ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

IT4552 – Web programming

Chapter 1-1.
Introduction to Web Programming

1

Content

1. The Internet and WWW
2. Uniform Resource Identifier (URI)
3. Web Application model

2

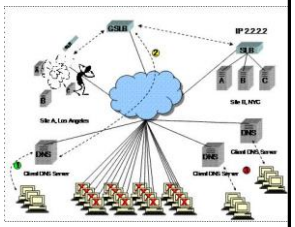
Content

1. **The Internet and WWW**
2. Uniform Resource Identifier (URI)
3. Web Application model

3

1.1. The Internet

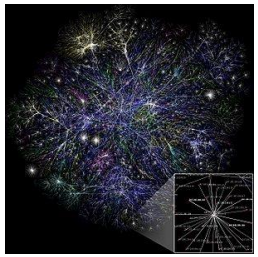
- A network of networks
- Began in 1969 as ARPAnet (Advanced Research Projects Agency)
- No central authority and thus impossible to state the precise size



4

1.1. The Internet (2)

- Free access to central servers that allow machines and people to locate other machines by their **Internet address.**
  - e.g. 100.99.88.32



5

1.2. The World Wide Web

- Developed by Tim Berners-Lee at CERN in 1990
- The idea of documents that contain hyperlinks to other documents on the Internet
- W3 or Web for short

NeXT Computer
The first Web server

6

1.2. The World Wide Web (2)

- World Wide Web
  - a system of interlinked hypertext documents accessed via the Internet
- HyperText Markup Language (HTML)
  - document layout language for all Web Documents
- HyperText Transfer Protocol (HTTP)
  - allows any machine to load a document via a hyperlink from any other machine

7

1.3. Web page or Webpage

- A Web document
  - a document or resource of information that is suitable for the WWW and can be accessed through a web browser and displayed on a computer screen
- Usually in HTML or XHTML format
  - XHTML (Extensible HTML): Intersection between HTML and XML
- Requested and served from web servers using HTTP.

8

1.4. Web site or Website

- A collection of related web pages, images, videos or other digital assets that are addressed with a common domain name or IP address in an Internet Protocol-based network
- Hosted on at least one web server, accessible via the Internet or a private local area network.

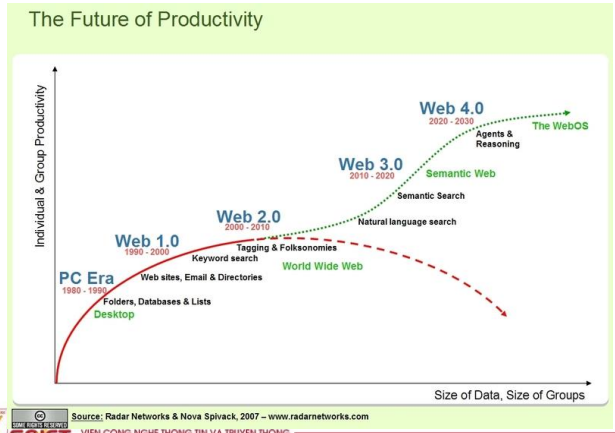
9

Web ≠ Internet

- Internet: a physical network connecting millions of computers using the same protocols for sharing/transmitting information (TCP/IP)
  - in reality, the Internet is a network of smaller networks
- World Wide Web: a collection of interlinked multimedia documents that are stored on the Internet and accessed using a common protocol (HTTP)
- Key distinction: Internet is hardware; Web is software along with data, documents, and other media
- Many other Internet-based applications exist e.g., email, telnet, ftp, usenet, instant messaging services, file-sharing services, …

10

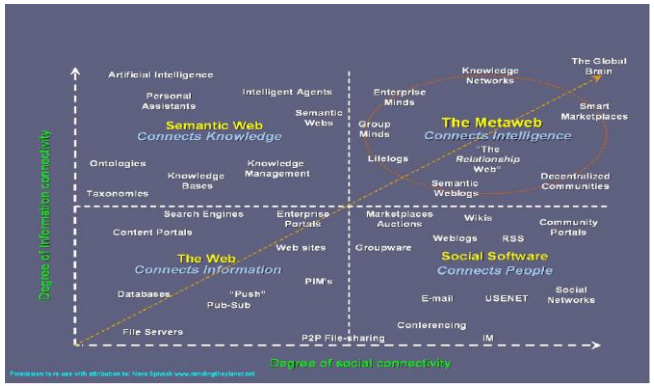
Development of the Web



Source: Radar Networks & Nova Spivack, 2007 – www.radarnetworks.com

11

Development of the Web (2)



