1}{T} \pi_{\frac{1}{T}}(f).$$

2.2.10 trasformata di un segnale periodico

Siamo finalmente in grado di esprimere la trasformata di un segnale periodico come il prodotto tra la F -trasformata di un suo periodo ed un treno di impulsi in frequenza:

$$X(f) = G(f) \frac{1}{T} \pi_{\frac{1}{T}}(f).$$

2.2.11 Codice a correzione d'errore

La robustezza delle trasmissioni digitali è in buona parte dovuta ai codici con correzione d'errore. Immaginiamo ad esempio di voler trasmettere le cifre binarie: 001

e di voler essere sicuri che siano state trasmesse correttamente: allora potremmo semplicemente trasmetterle due volte: 00 00 11

In questo caso se ricevessimo le cifre 00 01 11 sapremmo con certezza che c'è stato un errore nella trasmissione del secondo bit, e chi riceve potrebbe chiedere la ritrasmissione del messaggio. Si potrebbe fare anche di meglio, e trasmettere le cifre binarie tre volte, ad esempio 000 000 111, riuscendo così anche a correggere un eventuale errore: ricevendo le cifre 000 010 111 si capirebbe subito che c'è stato un errore nella trasmissione del secondo bit, e supponendo che la probabilità di avere due errori sia molto bassa, allora si potrebbe correggere l'errore e trovare che il secondo bit è uno 0. Questi

sono dunque due esempi di codici digitali: l'informazione viene codificata in modo da trovare un errore con il primo codice, e da correggerlo con il secondo.

In realtà i codici che abbiamo appena visto non sono efficienti, e pertanto non vengono utilizzati: di gran lunga più comuni sono quelli che utilizzano dei bit di parità per trovare e correggere gli errori. L'idea del bit di parità è la seguente: prendiamo una sequenza di bit come, ad esempio, 01100101. Se il numero di bit "alti" (cioè uguali a 1) è pari, allora aggiungiamo in fondo alla sequenza uno 0, altrimenti aggiungiamo un 1. Nel nostro esempio la sequenza di bit diventa 011001010, e se uno di questi bit viene trasmesso in modo errato allora uno 0 diventa 1 o viceversa, e cambia la parità della sequenza. In questo modo il bit di parità costituisce un meccanismo per rivelare un errore nella trasmissione, ma molto più efficiente di quello in cui ogni bit viene ripetuto (infatti in quel caso invece di 8 bit ne avremmo trasmessi 16, mentre con un bit di parità ne dobbiamo trasmettere solo 9). A questo punto è ovvio chiedersi come si possano utilizzare i bit di parità per correggere gli errori invece di accontentarsi di rivelarne la presenza. Una classe particolarmente ingegnosa di codici con correzione di errore che utilizzano dei bit di parità è stata inventata nel 1947 da Richard Hamming: l'idea è quella di dividere una sequenza di bit in sottosequenze sovrapposte, così che i bit di parità associati a ciascuna sottosequenza permettono di individuare il bit che è stato ricevuto in modo errato.

I codici di Hamming sono utili nel caso in cui non ci sia più di un errore in ciascuna sequenza di bit trasmessa; esistono tuttavia dei codici più complessi che permettono di affrontare anche altre situazioni. In generale, un codice con correzione di errore è definito da un certo numero M di sequenze di n simboli (ad esempio i simboli 0 e 1 nei codici binari) associate ai messaggi che si vogliono trasmettere, in modo che due sequenze qualunque differiscano tra loro per almeno d simboli. Ad esempio, se volessimo trasmettere soltanto le lettere A, B e C mediante un codice binario con sequenze lunghe $n = 6$ bit, allora potremmo prendere come messaggi validi le $M = 3$ sequenze:

A 000000

B 000111

C 111111

che differiscono una dall'altra per almeno $d = 3$ simboli, e quindi se ricevessimo la sequenza 001111 la interpreteremmo come una B, perché la sequenza associata alla B è quella più vicina al segnale ricevuto. Ovviamente tanto più grande è il numero di errori che ci si aspetta di trovare in una sequenza, tanto maggiore dev'essere la minima differenza d per poter associare le diverse sequenze ricevute a uno dei messaggi.

La teoria dei codici è una sana e robusta branca della matematica moderna, tuttavia, nel linguaggio comune, quando si parla di codice si intende spesso codice segreto. Anche i codici segreti sono ormai analizzati sempre più a fondo dai matematici.

