

Introducere în Internet

Mihai Gabroveanu

Istoric

- n **1969** – Departamentul de Aparare al Statelor Unite inițiază proiectul ARPAnet (Advanced Research Projects Agency network)
 - .. Scop: realizarea unei rețele de dimensiuni mari
- n **1972** – Prima demonstrație publică a rețelei ARPAnet – transmiterea de mesaje utilizând poșta electronică
- n **1973** – DARPA – rețele interconectate; Robert Metcalf (Harvard) dezvoltă tehnologia Ethernet care permite transferul de date pe cablu coaxial

Ce este Internetul?

- n **Rețea de calculatoare** = ansamblu de calculatoare capabile să comunice între ele prin intermediul unor canale de comunicație
- n **Internetul** = rețea de rețele, prin care utilizatorii unui calculator sau ai unei rețele pot accesa diverse sisteme aparținând altor rețele.

Istoric (cont.)

- n **1974** – Vincent Cerf și Robert Kahn propun protocolul de comunicare TCP (Transmission Control Protocol)
- n **1978** – suita de protocoale TCP/IP, standardizată via documentele RFC (Request For Comments)
 - .. <http://www.ietf.org/rfc.html>
- n **1980** – NFS (National Science Foundation) reproiectează modelul ARPAnet, noua arhitectură primind numele de **Internet**
- n Servicii: email, transfer de fișiere (FTP), Gopher, grupuri de știri, IRC (1988), World Wide Web (1990),...

World Wide Web (WWW)

- n 1990 – Tim Berners-Lee inventează World Wide Web (WWW sau WEB) la CERN.
- n **WEB** = *sistem de distribuție locală sau globală a informațiilor hipermedia*
- n Web pune la dispoziție un sistem global de standarde de comunicare multimedia, informațiile fiind organizate asociativ și fiind distribuite în funcție de cererile utilizatorilor
- n Utilizează modelul client/server



World Wide Web Consortium

- n 1994 – se constituie consorțiul World Wide Web Consortium (W3C) <http://www.w3.org/> condus de Tim Berners-Lee
- n Scop: elaborarea, dezvoltarea și menținerea de standarde pentru spațiul WWW

World Wide Web (WWW) (cont.)

- n Obiectivele principale ale sistemului WEB propus de Tim Berners-Lee:
 - .. accesul la distanță într-o rețea, acceptându-se diverse tipuri de conexiuni;
 - .. accesul la o multitudine de surse de informații stocate într-o varietate de forme (text, grafică, sunet etc.) cu ajutorul unor tipuri diferite de sisteme de operare situate pe diferite platforme hardware;
 - .. capacități de extindere și dezvoltare în ceea ce privește aria de acoperire prin creșterea numărului de legături, fără existența unui control central.

Clienți și servere Web

- n Tim Berners-Lee împreună cu echipa realizează primele patru componente cheie necesare serviciului Web:
 - .. serverul web
 - .. browserul
 - .. protocolul HTTP
 - .. HTML – limbaj de descriere a hipertextului

Server WEB

- n **Server WEB** = program ce rulează pe un calculator conectat la Internet și furnizează clienților la cerere diverse resurse Web (documente).
- n Tim Berners-Lee în 1990 concepe primul **server Web** rulând pe calculatoare NeXT.
- n *Apache Web Server* – cel mai popular server Web. Prima versiune a fost lansată în 1995

Browser WEB (cont.)

- n 1990 - Tim Berners-Lee dezvoltă primul browser - **WorldWideWeb** (redenumit ulterior în Nexus)
- n 1991 - Nicola Pellow, student intern la CERN, dezvoltă **Line Mode Browser** (browser in mod text)
- n 1994 - Mark Andreessen lansează unul dintre cele mai populare browsere grafice, **MOSAIC**, la National Center for Super Computing Applications (NCSA)
- n Cele mai populare browsere Web: Google Chrome, Safari, Firefox, Internet Explorer, Opera