12

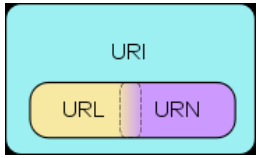
Content

1. The Internet and WWW
2. **Uniform Resource Identifier (URI)**
3. Web Application model

13

2.1. URI (Uniform Resource Identifier)

- A string of characters used to identify or name a resource on the Internet



- Classification
  - URN: a person's name
  - URL: that person's street-address

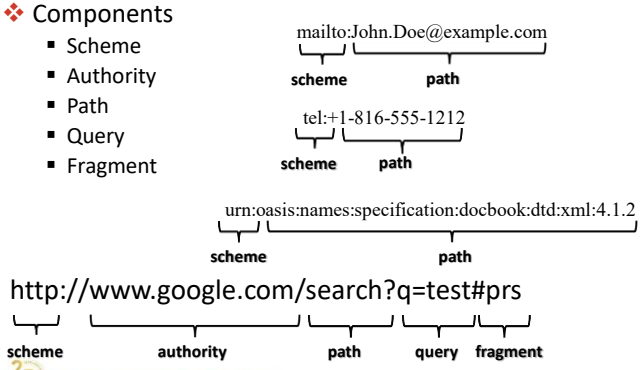
  →URN defines an item's identity

  →URL provides a method for finding it

14

URI Generic Syntax

- RFC 3986
- Components
  - Scheme
  - Authority
  - Path
  - Query
  - Fragment



15

2.2. Uniform Resource Locator (URL)

- Created in 1990 by Tim Berners-Lee as part of the URI
- Specifies where an identified resource is available and the mechanism for retrieving it
- Separated by dots, and the file path, separated by slashes, as a coherent hierarchical path

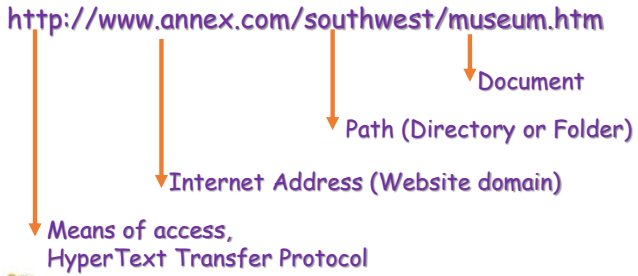
16

2.2. Uniform Resource Locator (URL) (2)

- Syntax

resource_type://domain:port/filepathname?query_string#anchor

- Example



17

2.3. Uniform Resource Name (URN)

- Globally unique and persistent name of a resource on the Internet
- Syntax: <URN> ::= "urn:" <NID> ":" <NSS>
  - <NID> is the Namespace Identifier
  - <NSS> is the Namespace Specific String
- Example
  - urn:isbn:0451450523
    - The URN for "The Last Unicorn", identified by its book number.
  - urn:isan:0000-0000-9E59-0000-O-0000-0000-2
    - The URN for "Spider-Man (film)", identified by its audiovisual number.
  - …

18

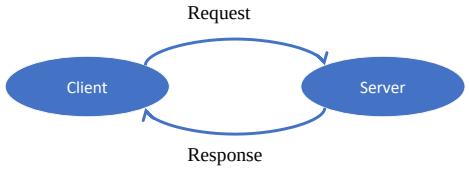
Content

1. The Internet and WWW
2. Uniform Resource Identifier (URI)
3. **Web Application model**

19

Client Server Model

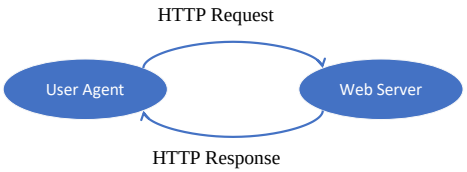
- A simple network model
- Used by various network applications



20

Client Server Model (Web)

- Client: User Agent
- Server: Web server



21

Web Browsers

- Primary tasks:
  - Convert web addresses (URL's) to HTTP requests
  - Communicate with web servers via HTTP
  - Render (appropriately display) documents returned by a server

22

Web Clients

- Many possible web clients:
  - Text-only "browser" (lynx)
  - Mobile phones
  - Robots (software-only clients, e.g., search engine "crawlers")
  - etc.

