Introduction to web

HCL

What is the Internet?

- Network is a group of connected computers/devices
- Internet is a network of networks
- Internet is the world wide web

HCL

Internet address

- Used to identify a computer connected to the Internet.
- Every address must be unique, since the computer represented by that address is unique to the Internet.
- The Internet Corporation for Assigned Names and Numbers (ICANN) is an internationally organized, non-profit corporation that has responsibility for Internet Protocol (IP) address space allocation, protocol identifier assignment.
- Internet addresses can be represented in terms of textual domain names, such as www.yahoo.com and mapped to corresponding IP addresses. The IP for www.yahoo .com is 205.132.48.237.

Protocol

- A *protocol* is a set of rules for communicating across the Internet. Both parties know and follow the rules for sending and receiving information, making meaningful communication possible.

  Ex:

  Hypertext Transfer Protocol (HTTP)

  File Transfer Protocol (FTP)

  Telnet

HTTP

- This protocol, the backbone of the World Wide Web, enables users to send and receive information from Internet servers in the form of documents, or *pages,* written using the Hypertext Markup Language (HTML).
- The user who receives the document, often called the *client*, can then use a *browser* or other form of software that recognizes the HTML language to view the contents of the document.

HTTPS

- This is a secure protocol which is combination of HTTP + SSL/TLS protocol.
- SSL – Secure Socket Layer, TLS – transport Layer Security. SSL is TLS's predecessor.
- The browser adds encryption layer of SSL/TLS to protect the traffic.

WWW (World Wide Web)

- An information system that brings together data from many of the other Internet services under one set of protocols.
- W3 Consortium was created for continuing to develop the standards. The consortium put together a set of protocols for the World Wide Web.

What it W3C?

- W3C Stands for the World Wide Web Consortium
- W3C was created in October 1994
- W3C was established by Tim Berners-Lee
- W3C has many members who work towards standardization of the Web
- W3C Standards are called W3C Recommendations

W3C Recommendations

- Each W3C Recommendation is developed by a working group consisting of members and invited experts.
- The group obtains its input from companies and other organizations, and creates a Working Draft and finally a Proposed Recommendation.
- The Proposed Recommendation is usually submitted to the W3C membership and Director
- On formal approval, it becomes a W3C Recommendation.

Web Paradigms

- Broadly classified into two;
  - Classic Web Paradigm-Synchronous
  - Asynchronous

Synchronous Web Paradigm

- Client browser requests data from the server
- Each time the client wants an update, it makes a request
- Also known as “pull”

Limitations of Synchronous Web Paradigm

- There are a growing class of applications that need visualization of real-time data
  - stock prices from trading sites

A solution…

- A polling technique
- Problem only partially resolved
  - The update frequency cannot be high. A synchronous paradigm (request/response) makes it impossible to receive data in real time.
  - The occupied network bandwidth is high, because with each response a whole page is transferred, instead of only the changed data.
  - The impact on Web server resources is huge, because the server needs to sustain a high load of page requests even though users are inactive.