3 Stato di internetworking

Questo strato offre i servizi di instradamento dei dati tra gli host delle varie reti utilizzando il protocollo standard Internet Protocol (IP). Tale protocollo risiede negli host terminali e nei dispositivi che interconnettono le differenti reti svolgendo solo funzioni di routing (router o gateway IP). L'unità di informazione nel protocollo IP è detta datagram. Il protocollo IP non è orientato alla connessione (connectionless): ogni datagram è trattato come unità di informazione indipendente ed il suo trasferimento non è preceduto dalla fase di apertura della connessione né è seguito dalla fase di chiusura della stessa. Inoltre il protocollo non garantisce il controllo di flusso, l'affidabilità dei dati e la loro sequenza.

In trasmissione i dati provenienti dal livello superiore vengono strutturati in datagram IP; nella relativa intestazione vengono inserite, ad esempio l'indicazione del mittente e del destinatario ed il livello di trasporto utilizzato. Viene quindi applicato l'algoritmo di routing per l'instradamento sulla stessa o verso altre reti. In ricezione viene esaminata l'intestazione del datagram per determinare se esso è giunto a destinazione: in caso affermativo si elimina l'intestazione e si individua il protocollo dello strato superiore cui passare, in forma opportuna i dati estratti; altrimenti si provvede a re-instradare il datagram.

Il campo dati è preceduto da un'intestazione contenenti gli indirizzi destinazione e sorgente ed un campo che specifica il tipo di trama trasportata. Il pacchetto termina con 4 bytes di Cyclic Redundancy Check.

L'architettura Internet, quindi consiste in un insieme di reti interconnesse che agiscono come un'unica unità. Le reti sono collegate con apparati, detti "gateway", che hanno la funzione di instradare i dati sulla base dell'indirizzo della rete destinataria.

Con riferimento alla figura 9, il percorso software di un messaggio che attraversa più reti interconnesse, tramite un'architettura Internet, prevede che ciascun pacchetto venga analizzato dai gateway intermedi e reindirizzato verso ulteriori reti fino a che non raggiunge il nodo destinatario.

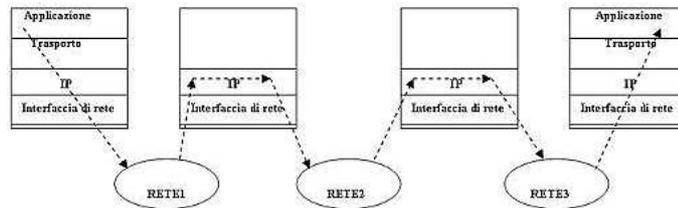


Figura 5: Percorso di un Datagram a livello IP

Le funzioni dello strato di internetworking sono presenti sia sugli host terminali che sui “gateway” e sono:

- omogeneizzare i servizi di rete;
- fornire il servizio di rete; tale servizio non è affidabile è connection-less;
- offrire un meccanismo di indirizzamento globale tra le reti;
- gestire le tabelle di instradamento (routing);
- gestire gli errori.

3.1 Il protocollo IP

Il protocollo IP definisce un meccanismo di consegna dati di tipo “non affidabile” e “non orientato alla connessione” (connectionless). Questo significa che non viene posta nessuna azione per garantire l’integrità del pacchetto, la consegna o il rispetto della sequenza d’invio.

L’unità trasmissiva definita in questo protocollo è detta “IP datagram” ed è formata da due campi: l’intestazione (header) ed il campo dati-utente (user-data).

Il campo dati è costituito da zero o più byte che provengono da protocolli di strati superiore e la sua dimensione massima è data dal parametro MTU (Maximum Transmission Unit) che varia da rete a rete, se lo strato riceve un numero di byte maggiore, deve provvedere alla frammentazione in modo da adattarli alla rete. Questi frammenti sono inoltre instradati in maniera indipendente e riuniti alla destinazione.

Nel campo intestazione, invece vengono inserite informazioni relative a:

- indirizzo IP sorgente;
- indirizzo IP destinatario;
- checksum intestazione: campo per il controllo dell’integrità dell’intestazione;
- protocollo: protocollo selezionato nello strato trasporto;
- tempo di vita: è un numero intero che identifica il numero di router che il datagram può attraversare prima di essere eliminato dalla rete;
- offset frammento: offset del frammento nell’ambito del totale dei dati da trasmettere;
- identificativo del datagram (utilizzato per gestire la frammentazione);
- flag: indica se il datagram rappresenta un frammento;
- lunghezza totale: lunghezza del datagram espressa in byte;

- tipo di servizio: identifica la qualità del servizio che un applicativo richiede alla rete;
- IHL: lunghezza dell'header.

Per quanto riguarda gli indirizzi ad ogni "host" collegato in rete viene assegnato un numero intero su 32 bit che rappresenta l'indirizzo Internet. Per facilitare la lettura di tali indirizzi, è convenzione scriverli nella cosiddetta notazione "decimale puntata", cioè rappresentarli con quattro numeri interi decimali separati da punti.

