CS631 - Advanced Programming in the UNIX Environment

HTTP; Group Project Discussions

Department of Computer Science
Stevens Institute of Technology
Jan Schaumann
jschauma@stevens.edu
https://stevens.netmeister.org/631/
HTTP

Hypertext Transfer Protocol

RFC 1945 (HTTP/1.0)

And then RFC 2616 (HTTP/1.1); and then RFC 7540 (HTTP/2); and then HTTP/3 (Draft); ...

A simple request/response protocol.
The Hypertext Transfer Protocol

HTTP is a request/response protocol:

1. client sends a request to the server
2. server responds
The Hypertext Transfer Protocol

HTTP is a request/response protocol:

1. client sends a request to the server
   - request method
   - URI
   - protocol version
   - request modifiers
   - client information

2. server responds
HTTP: A client request

$ telnet www.google.com 80
Trying 2607:f8b0:4006:800::2004...
Connected to www.google.com.
Escape character is '\]'.
GET / HTTP/1.0
The Hypertext Transfer Protocol

HTTP is a request/response protocol:

1. client sends a request to the server
   - request method
   - URI
   - protocol version
   - request modifiers
   - client information

2. server responds
   - status line (including success or error code)
   - server information
   - entity metainformation
   - content
HTTP: a server response

HTTP/1.0 200 OK
Date: Fri, 01 Nov 2019 02:34:00 GMT
Content-Type: text/html; charset=ISO-8859-1
Server: gws
X-XSS-Protection: 1; mode=block
X-Frame-Options: SAMEORIGIN
Set-Cookie: 1P_JAR=2019-11-01-02; expires=Sun, 01-Dec-2019 02:34:00 GMT; path=/; domain=.google.com;
Set-Cookie: NID=190=s7MFbpfUQ1u26sSNpa3NHj42Rg98cQDw1iggAt0MvpM0YTH8NfmWQt0FMruYwAV4I

<!doctype html><html itemscope="itemscope" itemtype="http://schema.org/WebPage">
<head><meta content="Search the..."
The Hypertext Transfer Protocol

Server status codes:

- **1xx** – Informational; Request received, continuing process
- **2xx** – Success; The action was successfully received, understood, and accepted
- **3xx** – Redirection; Further action must be taken in order to complete the request
- **4xx** – Client Error; The request contains bad syntax or cannot be fulfilled
- **5xx** – Server Error; The server failed to fulfill an apparently valid request
HTTP: A client request

$ telnet www.cs.stevens-tech.edu 80
Trying 155.246.51.11...
Connected to www.cs.stevens-tech.edu.
Escape character is '^[].
GET / HTTP/1.0

HTTP/1.1 301 Moved Permanently
Date: Fri, 01 Nov 2019 02:42:36 GMT
Server: Apache
Location: https://www.cs.stevens.edu/
Content-Length: 235
Connection: close
Content-Type: text/html; charset=iso-8859-1

<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
<html><head>
HTTP: A client request


GET / HTTP/1.1
Host: www.cs.stevens.edu

HTTP/1.1 302 Found
Date: Fri, 01 Nov 2019 02:43:41 GMT
Server: Apache
Location: https://www.stevens.edu/ses/cs
Vary: Accept-Encoding
Content-Length: 214
Content-Type: text/html; charset=iso-8859-1

<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML 2.0//EN">
HTTP: A client request

$ openssl s_client -crlf -servername www.stevens.edu -connect www.stevens.edu:443

[...] GET /ses/cs HTTP/1.1
Host: www.stevens.edu

HTTP/1.1 301 Moved Permanently
Date: Fri, 01 Nov 2019 02:45:11 GMT
Content-Type: text/html; charset=UTF-8
Transfer-Encoding: chunked
Set-Cookie: __cfduid=dfdf7109985d2521a4717fa0e0d62d8a21572576311; expires=Sat, 31-Oct-20 02:45:11 GMT
Location: https://www.stevens.edu/schaefer-school-engineering-science/departments/computer-science
Via: varnish
Server: cloudflare
HTTP: A client request

[...]
GET /schaefer-school-engineering-science/departments/computer-science HTTP/1.1
Host: www.stevens.edu

HTTP/1.1 200 OK
Date: Fri, 01 Nov 2019 02:46:52 GMT
Content-Type: text/html; charset=utf-8
Set-Cookie: __cfduid=dd5d6da1261bad72024f790c023a1a6561572576412; expires=Sat, 31-Oct-20 02:46:52 GMT;
Last-Modified: Fri, 01 Nov 2019 00:22:20 GMT
Via: varnish
X-Generator: Drupal 7 (http://drupal.org)
Server: cloudflare

<!DOCTYPE html>
<html lang="en" class="no-js">

HTTP, Coding Practices

November 4, 2019
HTTP: A client request

[Browser window showing network requests and timings for Stevens Institute of Technology's computer science page]
HTTP - more than just text

HTTP is a *Transfer Protocol* – serving *data*, not any specific text format.

- **Accept-Encoding** client header can specify different formats such as *gzip*, *Shared Dictionary Compression over HTTP (SDCH)* etc.
- **corresponding server headers**: *Content-Type* and *Content-Encoding*
HTTP - more than just static data

HTTP is a Transfer Protocol – what is transferred need not be static; resources may generate different data to return based on many variables.

- CGI – resource is executed, needs to generate appropriate response headers
- server-side scripting (ASP, PHP, Perl, ...)
- client-side scripting (JavaScript/ECMAScript/JScript,...)
- applications based on HTTP, using:
  - AJAX
  - RESTful services
  - JSON, XML, YAML to represent state and abstract information
HTTP in your final project

- protocol versions supported: 1.0
- request methods supported: GET, HEAD
- request headers supported: If-Modified-Since
- response headers included: Date, Server, Last-Modified, Content-Type, Content-Length
HTTP in your final project

Let’s create your building blocks:

- program startup and daemonization
- rough network logic
- input / request parsing and validation
- cgi handling
- user directory handling
- directory index generation
- regular file serving
Reading

HTTP etc.:

- RFC 1945, 2616, 2818, 3875, 7540
- https://httpd.apache.org/docs/current/
- https://www.w3.org/Protocols/
- REST: https://is.gd/leSvGa

Homework: Vote. (Sorry, US Citizens only.)
https://www.vote411.org/