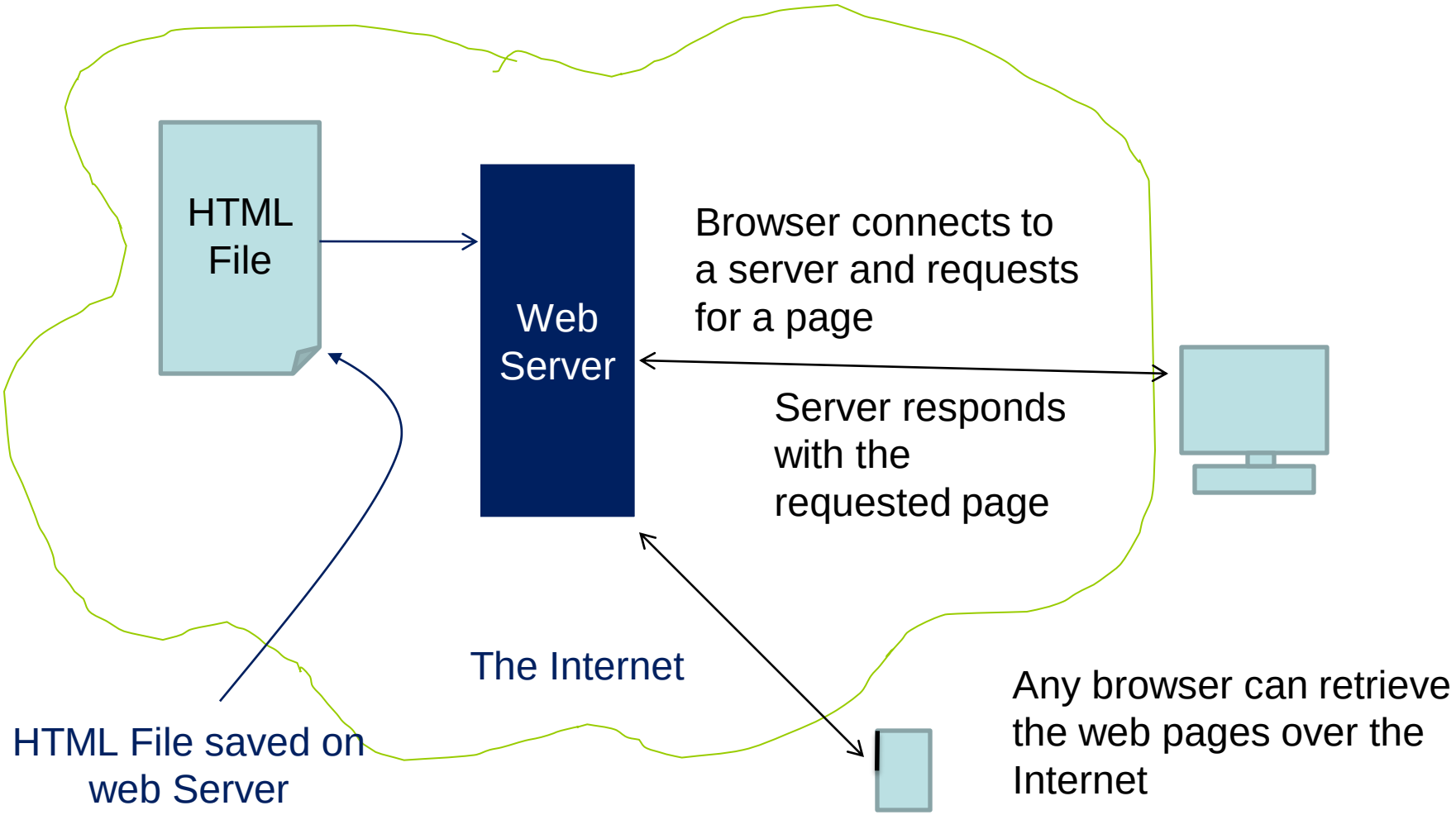
HCL

Asynchronous Web Paradigm

- To guarantee a very low latency between the generation of fresh data and its presentation to the end user within a common browser, a dedicated solution is necessary, namely Push Technology. This term was coined in 1996.
- In the push (or streaming) model, the client receives updates in an asynchronous manner at the server's discretion, in the form of a continuous data flow.

HCL

Basic Process



Behind the Scene!!!

- *Uniform resource locators*, or URLs, incorporate Internet addresses to indicate the network location of a Web page or other network resource
- The browser breaks the URL into three parts:
  - The protocol ("http")
  - The server name ("www.rediff.com")
  - The file name (“index.html")
- The browser communicates with a name server to translate the server name "www.rediff.com" into an IP Address, which it uses to connect to the server machine.
- The browser then establishes a connection with the server at that IP address on port 80.

- Following the HTTP protocol, the browser sends a request to the server, asking for the file http://www.rediff.com/index.html
- The server then sends the HTML text for the Web page to the browser.
- The browser reads the HTML tags, formats the page and displays it onto your screen.