Con questa convenzione, ad esempio, l'indirizzo Internet a 32 bit:

10000000000010100000001000011110

viene scritto come 128.10.2.30. L'aumento di leggibilità è evidente. Tale indirizzo, a livello concettuale, è rappresentato dalla seguente coppia ordinata:

$\langle \text{netid}, \text{hostid} \rangle$

dove netid (net identifier) identifica una rete, mentre hostid (host identifier) identifica un calcolatore collegato a tale rete.

Ogni gateway del sistema instrada il datagramma proprio sulla base dell'informazione netid. Relativamente alla parte netid, gli indirizzi IP possono essere ricondotti a differenti classi in funzione dei bit più significativi. In particolare si hanno le seguenti classi:

- A: bit più significativo a zero (0.0.0.0-127.255.255.255), 7 bit per l'indirizzo di rete e 24 bit per il campo host. Complessivamente risultano realizzabili 128 reti con 16777216 host ciascuna;
- B: primi due bit pari a "10" (128.0.0.0-191.255.255.255), 14 bit per l'indirizzo di rete e 16 bit per il campo host. Complessivamente risultano realizzabili 16384 reti con 65536 host ciascuna;
- C: primi tre bit pari a "110" (192.0.0.0-223.255.255.255), 21 bit per l'indirizzo di rete ed 8 bit per il campo host. Complessivamente risultano realizzabili 2097152 reti con 256 host ciascuna.

Riassumendo:

classe	numero bit	
	net	host
A	7	24
B	14	16
C	21	8

Schema delle classi di indirizzamento

I server del portale RAI appartengono ad una rete di classe B, ma vengono utilizzati solo 255 indirizzi, quindi è sottoposta ad un filtro che la rende subnet di classe C.

È evidente che con lo sviluppo di Internet l'identificazione degli host con indirizzamento IP a 32 bit sta diventando insufficiente, così per il futuro si prevede il passaggio dalla "notazione decimale puntata" nota come protocollo IPV4, al protocollo IPV6 ovvero indirizzamento a 64 bit.

Le principali funzioni del protocollo IP sono:

in trasmissione

- incapsulare in "datagram" i dati provenienti dallo strato di trasporto;
- preporre l'opportuna intestazione, in modo che il "datagram" contenga tutte le informazioni necessarie per essere consegnato all'entità Ip di destinazione;
- applicare l'algoritmo di "routing";
- inviare i dati verso l'opportuna interfaccia di rete;

in ricezione

- verificare la validità dei "datagram" in arrivo;
- esaminare l'intestazione;
- verificare se i dati sono destinati a sé (entità IP locale) oppure se sono da reinstradare verso un'entità remota (in questo caso li tratta come dati in trasmissione);
- se sono dati locali, scegliere l'opportuno protocollo dello strato di trasporto ed inviarli nel formato opportuno eliminando l'intestazione.

Quest'ultima funzione, molto importante, è chiamata demultiplazione (software demultiplexing) e permette ad uno stesso modulo software IP di servire differenti moduli software dello strato di trasporto.

4 Stato di trasporto

Lo scopo dello strato di trasporto è quello di fornire una comunicazione end-to-end nei confronti dei processi residenti sull'elaboratore indirizzato dall'IP. A tal fine vengono individuati un insieme di punti logici di accesso al servizio detti "port number". I primi 1023 "port number" sono riservati per servizi diffusi o di sistema. I programmi utente possono utilizzare numeri maggiori o uguali a 1024.

Nell'architettura Internet esistono due standard principali di protocolli che appartengono allo strato di trasporto:

- User Datagram Protocol (UDP);
- Transmission Control Protocol (TCP).

Il servizio di trasporto fornito dal protocollo UDP è datagram, mentre TCP fornisce un servizio di trasporto orientato alla connessione (connection-oriented) e garantisce l'affidabilità, la sequenza e la non duplicazione.

4.1 Il protocollo UDP (User Datagram Protocol)

Il protocollo UDP non garantisce l'affidabilità in quanto è di tipo connectionless e non ha meccanismi per la ritrasmissione dei dati. È opzionale un meccanismo di rilevazione degli errori basato su un checksum, ma nel caso di errore il datagram viene scartato.

Non c'è un meccanismo di controllo di flusso: nel caso di overflow sul buffer di ricezione il datagram viene scartato. orientato alla connessione (connection-oriented) e garantisce l'affidabilità, la sequenza e la non duplicazione.

4.2 Il protocollo TCP (Transmission Control Protocol)

Fornisce un servizio di trasporto affidabile e connection oriented, svolgendo le seguenti funzioni:

- controllo di errore;
- controllo di flusso;
- controllo di sequenza.

