

# 1 Lezione del 23-09-24

## 1.1 Introduzione

Il corso di progettazione web tratta i seguenti argomenti:

- Il linguaggio HTML 5.0
- I Cascading Style Sheet (CSS)
- Programmazione lato client: JavaScript (ECMAScript)
- Programmazione lato server: PHP
- Il protocollo HTTP

## 1.2 Storia del Web

Il Web, o World-Wide Web, è un'applicazione di Internet (come lo sono la posta elettronica, il File Transfer Protocol (FTP), ecc...).

### 1.2.1 Reti a commutazione

Le reti telefoniche (vecchio stile) forniscono un buon modello per il funzionamento di Internet. Le vecchie reti telefoniche erano reti a **commutazione di circuito**, ergo stabilivano una connessione fra due utenti attraverso un circuito fisico. Il filo di rame fisico che collegava i due utenti era riservato a loro, e non poteva servire nessun'altro durante il tempo della chiamata. Questo sistema aveva delle chiare problematiche:

- La connessione va stabilita e mantenuta per la durata della chiamata;
- C'è difficoltà a mantenere più chiamate simultaneamente;
- Si spreca banda in quanto anche i silenzi vengono trasmessi.

Un'alternativa è una rete a **commutazione di pacchetto**. La rete ARPANET non usava una rete a commutazione di circuito, ma una rete a commutazione di pacchetti. In una rete di questo tipo, il messaggio viene diviso in pacchetti, che possono prendere strade diverse per raggiungere il destinatario. Un pacchetto contiene informazioni sul mittente, il destinatario, e l'ordine del pacchetto stesso nel messaggio trasmesso.

Ad esempio, se Aldo vuole inviare a Bruno il messaggio "Didgeridoo è una parola difficile da dire", questo messaggio potrebbe essere diviso nei pacchetti:

- Aldo | Bruno | 1 | "Didgeridoo"
- Aldo | Bruno | 2 | "è una parola"
- Aldo | Bruno | 3 | "difficile da dire"

e questi pacchetti tramessi attraverso la rete dal mittente al destinatario.

Le reti a commutazioni di pacchetti sono più efficienti in quanto:

- Sono più robuste;
- Non richiedono la totalità delle risorse di trasferimento finché si mantiene la conversazione.

Internet nasce dal progetto ARPANET, e adotta questo tipo di design.

### 1.2.2 Protocollo TCP/IP

Uno dei primi problemi sorti nella progettazione di Internet è stato quello di unificare le regole di dialogo fra più reti. Una prima soluzione a questo problema, sorta nel 1981, è stata il Transmission Control Protocol / Internet Protocol (TCP/IP).

### 1.2.3 Web

L'invenzione del Web si attribuisce solitamente a Tim Berners-Lee. Dalla sua concezione iniziale ad oggi, le caratteristiche fondamentali del sistema sono rimaste invariate:

- Un Unique Resource Locator (URL) per identificare e locare risorse sul Web;
- Il protocollo HTTP per descrivere come le richieste e le risposte operano;
- Un programma, chiamato **server**, che risponde alle richieste HTTP;
- Il linguaggio HTML per pubblicare documenti;
- Un programma, chiamato **browser** (che fa la parte del **client**), che crea richieste HTTP dagli URL, e visualizza l'HTML che riceve.

Poco dopo la sua fondazione, Berners-Lee aiutò a fondare il Word Wide Web Consortium (W3C), che ha facilitato la crescita della rete attraverso la creazione di standard. Questo è stato permesso dalla decisione del CERN di rendere tutto il codice sorgente e i protocolli web sviluppati in questo periodo, effettivamente liberi.

Alcuni dei vantaggi delle applicazioni Web sono:

- Accessibilità da parte di ogni computer connesso a Internet;
- Utilizzabile con più browser e sistemi operativi;
- Più semplici da aggiornare;
- Archiviazione dei dati centralizzata.

alcuni contro, invece, sono:

- Richiesta di avere una connessione internet attiva.
- Problemi di sicurezza, dati dalla trasmissione di dati sensibili attraverso Internet;
- Problemi di inaderenza agli standard di alcuni siti su alcuni browser;
- Accesso limitato al sistema operativo.