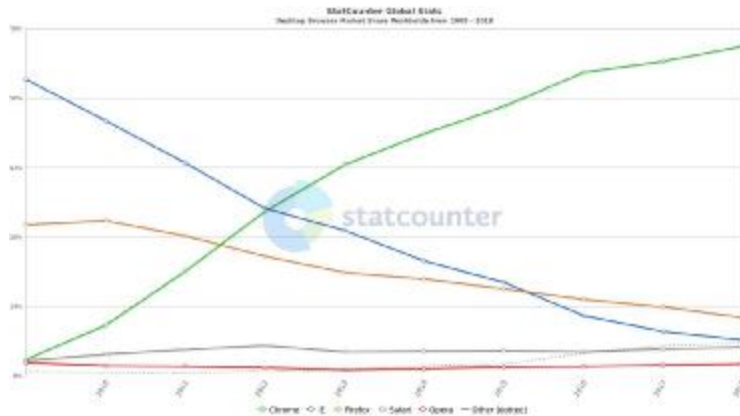
Browser WEB

- n **Browser WEB** = program software (client) ce se permite utilizatorilor să se conecteze la un server Web în vederea explorării resurselor găzduite de acesta (text, grafică, video, etc.);
- n Protocolul utilizat: **HTTP**
- n Resursele sunt identificate printr-un Uniform Resource Locator (URL)

WorldWideWeb – Primul Browser Web



Browsere Web. Popularitate



Introducere în Internet

13

Protocolul HTTP

- n **Protocol** = set de reguli de comunicare între client și server
- n **Protocolul HTTP (HyperText Transfer Protocol)** = set de reguli de comunicare între un server și browser web.
 - .. Dezvoltat în 1990 de Tim Berners-Lee

Introducere în Internet

15

Browsere Web. Caracteristici

- n Posibilitatea de a realiza interogări multiple către servere
- n Asigurarea securității transmișorilor de date
- n Stabilirea de liste a site-urilor web favorite
- n Memorarea istoricului navigării (history)
- n Posibilitatea de a folosi mai multe ferestre de navigare
- n Suport pentru cadre (frame-uri)
- n Asigurarea suportului pentru diverse limbaje de programare folosite la realizarea paginilor Web dinamice, foi de stiluri (CSS).

Introducere în Internet

14

Site-uri Web

- n **Site WEB** (eng. *website*) = o colecție de pagini statice (documente) WWW înrudite ale unei organizații, companii sau persoane, conectate între ele prin intermediul hiperlinkurilor găzduite de un calculator pe care rulează un server Web
- n Serverul ce găzduiește paginile Web este conectat de regula la Internet, paginile fiind accesibile prin intermediul numelui de domeniu asignat acestuia (ex: <http://inf.ucv.ro>)
- n Un server WEB poate gestiona (găzdui) simultan mai multe site-uri web (*virtual hosting*)

Introducere în Internet

16

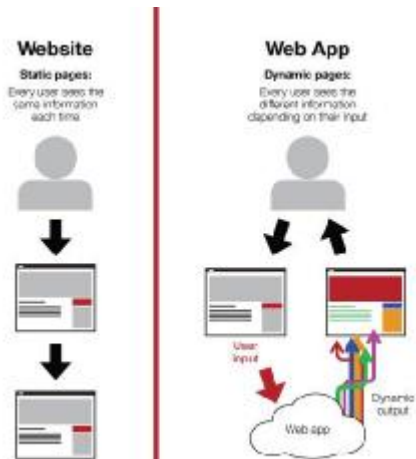
Aplicație WEB

- n O **aplicație web** (eng. *Web application*) reprezintă o colecție interconectată de pagini Web cu un conținut *dinamic* menită a oferi o funcționalitate specifică utilizatorilor.
- n Aplicația Web este accesibilă via unui web browser, paginile dinamice fiind obținute prin execuția unor “programe” pe server

Localizarea resurselor WEB

- n Resursa Web = orice entitate având o identitate (document, imagine, serviciu, colecție de alte resurse)
- n Localizarea resurselor web în Internet se realizează prin Identificatori Uniformi de Resurse (**URI** – *Uniform Resource Identifier*)

