## Chapter 2: outline

- 2.1 principles of network applications
- 2.2 Web and HTTP
- 2.3 electronic mail
  - SMTP, POP3, IMAP
- **2.4 DNS**

- 2.5 P2P applications
- 2.6 video streaming and content distribution networks
- 2.7 socket programming with UDP and TCP

### Web and HTTP

#### Related Standards and RFCs

- Related RFCs
  - RFC 1945: Hypertext Transfer Protocol, HTTP/1.0
  - RFC 2616: Hypertext Transfer Protocol, HTTP/1.1
  - RFC 3986: Uniform Resource Identifier (URI): Generic Syntax
  - RFC 1738: Uniform Resource Locators (URL)
- Web Pages
  - World Wide Web Consortium (W3C): <a href="http://www.w3.org/">http://www.w3.org/</a>
  - The Internet Archive: <u>www.archive.org</u>

Obsoleted by: <u>7230</u>, <u>7231</u>, <u>7232</u>, <u>7233</u>, <u>7234</u>, <u>7235</u>

Updated by: <u>2817</u>, <u>5785</u>, <u>6266</u>, <u>6585</u>

Network Working Group Request for Comments: 2616

Obsoletes: 2068

Category: Standards Track

R. Fielding
UC Irvine
J. Gettys
Compaq/W3C
J. Mogul
Compaq
H. Frystyk
W3C/MIT
L. Masinter
Xerox
P. Leach
Microsoft
T. Berners-Lee
W3C/MIT
June 1999

DRAFT STANDARD

Errata Exist

#### RFC 2616

https://tools.ietf.org/html/rfc2616

#### Hypertext Transfer Protocol -- HTTP/1.1

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (1999). All Rights Reserved.

Abstract

The Hypertext Transfer Protocol (HTTP) is an application-level protocol for distributed, collaborative, hypermedia information systems. It is a generic, stateless, protocol which can be used for

## Web and HTTP

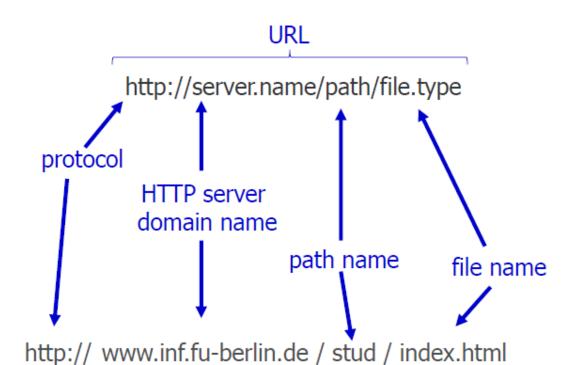
### First, a review...

- web page consists of objects
- object can be HTML file, JPEG image, Java applet, audio file,...
- web page consists of base HTML-file which includes several referenced objects
- each object is addressable by a URL, e.g.,

www.someschool.edu/someDept/pic.gif

host name

path name



URI, URI

URI, URN

SCHEME HOST PATH

https://www.wikipedia.org/index.html

## 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)

#### 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
- if server/client crashes, their views of "state" may be inconsistent, must be reconciled

## HTTP connections

### non-persistent HTTP

- at most one object sent over TCP connection
  - connection then closed
- downloading multiple objects required multiple connections

#### persistent HTTP

 multiple objects can be sent over single TCP connection between client, server

## Non-persistent HTTP

#### suppose user enters URL:

www.someSchool.edu/someDepartment/home.index

(contains text, references to 10 jpeg images)

- Ia. HTTP client initiates TCP connection to HTTP server (process) at www.someSchool.edu on port 80
- 2. HTTP client sends HTTP request message (containing URL) into TCP connection socket.

  Message indicates that client wants object someDepartment/home.index
- Ib. HTTP server at host
   www.someSchool.edu waiting
   for TCP connection at port 80.
   "accepts" connection, notifying client
- 3. HTTP server receives request message, forms response message containing requested object, and sends message into its socket

## Non-persistent HTTP (cont.)



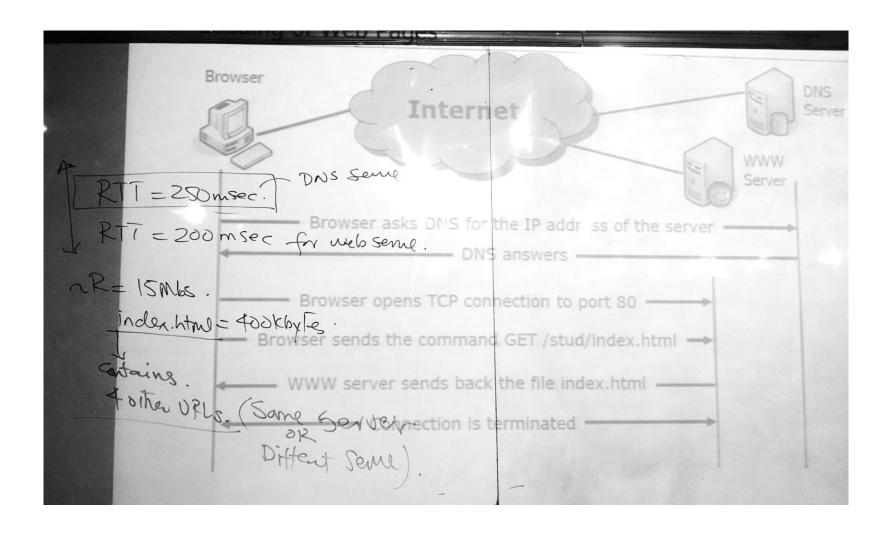
5. HTTP client receives response message containing html file, displays html. Parsing html file, finds 10 referenced jpeg objects

**4.** HTTP server closes TCP connection.



6. Steps 1-5 repeated for each of 10 jpeg objects

### Loading of Web Pages DNS Browser Server **Internet** WWW Server Browser asks DNS for the IP address of the server DNS answers Browser opens TCP connection to port 80 Browser sends the command GET /stud/index.html → WWW server sends back the file index.html . Connection is terminated



- 1. Calculate total time to transfer index.html given the values
- 2. Calculate total time to transfer index.html when it contains four other URLs each pointing to resources from different web servers.