L'unità informativa scambiata tra le entità TCP è detta segmento; il segmento viene imbustato in un datagram IP. Al termine della descrizione degli strati di internetworking e di trasporto è intuitivo come sia necessario avere delle conoscenze matematiche per gestire gli indirizzamenti degli host. Dopo aver dato un'idea degli indirizzamenti dei server è bene cercare di espletare su come questi server vengono disposti per realizzare una rete LAN che permette ad essi di comunicare tra loro e distribuire le attività per fornire il servizio finale ad Internet.

Qualsiasi portale viene realizzato attraverso la disposizione dei server in una LAN (Local Area Network), ed ogni LAN in linea di massima si può dividere in due parti:

- FRONT END (la parte che si affaccia direttamente su Internet)
- BACK-END (la parte più remota rispetto ad Internet e la parte più protetta)

Nel caso del portale RAI il Front-End si può dividere a sua volta in due sottoreti.

- FRONT-END esterno composto da web-servers, application servers ed observation server
- FRONT-END interno composto da Page Delivery e Docroot Server

Ad ogni parte della rete LAN corrispondono indirizzi IP diversi, in questo caso la rete Lan si può considerare l'insieme di tre sottoreti che comunicano tra di loro, e tra una rete ed un'altra vi sono degli apparati di protezione. Nella parte più remota il BACK-END in genere si posizionano i server che contengono i dati più sensibili, per esempio il Data Base.

Nel caso del portale RAI il Data Base è contenuto in una struttura chiamata cluster che garantisce l'alta affidabilità.

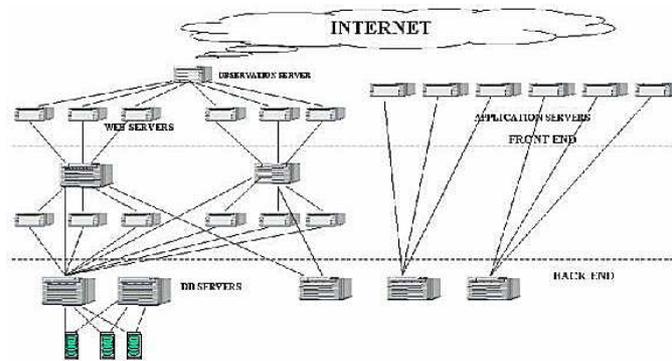


Figura 6: Schema architetturale

5 Sicurezza in rete

Uno degli aspetti che più preoccupano nell'affrontare un'interconnessione ad Internet è quello dei potenziali rischi che una tale interconnessione potrebbe portare.

I potenziali rischi possono essere fatti risalire a due classi principali di problemi:

- la possibilità di accessi esterni non autorizzati alla rete privata
- la possibilità che l'informazione che viaggia sulla rete venga utilizzata non conformemente a quanto desiderato dalle entità comunicanti

Il primo problema viene affrontato con adeguate soluzioni architetturali specifiche del mondo Internet (firewall), il secondo problema (che coinvolge la riservatezza dei dati scambiati, l'identificazione certa della controparte

con la quale si sta comunicando) coinvolge quegli aspetti più generali di sicurezza dei sistemi distribuiti (password crittografia). Il problema attualmente più sentito ed analizzato è il controllo degli accessi esterni alla rete privata.

Nel momento in cui si apre la propria rete all'esterno c'è la possibilità che un utente di internet possa penetrare nella nostra rete e quindi provocare danni dovuti al furto o alterazione dell'informazione in essa contenuta. Per non rendere difficoltoso l'utilizzo del computer all'utente finale è opportuno concentrare la gestione della sicurezza in pochi punti chiave ed in particolare nel punto di interconnessione della rete privata ad Internet. In tale punto si può tenere sotto controllo tutto il traffico (sia entrante che uscente) e quindi verificare se le attività dell'utenza è conforme alla security policy definita dall'organizzazione alla quale si appartiene. Il controllo del traffico in transito può essere eseguito a vari livelli:

- Si può innanzi tutto decidere a quali macchine rendere disponibile la rete. Tale controllo può essere facilmente eseguito analizzando gli indirizzi sorgente e destinatari dei pacchetti IP in transito, scartando quei pacchetti ritenuti illegali.
- Si potrebbe scartare tutto il traffico proveniente da una certa rete perchè ritenuta notoriamente inaffidabile.
- Si potrebbe permettere il transito di pacchetti provenienti solamente da un numero limitato di sedi.
- Si potrebbe rendere la rete privata totalmente inaccessibile dall'esterno permettendo solo il traffico uscente ed i corrispondenti pacchetti di risposta (controllando i flag del pacchetto TCP).

Tale approccio ha il vantaggio di una relativa semplicità di gestione, ma ha lo svantaggio di dover negare o permettere l'interazione con una certa rete IP.

Per ottenere un controllo più efficiente delle comunicazioni da e verso Internet è opportuno analizzare il traffico anche a livelli superiori a quelli IP. In particolare è possibile decidere quale servizio abilitare e quali disabilitare lasciando transitare o meno pacchetti con determinati indirizzi a livello UDP/TCP.

