Web Essentials

February 8, 2012

Web Technologies

M.E. (CSE) Semester 2

R S Milton

Department of Computer Science and Engineering

Contents

Woe be to him that reads but one book.

George Herbert (1593–1633)

Books are a triviality. Life alone is great.

Thomas Carlyle (1795–1881)

The mere practical architect is not able to assign sufficient reasons for the forms he adopts; and the theoretic architect also fails, grasping the shadow instead of the substance. He who is theoretic as well as practical, is therefore doubly armed; able not only to prove the propriety of his design, but equally so to carry it into execution

---

Vitruvius (80–15 BC)

1. Text

1. Jeffrey Jackson

   Web Technologies, A Computer Science Perspective

   Pearson Education, 2005

2. Chist Bates

   Web Programming, Building Internet Applications, Second edition

   Wiley-dreamtech India Pvt Ltd

3. Uttam Roy

   Web Technolgies

   Oxford University Press, 2010

2. Syllabus

Unit I 9

Web essentials — clients — servers — communication; Markup languages: XHTML — simple XHTML pages; Style sheets — CSS

Unit II 9

Client side programming — JavaScript language — JavaScript objects — Host objects: Browsers and the DOM

Unit III 9

Server side programming — Java servlets – Basics — simple program — separating programming and presentation — ASP/JSP basics — ASP/JSP objects

— simple ASP/JSP pages

Unit IV 9

Representing Web data — data base connectivity — JDBC-Dynamic Web pages — XML — DTD — XML schema — DOM — SAX-Xquery

Unit V 9

Building Web applications — cookies — sessions — open source environment — PHP–MYSQL — case studies

3. Internet — information management technology

- Server
  - Computer wishing to provide information
  - Run server software
- Client
  - Computer wishing to access the information
  - Run client software
- Use a communication protocol on top of TCP/IP

4. Word wide web

- Several technologies — Usenet, FTP, IRC
- Web — one of the technolgies
- Client requests a document from a server
- Server returns the requested document
- Collection of web servers on the Internet
  - Provide information via HTTP
  - HTML documents

Advantages

- HTTP — generic protocol — flexible
- HTML — hyperlinks,page laypout, images

5. HTTP — Hypertext Transport Protocol

- How web clients and servers communicate
- Request-response model
- Client initiates: sends a request message to the server
- Server sends back a response message to the client
- Message format fixed
- TCP-style connection between client and server

6. Browse — what it means?

- Type the web site address in the location bar and press Enter
- Browser creates a HTTP request message
- Uses DNS to convert the web site name to an IP address
- Creates a TCP connection with the host at the IP address
- Sends the HTTP request message over the TCP connection
- Receives back a HTTP response message
- Displays the information contained in the message in the client area of the browser

7. HTTP message formats

Feature

Plain text — printable, readable form

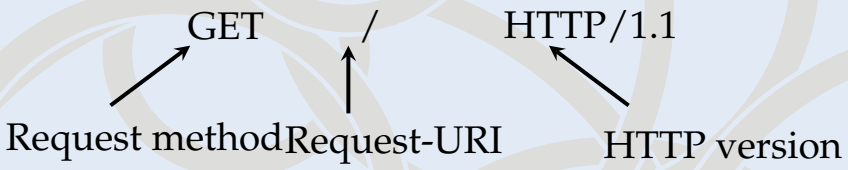
8. HTTP request message

Example

```
GET / HTTP/1.1
Host:  www.example.org
```

- Start line
- Header field(s)
- Blank line
- Message body (optional)

9. HTTP request message — Start line



- Request method
- Request-URI portion of web address
- HTTP version

10. Start line: Request method

Method	Reqeusts server to
GET	return the resource specified by the Request-URI as the response message body
POST	pass the request message body on as data to be processed by the resource specified by the Request-URI
HEAD	return the same HTTP header fields that would be returned if a GET method were used, but not return the message body