23

Web Servers

- Basic functionality:
  - Receive HTTP request via TCP
  - Map host header (domain name) to specific virtual host (one of many host names sharing an IP address)
  - Map Request-URI to specific resource associated with the virtual host
    - File: Return file in HTTP response
    - Program: Run program and return output in HTTP response
  - Map type of resource to appropriate MIME type and use to set Content-Type header in HTTP response
  - Log information about the request and response

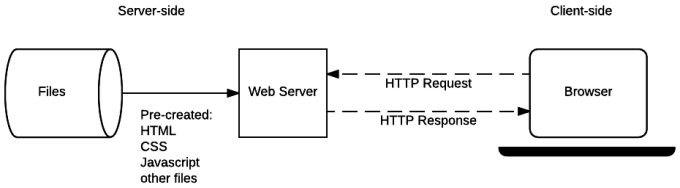
24

Web Servers

- httpd: UIUC, primary Web server c. 1995
- Apache: “A patchy” version of httpd, now the most popular server (esp. on Linux platforms)
- IIS: Microsoft Internet Information Server
- Tomcat:
  - Java-based
  - Provides container (Catalina) for running Java servlets (HTML-generating programs) as back-end to Apache or IIS
  - Can run stand-alone using Coyote HTTP front-end

25

Web Application Evolution – Static

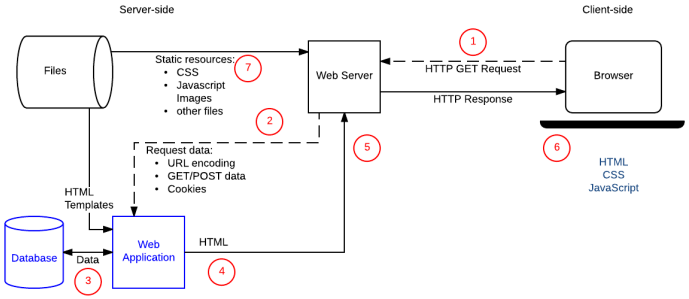


- Organizations want to make their information available to as many people in the world as possible
- This can be achieved by using the Web, delivering the information as static HTML pages

https://developer.mozilla.org/en-US/docs/Learn/Server-side

26

Web Application Evolution – Dynamic

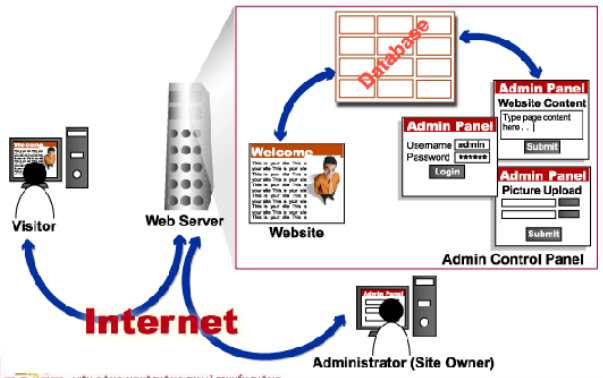


- Requests for dynamic resources are instead forwarded (2) to server-side code (shown in the diagram as a Web Application).
- For "dynamic requests" the server interprets the request, reads required information from the database (3), combines the retrieved data with HTML templates (4), and sends back a response containing the generated HTML (5,6).

https://developer.mozilla.org/en-US/docs/Learn/Server-

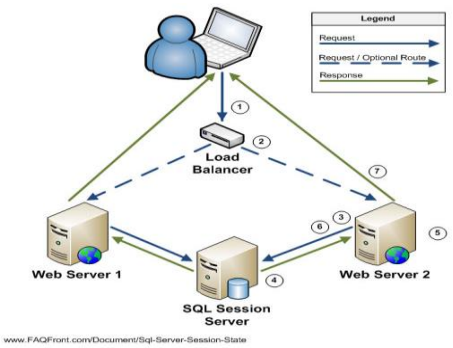
27

Web Application – Dynamic



28

More than one Web server?