Si può fare un esempio si può decidere di far passare tutto il traffico relativo alla posta elettronica (port TCP pari a 25) e bloccare tutto il traffico relativo all'emulazione di terminale (port TCP pari a 23) notoriamente più pericoloso.

L'indirizzamento dei processi da parte dei protocolli di trasporto (UDP/TCP) è basato su un numero intero che identifica univocamente una struttura dati (socket) che consente l'invio e la ricezione dei messaggi da parte dei processi stessi. In questo caso ad ogni servizio viene associato un numero di port che identifica il lato server dell'applicazione ad esempio:

- servizio di collegamento remoto (telnet) è assegnato il port 23;
- servizio di trasferimento file (ftp) è assegnato il port 21;
- servizio di e-mail è assegnato il port 25.

Ovviamente, il controllo sul port TCP può essere combinato con quello sull'indirizzo IP.

Si fa qui di seguito un esempio di configurazione di un router di filtraggio per evidenziare come alla base vi sia una logica matematica. È un esempio che mostra l'implementazione di una specifica linea di condotta. La linea di condotta è differente da progettista a progettista. In questo esempio, una compagnia ha una Classe B di indirizzi di rete 128.88.0.0 e sta usando 8 bit per le subnet. La connessione Internet è sulla sottorete rossa 128.88.254.0. Tutte le altre sottoreti sono considerate di fiducia o sottoreti blu.

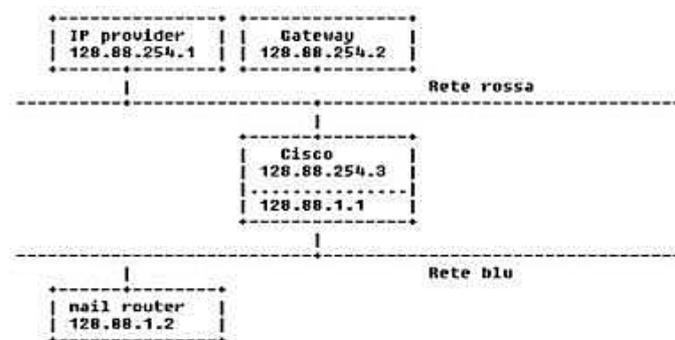


Figura 7: Esempio di router di filtraggio

1. Vengono filtrati i soli pacchetti in uscita
2. Le regole sono testate in ordine e si fermano quando la prima occorrenza è stata trovata.
3. C'è una implicita regola di rifiuto alla fine di una lista di accesso che nega ogni cosa.

L'esempio sotto tratta il filtraggio delle porte di una configurazione. I numeri di linea e la formattazione sono state aggiunte per leggibilità.

Le linee guida per l'implementazione sono: - Niente non esplicitamente permesso è negato.

- Il traffico tra la macchina esterna che fa da gateway e l'host della rete blu è permesso.- I servizi permessi sono originati dalla rete blu.- Assegna un range di porte alla rete blu per le informazioni di ritorno della connessioni dati FTP.

1. no ip source-route
2. !
3. interface Ethernet 0
4. ip address 128.88.1.1 255.255.255.0
5. ip access-group 10
6. !
7. interface Ethernet 1
8. ip address 128.88.254.3 255.255.255.0
9. ip access-group 11
10. !
11. access-list 10 permit ip 128.88.254.2 0.0.0.0 128.88.0.0 0.0.255.255
12. access-list 10 deny tcp 0.0.0.0 255.255.255.255 128.88.0.0 0.0.255.255
lt 1025
13. access-list 10 deny tcp 0.0.0.0 255.255.255.255 128.88.0.0 0.0.255.255
gt 4999
14. access-list 10 permit tcp 0.0.0.0 255.255.255.255 128.88.0.0 0.0.255.255
15. !
15. access-list 11 permit ip 128.88.0.0 0.0.255.255 128.88.254.2 0.0.0.0
16. access-list 11 deny tcp 128.88.0.0 0.0.255.255 0.0.0.0 255.255.255.255
eq 25
17. access-list 11 permit tcp 128.88.0.0 0.0.255.255 0.0.0.0 255.255.255.255

Linea Spiegazione

=====

- 5 - Ethernet 0 è nella rete rossa. La lista estesa 10 di permessi di accesso sarà applicata all'output su questa interfaccia. Puoi anche pensare l'output dalla rete rossa come input nella rete blu.
- 9 - Ethernet 1 è nella rete blu. La lista estesa 11 di permessi di accesso sarà applicata all'output su questa interfaccia.
- 11 - Permette tutto il traffico dalla macchina gateway alla rete blu.