GET

- Type a URL in location bar
- Click on a link
- Downloads image

11. Request-URI, URI, URL

Uniform Resource Identifier — URI

> Scheme + Host + Request-URI

http://www.example.org/

> http:// + www.example.org + /

- ftp, telnet, mailto, https, file
- URL = URI when the scheme is `http`

12. MIME types

Content type

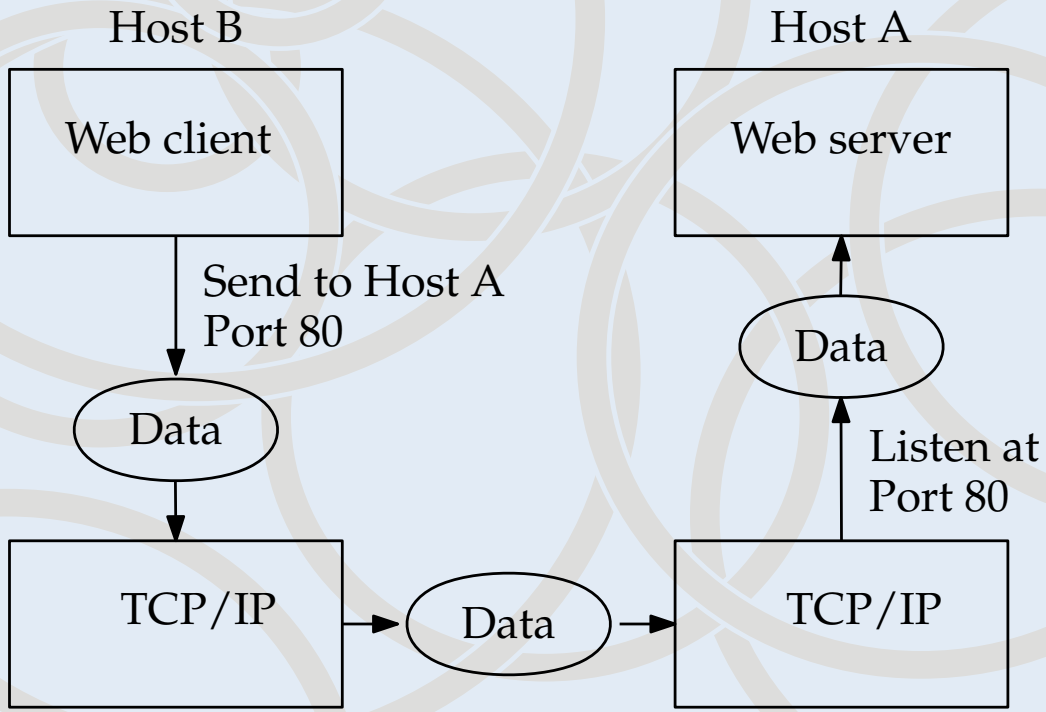
top-level type + subtype

Top-level content type	Document content
text	Displayable text
audio	Audio data; subtype defnes audio format
image	Image data; subtype defines format
video	Animated image, possibly with sync sound
application	To be processed by an application software

13. HTTP response message

- Status line
- Header field(s)
- Blank line
- Message body (optional)

14. TCP/IP protocol



- Internet connected computers

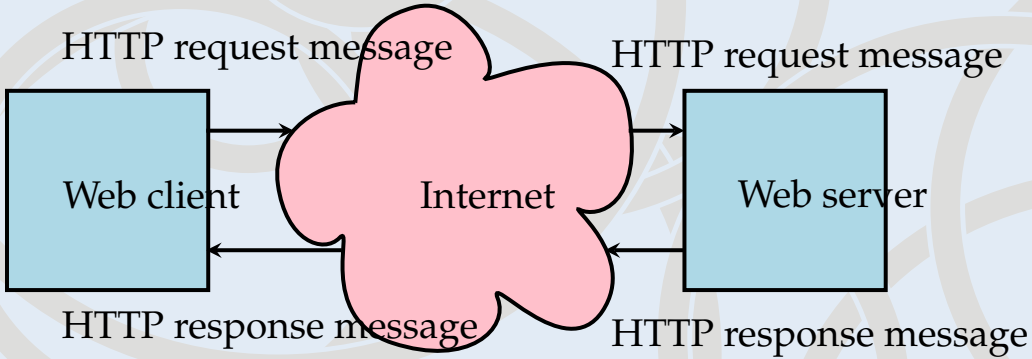
- Different hardware
- Different operating systems

- Communicate over TCP/IP protocol
- Client has to specify the server host (computer) IP address
  - Host name $\rightarrow$ Host IP address
- In the server host, several server processes may be running
- Client has to specify the port at which the server process is listening at the server host

15. World wide web

- Collection of web servers on the Internet
  - Provide information via HTTP
  - HTML documents
- Web clients can access the information