### 1.2.4 Intranet

Un'intranet è una rete internet limitata ad un'organizzazione o un'azienda. A differenza delle reti internet, le reti intranet sono private e limitate agli utenti che le usano (impiegati o clienti). Il controllo sulle connessioni in entrata e in uscita di una rete internet o intranet è solitamente eseguito da un **firewall**.

### 1.2.5 Siti Web statici

Un sito web statico è formato da pagine HTML che appaiono identiche a tutti gli utenti in tutti i momenti.

Storicamente, questi contenuti venivano mantenuti dal cosiddetto webmaster, che doveva solo modificare il codice in HTML e, eventualmente, immagini o altri contenuti multimediali.

Il principio di funzionamento di un sito statico è il seguente:

- L'utente (il client) richiede una pagina al server;
- Il server risponde alla richiesta erogando la pagina desiderata;
- L'utente riceve la pagina e la visualizza.

### 1.2.6 Siti Web dinamici

Dopo alcuni anni dall'invenzione del Web, alcuni siti iniziarono a diventare più complicati e ad usare programmi che giravano sui web server per creare contenuti in maniera dinamica.

Secondo questo modello, il principio di funzionamento di un sito statico è il seguente:

- L'utente (il client) richiede una pagina al server;
- Il server risponde alla richiesta erogando la pagina desiderata e eventualmente riempiendola di informazioni generate da script, programmi, ecc... eseguiti sul momento;
- L'utente riceve la pagina e la visualizza.

### 1.2.7 Web 2.0

Web 2.0 è una buzzword che si riferisce principalmente a due cose:

- Per gli utenti, Web 2.0 significa applicazioni interattive dove si possono consumare ma anche creare nuovi contenuti;
- Per gli sviluppatori, Web 2.0 significa un nuovo modo di creare siti Web dinamici: la logica di programmazione si spostava dal server al client, attraverso il linguaggio Javascript. Questo introduceva diversi problemi per quanto riguarda l'esecuzione di comunicazioni asincrone.

## 1.3 Protocolli Internet

Un protocollo è una serie di regole che vengono usate da più calcolatori che comunicano fra loro. I protocolli TCP/IP vennero inizialmente pensati come uno stack di quattro livelli, anche se alcune astrazioni dividono questo stack ulteriormente in quattro o cinque livelli. Vediamo quindi i quattro livelli fondamentali, dal basso verso l'alto:

1. **Link Layer:** si occupa della trasmissione fisica di bit di informazione.

Esempio: MAC;

2. **Internet Layer:** stabilisce connessioni, routing e indirizzamento.  
Esempio: **IPv4, IPv6;**
3. **Transport Layer:** si assicura che le trasmissioni arrivino in ordine e senza errori.  
Esempio: **TCP, UDP;**
4. **Application Layer:** protocolli che permettono alle applicazioni di interagire con il Transport Layer.  
Esempio: **HTTP, FTP, POP.**

### 1.3.1 Link Layer

Il link layer permette il controllo sull'hardware che stabilisce un collegamento con i mezzi di trasmissione cablata (vedi Ethernet) o wireless (vedi Wi-Fi), ovvero il Media Access Control (MAC), e fornisce gestione di errori e di flusso, ovvero il Logical Link Control (LLC).

### 1.3.2 Internet Layer, protocollo IP

Il protocollo IP identifica destinazioni attraverso indirizzi IP. Ogni dispositivo connesso a internet ha un indirizzo IP proprio, ovvero un codice di 32 bit che lo identifica univocamente.

In particolare, gli indirizzi IPv4 sono quelli del protocollo TCP/IP originale. La rappresentazione comune degli indirizzi IPv4 (192.168.0.1) racchiude fra punti interi di 8 bit per volta dei 32 dell'indirizzo.

Gli indirizzi IPv6 appartengono ad un protocollo aggiornato, attualmente in corso di adozione, che adottano 8 interi da 16 bit. Solitamente gli interi del protocollo IPv6 vengono scritti in esadecimale.

L'indirizzo IP viene solitamente assegnato da un Internet Service Provider (ISP). Su reti locali, più computer possono condividere un'indirizzo IP, attraverso il NAT. NAT sta per Network Address Translation. All'interno della rete locale, ogni dispositivo usa un indirizzo diverso e privato (10.0.0/24). In uscita dalla rete locale, tutti i datagrammi hanno lo stesso indirizzo NAT IP di mittente, e diversi numeri di port. Questo è permesso da un meccanismo interno al router NAT che converte automaticamente indirizzi IP e numeri di porta in indirizzi IP locali, e viceversa.