- 12-14 - Permette connessioni originarie della rete rossa che entrano tra le porte 1024 e 5000. Ciò avviene per permettere alle connessioni dati in FTP di dialogare alla rete blu. ; ; N.d.T.: permessi e richieste nella connessione FTP avvengono su una porta diversa da quella dove vengono scambiati i dati ;5000 è stato scelto come limite superiore.
- 16 - Permette l'accesso a tutti i pacchetti della rete blu alla macchina gateway.
- 17 - Nega l'SMTP (porta tcp 25) mail alla rete rossa.
- 18 - Permette a tutto l'altro traffico TCP di accedere alla rete rossa.

Il protocollo utilizzato per il trasferimento dei dati in WWW è l'http. L'attuale software http non supporta sistemi di sicurezza se non in modo molto blando. I due modi per migliorare la sicurezza in www sono SSL ed S-HTTP

6 Strato dei protocolli applicativi

I servizi Internet sono classificabili come applicazioni distribuite strutturate secondo il modello client/server. Per applicazione distribuita si intende un insieme di processi cooperanti in esecuzione su elaboratori distinti interconnessi tramite una rete. In un'applicazione di tipo client/server, il processo server si configura come fornitore di una risorsa/servizio nei confronti di uno o più processi client. Il processo client interfaccia l'utente a terminale ne interpreta i comandi (relativi al particolare servizio) invia una richiesta al processo server fornendo i parametri di servizio impostati dall'utente, attende i risultati dal server e li rende disponibili all'utente. Il processo client è dunque responsabile dell'interfaccia (più o meno amichevole) offerta all'utente. È evidente la necessità per i processi client e server, in generale per i processi cooperanti, di disporre di un meccanismo di comunicazione che consenta lo scambio delle informazioni necessarie per espletare il servizio a loro demandato. Il meccanismo di comunicazione viene reso disponibile da una pila protocollare seguendo due possibili modelli:

- message passing
- Remote Procedure Call

6.1 Message Passing

Message passing si lega direttamente ai servizi resi disponibili dal livello di trasporto di una qualsiasi architettura di comunicazione. Il livello di trasporto rappresenta il primo livello in cui è possibile l'indirizzamento e quindi la comunicazione tra processi. Tale livello fornisce le primitive che consentono lo scambio di messaggi tra processi (primitive send/receive).

Facendo riferimento all'architettura Internet, l'accesso ai servizi dello strato di trasporto può avvenire utilizzando l'interfaccia socket.

Il socket rappresenta il mezzo per interfacciare direttamente un'applicazione al livello di trasporto. È una struttura dati creata dinamicamente da un processo nell'istante in cui ha la necessità di eseguire uno scambio di informazioni: logicamente può essere vista come composta da due code in cui vengono depositati rispettivamente i messaggi ricevuti e quelli da trasferire. Nel caso del portale RAI i prodotti software utilizzati seguono questa modalità protocollare di comunicazione. Si citano per esempio i prodotti utilizzati:

VIGNETTE ambiente distribuito utilizzato per la realizzazione di siti dinamici

DOUBLECLIK prodotto utilizzato per la pubblicazione e gestione del reporting dei banner pubblicitari

AUTONOMY prodotto utilizzato per realizzare il motore di ricerca all'interno del sito

Questi prodotti software utilizzati come esempio hanno una forte dipendenza dalla struttura architeturale della rete e dalla comunicazione tra i server che costituiscono la rete.

Si prende come esempio uno di questi prodotti ed esattamente Vignette, e per esso si evidenzia graficamente la sua forte dipendenza architeturale.

