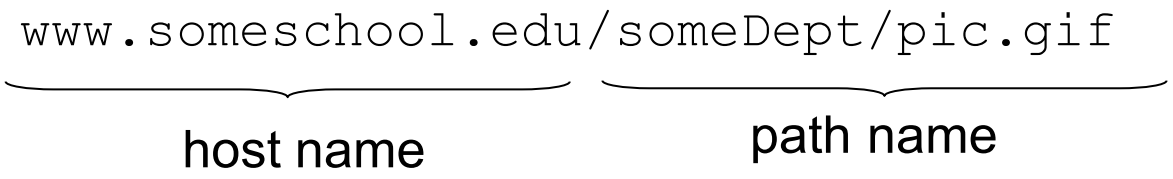
Web and HTTP

Roadmap

- Web and HTTP
- HTTP Overview
- Uniform Resource Locators (URL)
- HTTP Connections
- HTTP Messages
- HTTP Request Format
- HTTP Methods
- HTTP Response Format
- HTTP Status Codes

Web and HTTP

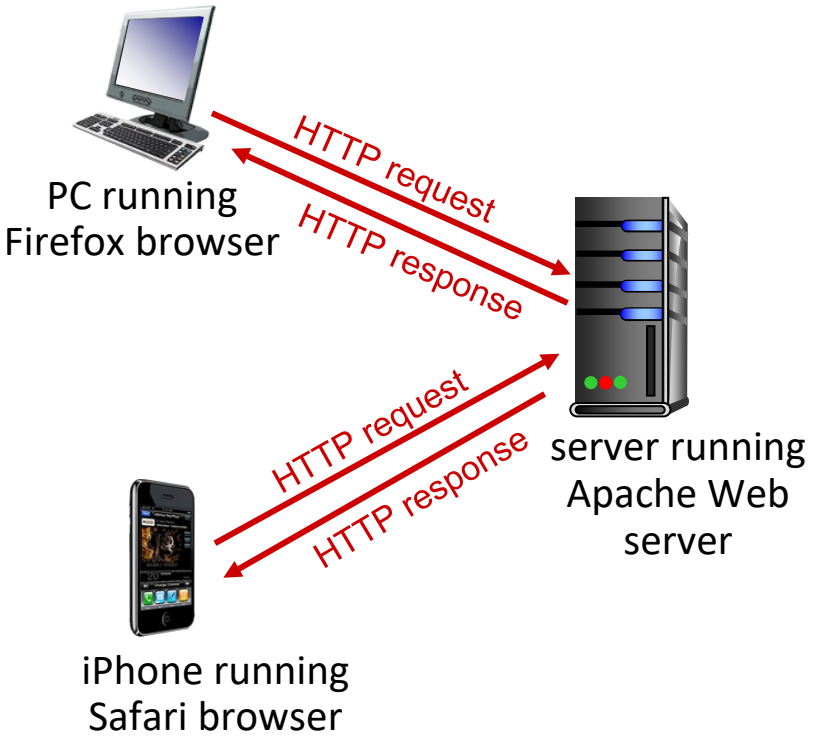
- web page consists of *objects/resources,* each of which can be stored on different Web servers
  - object can be HTML file, JPEG image, Java applet, audio file,…
  - Web page contains some text and 5 images … 6 objects
- web page consists of *base HTML-file* which includes *several referenced objects, each* addressable by a *URL,* e.g.,



HTTP overview

HTTP: hypertext transfer protocol

- Web's application-layer protocol
- client/server model:
  - *client:* browser that requests, receives, (using HTTP protocol) and "displays" Web objects
  - *server:* Web server sends (using HTTP protocol) objects in response to requests



HTTP overview (continued)

HTTP uses TCP:

- client initiates TCP connection (creates socket) to server, port 80
- server accepts TCP connection from client
- HTTP messages (application-layer protocol messages) exchanged between browser (HTTP client) and Web server (HTTP server)
- TCP connection closed

HTTP is "stateless"

- server maintains *no* information about past client requests

aside

protocols that maintain "state" are complex!

- past history (state) must be maintained

HTTP connections: two types

Non-persistent HTTP

1. TCP connection opened
2. at most one object sent over TCP connection
3. TCP connection closed

downloading multiple objects required multiple connections

Persistent HTTP

- TCP connection opened to a server
- multiple objects can be sent over *single* TCP connection between client, and that server
- TCP connection closed
- HTTP 1.1

HTTP: Uniform Resource Identifier (URI)

- The server resource name is called a Uniform Resource Identifier, or URI
- URIs are like the postal addresses of the Internet, uniquely identifying and locating information resources around the world
- URIs come in flavour, called URLs
- The uniform resource locator (URL) is the most common form of resource identifier
  - a URL tells precisely where a resource is located and how to access it

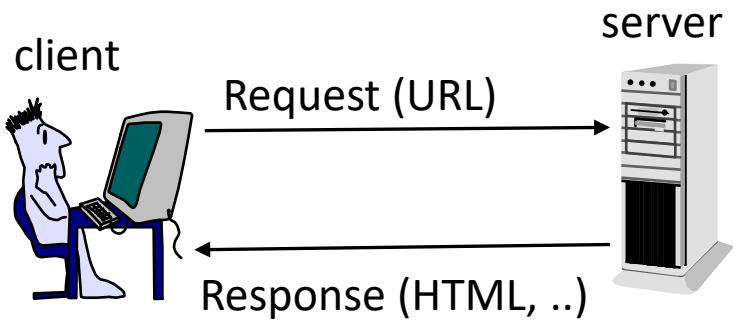
https://image.freepik.com/free-vector/beautiful-nature-wood-scene_1308-24813.jpg

HTTP: Uniform Resource Locators (URL)

- Used to identify resources on the Internet
- URL has the following general format
  - Scheme:object-address
  - Schemes are protocols and they can be
    - http, https, ftp, telnet, etc.
  - Object-address
    - //fully-qualified-domain-name/document-path
    - E.g. https://www.bradford.ac.uk/ei/computer-science/
  - Port- A port number to be used to access the resource
    - If omitted, the default value for the specified protocol will be used

HTTP: Client and Server communication

- User uses HTTP client (web browser)
- Enters a URL
- Makes a request to the server
- Client initiates TCP connection (creates socket) to server, **port 80**
- Server accepts TCP connection from client
- HTTP messages (application-layer protocol messages) exchanged between browser (HTTP client) and Web server (HTTP server)
- TCP connection closed



HTTPS

- HTTPS stands for Hypertext Transfer Protocol **Secure** (also referred to as HTTP over TLS or HTTP over SSL)
- HTTPS uses **port 443** for sending and receiving data

Acknowledgements and References

- James F. Kurose and Keith W. Ross. Computer Networking: A Top-Down Approach, 8th edition, Pearson, 2021.

- Some parts of the content are adapted from:
  - Professor F. Ricci's lecture on HTTP