29

Hypertext Transport Protocol (HTTP)

- HTTP is based on the request-response communication model:
  - Client sends a request
  - Server sends a response
- HTTP is a stateless protocol:
  - The protocol does not require the server to remember anything about the client between requests.

30

Hypertext Transport Protocol (HTTP) (2)

- Normally implemented over a TCP connection (80 is standard port number for HTTP)
- Typical browser-server interaction:
  - User enters Web address in browser
  - Browser uses DNS to locate IP address
  - Browser opens TCP connection to server
  - Browser sends HTTP request over connection
  - Server sends HTTP response to browser over connection
  - Browser displays body of response in the client area of the browser window

31

HTTP Request

- Structure of the request:
  - start line
  - header field(s)
  - blank line
  - optional body

32

HTTP Request

- Start line
  - Example: GET / HTTP/1.1
- Three space-separated parts:
  - HTTP request method
  - Request-URI (Uniform Resource Identifier)
  - HTTP version
    - We will cover 1.1, in which version part of start line must be exactly as shown

33

HTTP Request

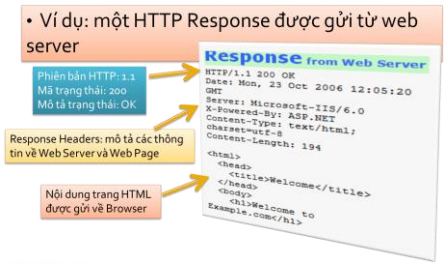
- Uniform Resource Identifier (URI)
  - Syntax: *scheme* : *scheme-depend-part*
    - Ex: In `http://www.example.com/`
    - the scheme is `http`
  - Request-URI is the portion of the requested URI that follows the host name (which is supplied the required Host header field)
    - Ex: `/` is Request-URI portion of `http://www.example.com/`

34

HTTP Response

- Structure of the response:
  - status line
  - header field(s)
  - blank line
  - optional body

- Ví dụ: một HTTP Response được gửi từ web server



35

HTTP Response

- Status line
  - Example: HTTP/1.1 200 OK
- Three space-separated parts:
  - HTTP version
  - status code
  - reason phrase (intended for human use)

36

HTTP Response

- Status code
  - Three-digit number
  - First digit is class of the status code:
    - 1=Informational
    - 2=Success
    - 3=Redirection (alternate URL is supplied)
    - 4=Client Error
    - 5=Server Error
  - Other two digits provide additional information
  - See http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html

37

HTTP Response

- Common header fields:
  - Connection, Content-Type, Content-Length
  - Date: date and time at which response was generated (required)
  - Location: alternate URI if status is redirection
  - Last-Modified: date and time the requested resource was last modified on the server
  - Expires: date and time after which the client's copy of the resource will be out-of-date
  - ETag: a unique identifier for this version of the requested resource (changes if resource changes)

38

Web Tools and Environments

- Content Management Systems (CMS)
  - Wordpress
  - Joomla
  - Drupal
- Text editors
  - Notepad
  - Text Wrangler
  - Sublime Text
- Integrated Development Environment (IDEs)
  - Eclipse
  - Aptana Studio
  - Cloud9
- WYSIWYG editors (drag and drop)
  - Google Web Designer
  - MS Visual Studio
  - Adobe Dreamweaver
- Image Editors
  - Illustrator
  - Photoshop
  - FireWorks

39

Web Tools and Environments

- Browsers
- https://caniuse.com/
  - Gives us up to date information on which features are currently supported by each browser

40

Web Tools and Environments

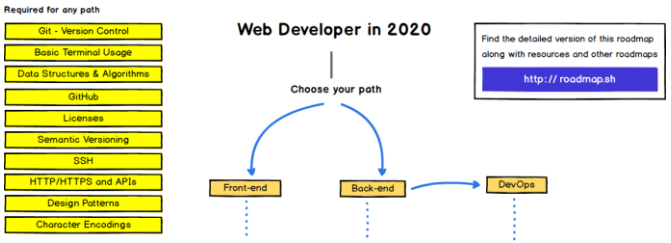
- Debugging
  - View Source
  - Firebug
  - Inspect Element

41

Question?