Website vs. Web Application



URI – *Uniform Resource Identifier*

- n Identificatorii uniformi de resurse se împart în două categorii:
 - URL – Uniform Resource Locator
 - n **localizează** resursa prin descrierea mecanismului de acces
 - n ex: <http://inf.ucv.ro/~mihaiug/index.html>
 - URN – Uniform Resource Name
 - n identifică resursele prin nume, într-o formă persistentă, chiar dacă resursa a dispărut
 - n ex: urn:isbn:0451450523

URL – Uniform Resource Locator

n Sintaxa

scheme://host[:port#]/path/. . ./[:url-params][?query-string][#anchor]

n Exemplu

<http://inf.ucv.ro:8080/optionale?action=view&idDisciplina=C102>

n Porturi implicite

- HTTP: 80
- HTTPS: 443
- FTP: 21

URL – Uniform Resource Locator

host

Identifică serverul web ce găzduiește resursele

port

Identifică portul pe care rulează serverul web. Opțional

path

Identifică calea pe server unde se găsește resursa

query-string

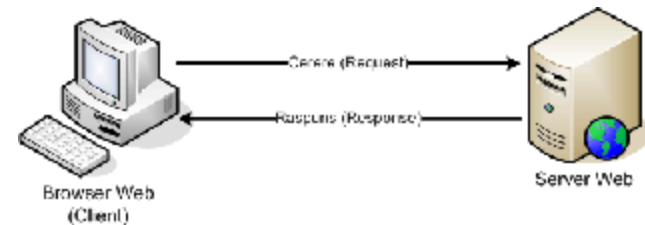
parametri utilizați în generarea pagini web. Opționali

URL – Uniform Resource Locator

scheme

- Constă dintr-o secvență de caractere care începe cu o literă urmată de orice combinație de litere, cifre, plus ("+"), punct ("."), sau minus ("-").
- case-insensitive
- oficiale:
 - [file](#), [ftp](#), [http](#), [https](#), [ldap](#), [mailto](#), [telnet](#), [urn](#), etc.
 - ex: <http://inf.ucv.ro>, <mailto:mihaiug@central.ucv.ro>
- neoficiale:
 - fish, skype, callto, view-source, ymsgr, etc.
 - ex:
<skype:username?call>
view-source:http://en.wikipedia.org/wiki/URI_scheme

Protocolul HTTP



n HTTP = HyperText Transfer Protocol

n Paradigma Client/Server

n Patternul Request/Response

Protocolul HTTP (cont.)

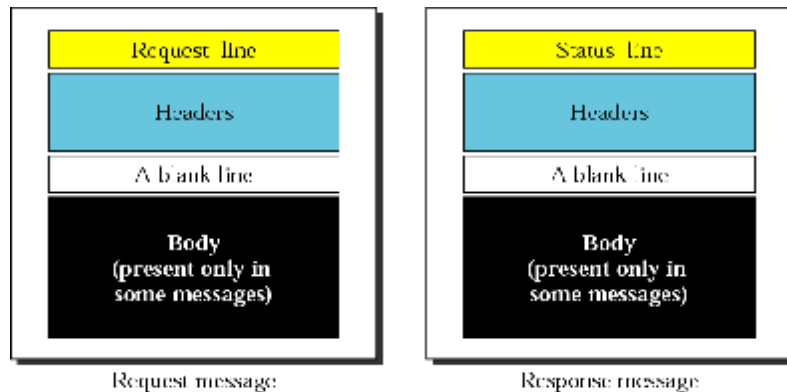
- n Standardizat de Internet Engineering Task Force (IETF) prin RFC 2616
- n Este un protocol stateless (nu se menține o stare a sesiunii de la o cerere la alta) și asincron (un document HTML este încărcat de browser, de îndată ce părți ale acestuia sunt disponibile)
- n Ultima versiune disponibilă este HTTP/2.0 (RFC 7540, Mai, 2015), însă versiunea HTTP/1.1 (RFC 7230-7235, Iunie, 2014) este cea mai utilizată
- n Rulează deasupra protocolului TCP. Portul standard pe care rulează fiind portul 80