Architettura Vignette Front-End

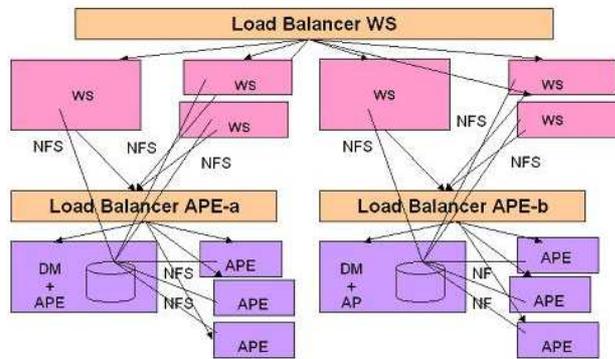


Figura 8: Architettura ambiente Vignette Front-End

Architettura Vignette – Back-End

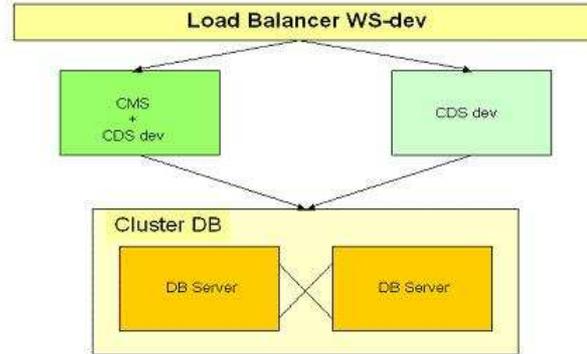


Figura 9: Architettura Ambiente Vignette Back-End

6.2 Remote Procedure Call

Con il termine Remote Procedure Call (RPC) si indica un meccanismo di comunicazione per la programmazione di applicazioni distribuite. Lo scopo principale è quello di facilitare la realizzazione di tali applicazioni, permettendo al programmatore di astrarsi dai problemi di comunicazione e gestione della connessione tra processi remoti, cioè dall'interazione con il livello di trasporto e sviluppare applicazioni distribuite come insieme di procedure, alcune delle quali remote. Per chiamata di procedura si intende un costrutto linguistico che permette di trasferire il controllo dell'elaborazione fra un'entità chiamante (client), che richiede un determinato servizio, ed un'entità chiamata, che lo espleta (server).

6.3 Primi Servizi Sviluppati In Rete

I primi servizi sviluppati sulla rete sono: telnet: "terminale virtuale" nella terminologia Internet. Quello che si voleva ottenere con il telnet era di consentire ad un qualunque terminale connesso su una certa macchina di essere visto come collegato ad una qualsiasi macchina raggiungibile in rete. Dal punto di vista funzionale questo servizio è relativamente semplice: basta inviare su di un collegamento aperto con la macchina destinazione quanto digitato dall'utente per poi interpretarlo come se generato da un terminale locale. Dal punto di vista concettuale invece il telnet è il primo servizio che ha fatto crollare il concetto di host-centring computing per passare ad un concetto di network-centric computing: una visione nella quale macchine connesse in rete si scambiano liberamente informazioni.

ftp: "file transfer protocol". Il servizio ftp rende possibile il trasferimento

di file tra due macchine connesse in rete, l'utente sfrutta le risorse della macchina con quella che viene definita "un'interfaccia a carattere".

6.4 Servizio www (world wide web)

La realtà attuale è quella che vede l'utente accedere alle risorse della macchina e della rete tramite interfacce grafiche, che nascondono sia le caratteristiche delle risorse che le primitive richieste per accedervi. In questa ottica va inquadrato il servizio WWW (World Wide Web).

Il fenomeno WWW è l'elemento di maggior impatto per spiegare il successo di Internet e per valutare le sue potenzialità.

Il servizio WWW può essere visto come un database di informazioni condivise (shared database) che, in prima istanza può essere usato come strumento per esporre informazioni (Reference Publishing System), ma che, attraverso opportuni mecenismi, è in grado di evolvere verso servizi di tipo interattivo.

Il servizio si basa essenzialmente sul concetto di ipertesto portato alle sue estreme conseguenze e potenzialità. Il server, nella sua forma più semplice è ancora un semplice contenitore di file in formato definito in un linguaggio specifico: l'HyperText Markup Language (HTML). I visualizzatori di tali file (browser) sono client (per esempio Netscape o Explorer) e costituiscono l'interfaccia utente: le informazioni di tipo multimediale vengono presentate dal browser interpretando il file in formato HTML e visualizzandole opportunamente.

Il file ipertestuale è un file di testo che contiene delle parole chiave (link) che vengono evidenziate dal browser in maniera diversa rispetto alle altre parole. Ad ogni link viene associato un comando che consente lo spostamento non lineare del testo. Nel web, questo concetto è stato esteso per consentire lo spostamento della visualizzazione su un altro file dello stesso server o su un file residente su un altro server, realizzando quella ragnatela di informazioni che è attualmente internet.

Cercando di semplificare ci sono essenzialmente tre tecnologie legate al servizio WWW su cui si basano le interazioni tra un WEB client ed un WEB server connessi su una rete TCP/IP:

1. HTTP (HyperText Transfer Protocol): rappresenta il protocollo che governa la comunicazione tra client e server WEB.
2. HTML (HyperText Markup Language): è il formato dei documenti WEB. Esso contiene delle direttive che consentono di definire l'impostazione del testo, le parole chiave e di comandi ad esse associati.
3. URL (Uniform Resource Locator): è un puntatore ad una risorsa Internet, viene utilizzato per indirizzare le pagine memorizzate in un WEB server.

Le informazioni da inserire sul client per accedere ad una risorsa WEB vanno specificate secondo il formato URL che nel nostro caso è il seguente:

<http://www.rai.it/portale> dove:

- `http` rappresenta il protocollo
- `www.rai.it` è l'indirizzo del server
- `portale` è il percorso che individua il file nella memoria di massa del server.