16. HTTP — Hypertext Transport Protocol



- Request-response model on top of TCP/IP
- Web clients and servers communicate using HTTP
- Web client sends a request message to the server
- Web server sends back a response message to the client
- Message format fixed

16.1. HTTP message format

HTTP request message	HTTP response message
Start line	Status line
Header fields	Header fields
Blank line	Blank line
Message body	Message body

```
milton@milton-laptop:~/course/web/notes$ telnet localhost 8080
Trying ::1...
Connected to localhost.
Escape character is '^]'.
GET / HTTP/1.1
Host: localhost

HTTP/1.1 200 OK
Server: Apache-Coyote/1.1
Accept-Ranges: bytes
ETag: W/"576-1265170439000"
Last-Modified: Wed, 03 Feb 2010 04:13:59 GMT
Content-Type: text/html
Content-Length: 576
Date: Wed, 03 Feb 2010 04:14:46 GMT

<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
   "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
<head>
    <title>Apache Tomcat</title>
</head>

<body>
<h1>It works !</h1>

<p>If you're seeing this page via a web browser, it means you've setup Tomcat successfully. Congratulations

<p>This is the default Tomcat home page. It can be found on the local filesystem at: <code>/var/lib/tomcat6/
</body>
</html>
```

17. HTTP request message

Status line

GET / HTTP/1.1

Start line : `GET / HTTP/1.1`

Header field : `Host: localhost`

Blank line :

Message body : Empty

17.1. HTTP request message

Status line : `HTTP/1.1 200 OK`

Header fields :

- `Server: Apache-Coyote/1.1`
- `Accept-Ranges: bytes`

- ETag: W/"576-1265170439000"
- Last-Modified: Wed, 03 Feb 2010 04:13:59 GMT
- Content-Type: text/html
- Content-Length: 576
- Date: Wed, 03 Feb 2010 04:32:04 GMT

Blank line :

Message body : <?xml> ... </html>

17.2. HTTP request message: Request-URI

```
/var/lib/tomcat6/webapps/ROOT/
|-- META-INF
|   `-- context.xml
```

```
|-- index.html
`-- _index.html
```

- Locate the resource in the server host's file-system
- Resources organized in a subdirectory in the server host's file-system
- Root of the web server
- URL = Scheme :// Hostname Request-URI
- URL = Scheme :// Hostname:Port Request-URI
- URL = Scheme :// Authority Request-URI

17.3. HTTP request message: Request method

Method	Requests server to
GET	return the resource specified by the Request-URI as the response message body
POST	pass the request message body on as data to be processed by the resource specified by the Request-URI
HEAD	return the same HTTP header fields that would be returned if a GET method were used, but not return the message body

- ▷ `GET` commonly used method
- ▷ `POST` for processing data input through form in the browser
- ▷ `HEAD` for cache control

18. HTTP response message

Status line

HTTP/1.1 200 OK

- Status code
- Reason phrase

Status class	Use
1	Informational
2	Success
3	Redirection
4	Client error
5	Server error

18.1. HTTP response message: Status code

Status code	Reason phrase	Use
200	OK	Request processed normally
301	Moved permanently	URI for the requested resource has changed; Location header field of the response message contains the new URI
401	Unauthorised	Resource is password protected, user has not supplied a valid password
403	Forbidden	Resource is read protected
404	Not found	No resource corresponding to the Request-URI

18.2. Request header fields

Field name	Use
Host	Authority portion of the URL (host name + port number)
User-Agent	Identify the browser
Accept	MIME types for the response body the browser can accept, possibly with preference
Accept-Language	Preferred languages for the response body
Accept-Encoding	Preferred encoding for the response body
Accept-Charset	Preferred character sets
Connection	Should the TCP connection be kept open after the response is sent?`keep-alive` (default), `close`
Keep-Alive	Number of seconds the TCP connection to be kept open?

18.3. Request header fields

Field name	Use
Content-Type	MIME type of the request message body
Content-Length	Number of bytes in the request message body

18.4. Response header fields

Field name	Use
Date	Time at which the response was generated by the server
Server	Identify the server
Last-Modified	Time at which the requested resource was last modified in the server
Expires	The requested resource may be modified in the server after this period
ETag	Hash code of the resource
Accept-Ranges	Client can request only a portion (range) of the resource using its Range header field. Server specified the units that may be used by the client in its Range field, or `none`.

18.5. Response header fields

- Connection
- Keep-Alive
- Content-Type
- Content-Length