HTTP Request

n Sintaxa unei cereri HTTP

METHOD /path-to-resource HTTP/version-number
Header-Name-1: value
Header-Name-2: value
[optional request body]

Protocolul HTTP – Sintaxa Mesajelor



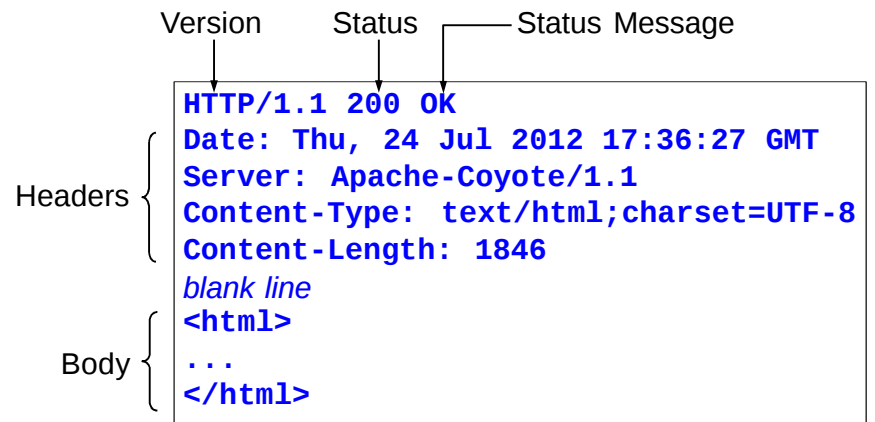
HTTP Request- Exemplu



HTTP Request (cont.)

- n METHOD poate să aibe una din următoarele valori:
 - .. **GET** – Solicită un document/resursă de pe server
 - .. **HEAD** – Solicită informații despre un document/resursă de pe server
 - .. **POST** – Metodă utilizată pentru a transmite date de intrare către server
 - .. **PUT** – Metodă utilizată pentru a depune un document/resursă pe server
 - .. **DELETE, OPTIONS, TRACE, CONNECT**

HTTP Response - Exemplu



HTTP Response

- n Sintaxa unui răspuns HTTP

HTTP/version-number status-code message

Header-Name-1: value

Header-Name-2: value

[response body]

Răspuns HTTP (cont.)

- n Status code
 - .. 200 OK
 - .. 301 Moved Permanently
 - .. 400 Bad Request
 - .. 401 Unauthorized
 - .. 403 Forbidden
 - .. 404 Not Found
 - .. 500 Internal Server Error
 - .. 503 Service Unavailable

Headere HTTP

- n Server Tipul serverul
- n Date Data și ora în format GMT)
- n Content-Length Numarul de octeți
- n Content-Type Tipul MIME al continutului
- n Content-Language Limba continutului
- n Last-Modified Data și ora ultimei modificări
- n Expires Data când fișierul devine invalid

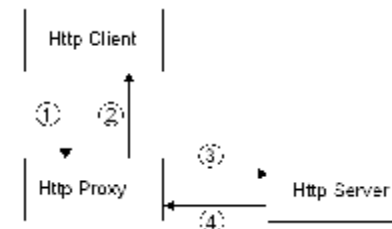
HTTP Proxy Server

- n Un server (aplicatie) ce acționează ca un intermediar pentru cererile HTTP primite de la client.

Headere HTTP (cont.)

- n Host Numele Hostului
- n Accept MIME Tipul acceptat ca raspuns (ex: text/html)
- n User-Agent Identificatorul browserului (ex: Mozilla/4.0)
- n Referer Pagina de unde s-a initiat cererea

HTTP Proxy Server (cont.)



- n Clientul se conectează la serverul proxy, solicită o resursă prin intermediul protocolului HTTP, proxy-ul analizează cererea și solicită resursa de pe serverul HTTP pe care este localizată, o aduce la nivelul serverului proxy și apoi o servește clientului



Utilizări ale unui HTTP Proxy Server

- n Caching - Resursele solicitate frecvent sunt stocate temporar pe serverul proxy pentru a reduce traficul/încărcarea
- n Protejarea identității solicitantului
- n Filtrarea conținutului