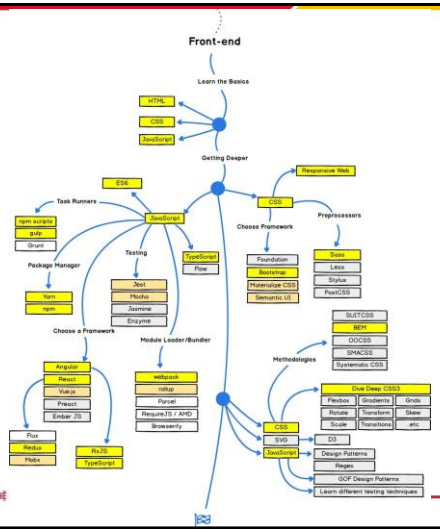
42

Web Developer Roadmap 2020

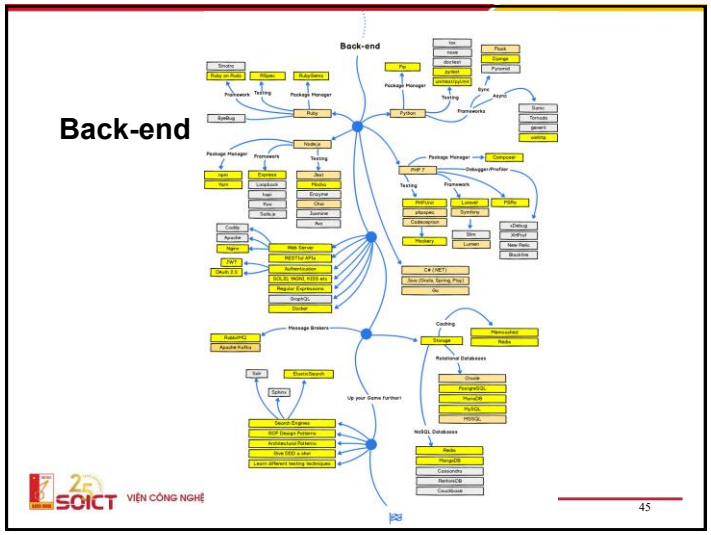


43

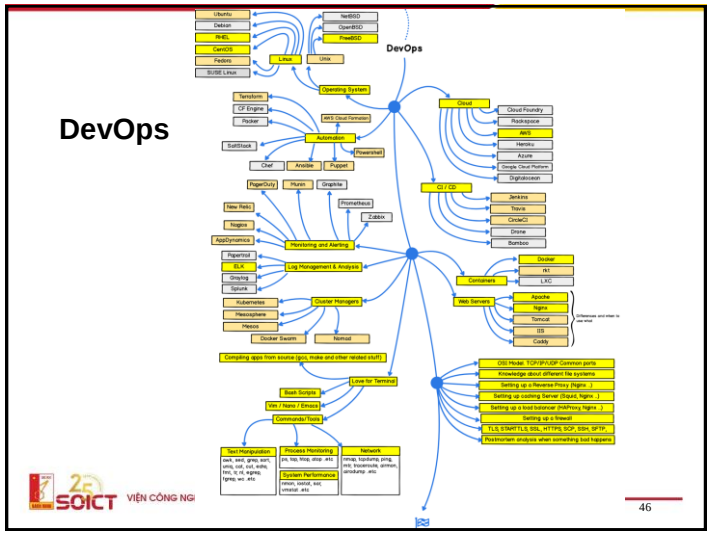
Front-end



44


Back-end
Back-end
VIỆN CÔNG NGHỆ
45

45


DevOps
DevOps
VIỆN CÔNG NGỆ
46

46