### 1.3.3 Transport Layer, protocollo TCP

L'IP ha diverse limitazioni:

- La consegna dei pacchetti non è garantita;
- I pacchetti potrebbero arrivare in disordine;
- La comunicazione avviene fra dispositivi;
- Nessun controllo sul flusso o sulle congestioni.

Il Transport Layer si assicura che ogni trasmissione arrivi in ordine e senza errori, attraverso dei meccanismi di sicurezza:

1. Prima i dati sono divisi in pacchetti formattati secondo il protocollo TCP;

2. In seguito, ogni pacchetto viene restituito al mittente attraverso l'ACK (acknowledge).

In questo modo, il mittente sa se i suoi dati sono stati effettivamente ricevuti, e può reinviare i pacchetti perduti.

Il protocollo TCP fornisce una serie di altre funzionalità: la gestione del flusso e della congestione. Questi meccanismi riguardano la diminuzione della frequenza di trasmissione in caso di sovraccarico del destinatario.

### 1.3.4 Application Layer

L'application layer è un livello di astrazione. Fanno parte di questo livello il protocollo di trasferimento di pagine web, ovvero l'Hypertext Transfer Protocol (HTTP), i protocolli di trasmissione di file, ovvero il File Transfer Protocol (FTP), o i protocolli di trasmissione di posta elettronica, come il Post Office Protocol (POP), fra gli altri.

## 1.4 Modello client-server

Nel modello client-server, i dispositivi si dividono in client e server. Il server è attivo e pronto ad erogare servizi, mentre il client si occupa di fare richieste a cui il server dovrà adempire. Questo meccanismo forma il cosiddetto request-response loop.

### 1.4.1 L'alternativa peer-to-peer

In una rete peer-to-peer non esiste un'unità centrale (il server), ma ogni nodo è in grado di inviare e ricevere indipendentemente dagli altri (peer).

Questo sistema è molto comodo per la trasmissione in massa di dati senza la possibilità di arrestare il processo a partire da una sola macchina (vedi Napster).

### 1.4.2 Tipi di server

Prima, abbiamo mostrato il server come una macchina singola. In verità, il server può essere formato da più macchine che si dividono il carico delle richieste dei client, e che rispondono dallo stesso URL.

Un'altra alternativa sono le cosiddette **server farm**. Una server farm distribuisce il carico delle richieste (attraverso i **load balancer**). Le server farm forniscono inoltre **fail-over redundancy**, ovvero la sicurezza che il servizio non si arresti anche nel caso di fallimento di singole macchine.

Le server farm sono tipicamente collocate all'interno di specifiche strutture dette **data center**. Non è inusuale archiviare lo stesso sito su più macchine, in parallelo. Può accadere anche il contrario: più pagine possono essere archiviate sullo stesso server.

## 1.5 Nozioni sull'infrastruttura

Oltre ai server, che abbiamo visto, altri componenti vanno a formare la rete Internet: innanzitutto il modem (che può essere di tipo Digital Subscriber Line (DSL) se sfrutta la linea telefonica), che funge da ponte fra la rete interna a aziende o abitazioni, e la rete esterna, gestita da un Internet Service Provider (ISP) o un'altra autorità.

Da qui il router, che si occupa di indirizzare i pacchetti secondo determinate tabelle chiamate routing tables, che includono per tutti gli IP di destinazione noti il prossimo passo (hop) da effettuare, ovvero il prossimo IP a cui inviare il pacchetto.

I cavi ottici arrivano infine all'head-end dell'ISP, dove i dati vengono gestiti da apparecchiature come il Cable Modem Termination System (CMTS), o il Digital Subscriber Line Access Multiplexer (DSLAM) per le trasmissioni DSL.

Da qui i dati potrebbero dover viaggiare oltre, tipicamente arrivando ad altre reti, che possono avere estensione anche locale. Per rendere più veloci le trasmissioni fra più reti, oggi si adoperano gli Internet Exchange Point (IX o IXP), che permettono a più reti di unirsi in strutture condivise.

Spesso, soprattutto di recente, le grandi compagnie decidono di installare la loro infrastruttura (server e data center) il più vicino possibile agli IXP, per velocizzare la trasmissione portando i dati più vicini agli utenti.