



Web Engineering

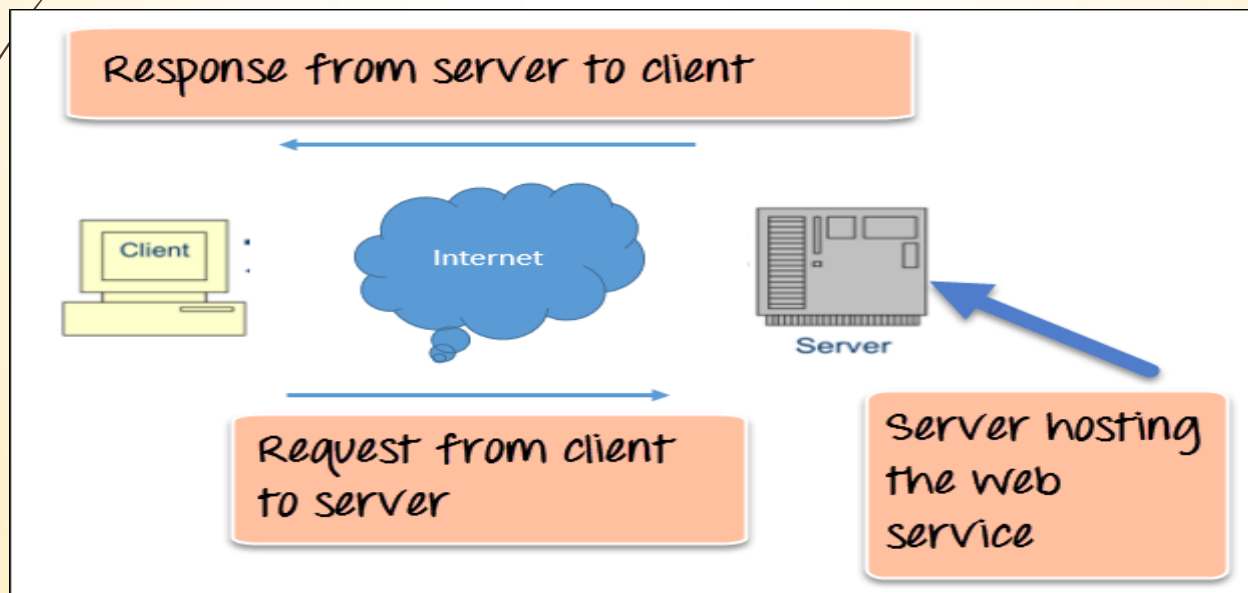
Web Service

❑ **A web service is, any service that:**

- Is a standardized medium to propagate communication between the client and server applications on the World Wide Web.
- Uses a standardized XML messaging system
- Is not tied to any one operating system or programming language

Web Service

- **Client:** web browsers, used to surf the Web
- **Server** systems: used to supply information to these browsers
- Computer **networks:** used to support the browser-server communication



HTTP

- HTTP is CONNECTIONLESS: after making the request the client disconnected from the server, then when the server is ready the server re-establish the connection again and deliver the response
- The http can deliver any sort of data
- HTTP is STATELESS: the client and server know each other just during the current request

http request

- A request line to get a required resource, for example a request GET /content/page1.html is requesting a resource called /content/page1.html from the server.
- Headers (Example – Accept-Language: EN).
- An empty line.
- A message body which is optional.

http response

- ▶ HTTP Status Code (For example HTTP/1.1 301 Moved Permanently, means the requested resource was permanently moved and redirecting to some other resource).
- ▶ Headers (Example – Content-Type: html)
- ▶ An empty line.
- ▶ A message body which is optional.

HTTP Status Codes

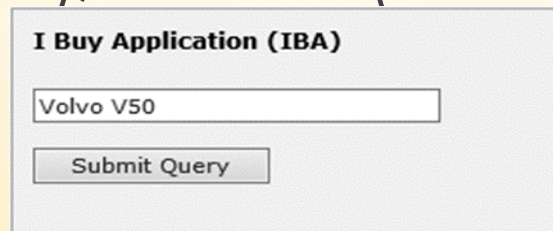
- 1xx series - Informational Message:- request is received from the client successfully and is under processing at the server end.
- 2xx - Success Message:- request is received and processed successfully
- 3xx - Redirection Message
- 4xx - Error Messages Related to Client:-
eg. 404 - Page Not Found
- 5xx - Error Messages Related to Server:-
eg. 504 - Gateway Timeout

Semantic web

- The Semantic Web is a vision about an extension of the existing World Wide Web, which provides software programs with machine-interpretable metadata of the published information and data.
- In other words, we add further data descriptors to otherwise existing content and data on the Web. As a result, computers are able to make meaningful interpretations similar to the way humans process information to achieve their goals.

Semantic Web Example

- Buying And Selling Used Cars
- Suppose a semantic web system was built to administer the selling and buying of used cars over the Internet.
- The system would contain two main applications:
- One for people who wanted to buy a car
- One for people who wanted to put up a car for sale
- Let's call the Internet applications for IBA (I Buy Application), and ISA (I Sell Application).



I Buy Application (IBA)

Volvo V50

Submit Query

RSS and Atom feed

- ▶ are small text files that provide information about content on websites. When content is updated, the feed text file is also updated, either manually or programatically.
- ▶ Applications called “readers” or “aggregators” can then check these small text files and notify someone when new content is available.

RSS feed

- RSS stands for both *Rich Site Summary* and *Really Simple Syndication*
- It is a mean of transmitting and updating news in an automated way.
- Most news sites (including all blogs) will publish what is called an RSS feed which is regularly updated with the latest available headlines and/or articles.
- An RSS aggregator makes it very convenient to follow up on news from a large number of sources in a single place.
- SharpReader is an example of an RSS Reader.

Atom feed

- An Atom feed is very similar to an RSS feed in that it is a lightweight XML format allowing for easy syndication of web content.
- The primary difference between the two formats is the XML tags that are used.

CAPTCHA

- ▶ A CAPTCHA is a program that protects websites against bots by generating and grading tests that humans can pass but current computer programs cannot. For example, humans can read distorted text as the one shown below, but current computer programs can't:
- ▶ Captcha is a verification process that requires users to enter a pre-determined code. CAPTCHA stands for Completely Automated Public Turing test to tell Computers and Humans Apart.



Types of captcha

- ▶ Mathematical calculations
- ▶ Images
- ▶ Distorted texts

Applications of CAPTCHAs

- **Preventing Comment Spam in Blogs.** Most bloggers are familiar with programs that submit bogus comments, usually for the purpose of raising search engine ranks of some website / Search engine bots-to make indexed
- **Protecting Website Registration.** Several companies (Yahoo!, Microsoft, etc.) offer free email services. Up until a few years ago, most of these services suffered from a specific type of attack: "bots" that would sign up for thousands of email accounts every minute.

Applications of CAPTCHAs

- **Protecting Email Addresses From Scrapers.** Spammers crawl the Web in search of email addresses posted in clear text. A free and secure implementation that uses CAPTCHAs to obfuscate an email address can be found at reCAPTCHA MailHide.
- **Online Polls.** asking which was the best graduate school in computer science. As is the case with most online polls, IP addresses of voters were recorded in order to prevent single users from voting more than once.

Applications of CAPTCHAs

- **Preventing Dictionary Attacks.** prevent a computer from being able to iterate through the entire space of passwords by requiring it to solve a CAPTCHA after a certain number of unsuccessful logins.
- **Worms and Spam.** CAPTCHAs also offer a reasonable solution against email worms and spam: "I will only accept an email if I know there is a human behind the other computer." A few companies are already marketing this idea.