Un approccio del WEB basato esclusivamente sul linguaggio HTML, ha evidenziato i suoi limiti, in quanto si potevano utilizzare solamente informazioni contenute staticamente sul server. Si è sentita quindi l'esigenza di riferenziare attraverso un link non solo una pagina WEB, ma anche dei programmi, attivati sul server tramite una modalità standard, detta Common Gateway Interface (CGI). Si può pensare a un programma CGI come ad un programma che è in grado di generare dinamicamente una pagina HTML. La trasformazione è intuitiva si è passati da un semplice strumento che visualizza informazioni già presenti su un server WEB (electronic publishing) ad un servizio che è in grado di generare automaticamente le informazioni in formato WEB (dynamic publishing).

In sostanza il CGI program restituisce all'utente una pagina WEB creata dinamicamente dove sono inserite indicazione circa l'esito della query inoltrata attraverso il server.

Uno svantaggio dell'utilizzo delle CGI era quello di dover gestire sulla macchina dell'utente tanti programmi client quante sono le applicazioni che utilizza, si è fatta un'ulteriore evoluzione.

Il nuovo approccio si basa sul concetto di software agent: il client ed il server scambiano non soltanto dati, ma anche i programmi da eseguire sul client (software agent) vengono trasferiti sulla macchina dell'utente dal server.

Il problema si sposta dunque sui meccanismi con cui garantire la mobilità dei software agent e sulla definizione di un linguaggio comune con il quale descrivere le funzionalità di questi programmi. La soluzione in prima istanza fu quella di apportare delle modifiche al linguaggio HTML (direttiva `applet`), con tale modifica è possibile richiedere dal browser al server il trasferimento di un programma (software agent) scritto in linguaggio Java.

Il passo successivo è stato quello di far eseguire i programmi java sul server e fornire i risultati al client (direttiva `servelet`).

Nel caso del portale RAI, esso è stato realizzato con un apposito prodotto chiamato Vignette che permette la realizzazione di portali dinamici, ovvero portali in cui i contenuti variano molto velocemente anche ogni 5 minuti, pensate alla sezione del portale in cui si svisualizzano le notizie di televideo o rainetnews. Come linguaggio di programmazione si è usato un linguaggio simile all'HTML chiamato TCL, vengono utilizzate anche CGI e servlet, alcuni siti del portale invece sono statici e per realizzare quest'ultimi si è utilizzato il linguaggio HTML. In entrambi i casi, alla base della realizzazione

vi è una tecnica di programmazione e software engineering, che implica l'utilizzo della teoria dei grafi.

Infine i dati sensibili relativi ai contenuti del portale sono organizzati in un Data Base posizionato su due server in Back-End facenti parte di un cluster per garantire l'alta affidabilità dei dati. Un affermato standard nella progettazione dei Base Dati è il modello Entità-Relazionale, esso è un linguaggio grafico semi formale per la rappresentazione di schemi concettuali.

Vorrei inoltre sottolineare l'utilizzo di teorie statistiche, e formule di estrapolazione dei dati nel controllo continuo del corretto funzionamento dei server e delle applicazioni su essi funzionanti. Si riporta come esempio un controllo dell'occupazione di CPU. Infine si può valutare il successo o il

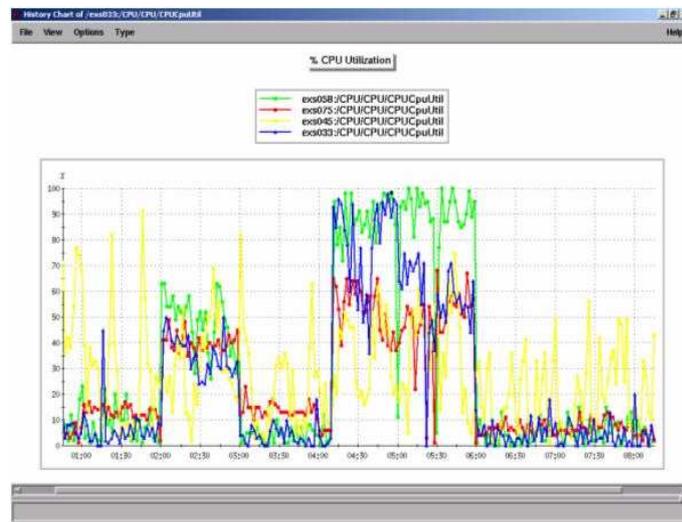


Figura 10: Occupazione di CPU su alcuni server sia di front-end che di back-end

fallimento della realizzazione di un progetto di questo tipo, in particolare un portale di informazione quale può essere www.rai.it, calcolando le page views, o ancora meglio gli hits dei banner pubblicitari, utili per quantificare il business del progetto. Anche in questo caso, sono necessarie conoscenze matematico-statistiche.

Riferimenti bibliografici

- [1] www.rai.it