

CS 181AG
Lecture 22

DNS & HTTP

Arthi Padmanabhan

Nov 30, 2022

Upcoming Schedule

- Assignment 10 released **tomorrow** by 12pm, due Friday Dec 9 by 10pm (note that this is the same time that Assignment 8 is due)
 - Please see me *early* if you don't have the client-server communication working
- Final: released Saturday Dec 10th at 8am, due Friday Dec 16th at 10pm

Our Socket Program

- We sent directly to an IP address and port
- Humans aren't good at remembering IP addresses

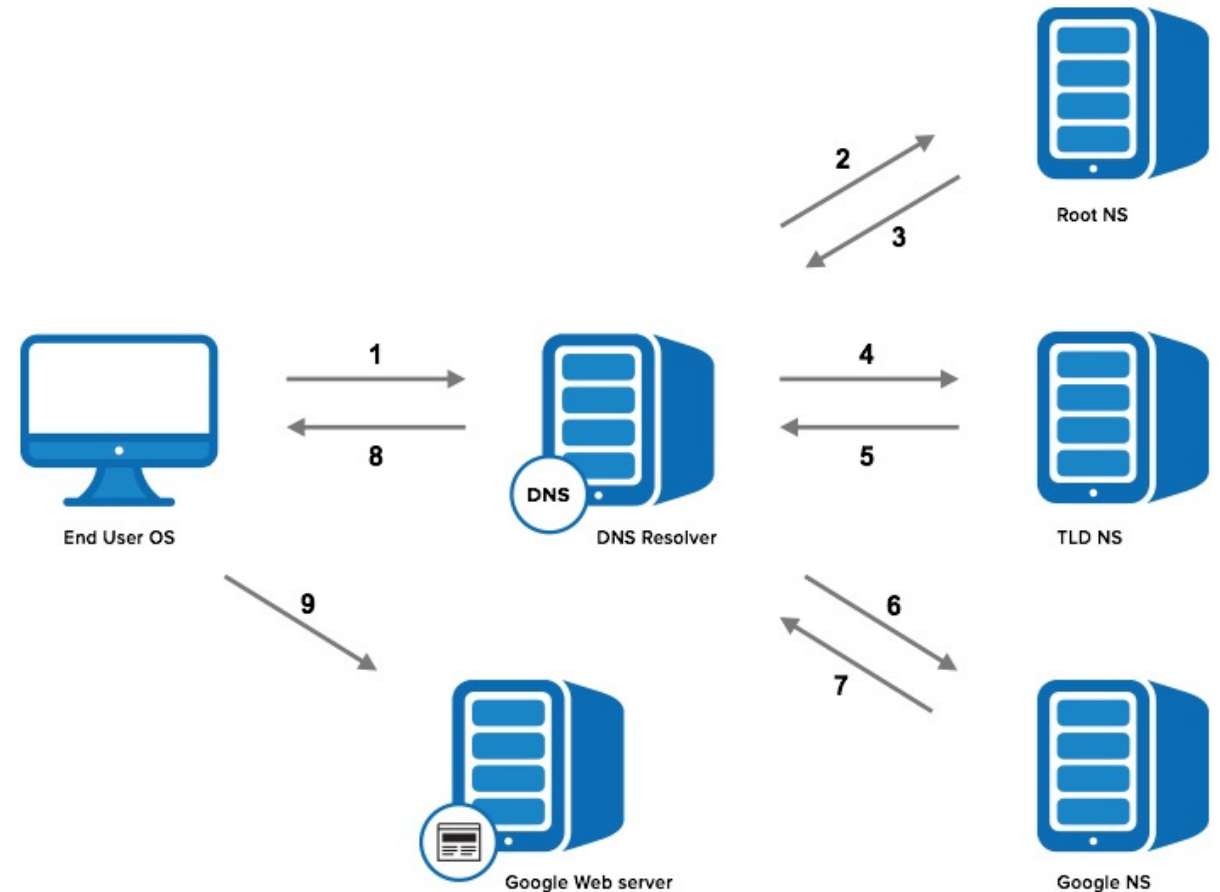
```
HOST = "172.28.90.8"  
PORT = 65432  
  
s = socket.socket(socket.AF_INET, socket.SOCK_STREAM)  
  
s.connect((HOST, PORT))  
s.sendall("Hello world".encode())
```

Domain Name System

- Helps us get the IP address, given a domain name (www.hmc.edu)

DNS Lookup

- Device first goes to resolver (a DNS server often set up by your ISP)
- UDP is used for device – resolver, resolver – root, resolver – TLD, and resolver – authoritative
- Once device has the IP, it uses whatever protocol it wants (usually TCP)



HTTP

- One issue with our socket program was remembering IPs
- Another is that we typically don't just send strings back and forth - we send "resources"
- Resources can be files, audio, video, etc (they can be dynamic)

URL

- Resources are identified by url (universal resource location)
- <http://www.example.com/file/dir.html>

URL

- Resources are identified by url (universal resource location)
- <http://www.example.com/file/dir.html>
- Scheme: what protocol is this? (http, ftp, dns, etc)

URL

- Resources are identified by url (universal resource location)
- <http://www.example.com/file/dir.html>
- Domain name: what users type into browser to reach website

URL

- Resources are identified by url (universal resource location)
- <http://www.example.com/file/dir.html>
- Sub-domain: domain is reachable via world wide web

URL

- Resources are identified by url (universal resource location)
- <http://www.example.com/file/dir.html>
- Path to resource within example.com

URL

- Resources are identified by url (universal resource location)
- <http://www.example.com/file/dir.html>
- Optimal parameters, usually of form key:value

HTTP

- Client sends a request for a resource to the server
- Server responds and then closes connection
- HTTP is stateless
- HTTP exchanges run over TCP

Some Types of HTTP Messages

- GET – request data
- POST – here is data to store (for uploading, create another copy if already exists)
- PUT – here is data to store (for uploading, replace if already exists)
- HEAD – give me information about the resource but don't actually send it

Form of HTTP Request

- Initial line
- Header lines
- Message body, if applicable
- Initial line: GET <path to file> <HTTP version>
- GET path_to_file/index.html HTTP/1.0
- Ex/ use Wireshark to go to gaia.cs.umass.edu/wireshark-labs/HTTP-wireshark-file1.html

Form of HTTP Response

- <HTTP version> <Status code> <More information>
- HTTP/1.0 200 OK
- HTTP/1.0 404 Not Found
- Code is 3 digit number, where 1st is general category
- 1: informational message
- 2: success
- 3: redirecting
- 4: error on client's part
- 5: error on server's part

Resource that Include Resources?

- Html file might need stylesheet, so it would specify the path to the stylesheet
- Browser sees this and automatically requests the stylesheet