

BUCKNELL UNIVERSITY
Computer Science

CSCI 363 Computer Networks

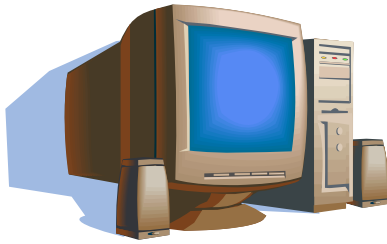
HTTP and the World Wide Web

Source: Computer Networks, Andrew Tanenbaum, 4th Edition.

Uniform Resource Locator (URL)

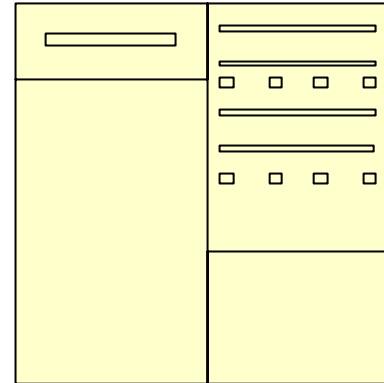


Architecture



Client: **web-browser**

Firefox
Mozilla
Opera
Safari
IE
...



Server: **web-server**
Well-known service: **port 80**

Talking to a Web Server

```
myhost% telnet www.eg.bucknell.edu 80
Trying 134.82.131.16...
Connected to hera.eg.bucknell.edu.
Escape character is '^]'.
GET http://www.eg.bucknell.edu/~perrone/ HTTP/1.0

HTTP/1.1 200 OK
Date: Fri, 29 Apr 2005 10:43:10 GMT
Server: Apache/1.3.29 (Unix) mod_ssl/2.8.16 OpenSSL/0.9.7d PHP/4.3.10
Last-Modified: Fri, 03 Sep 2004 18:35:07 GMT
ETag: "986397-76e-4138b95b"
Accept-Ranges: bytes
Content-Length: 1902
Connection: close
Content-Type: text/html

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">
<html>

<head>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1">
...

</html>
Connection to hera.eg.bucknell.edu closed by foreign host.
```

Server Side

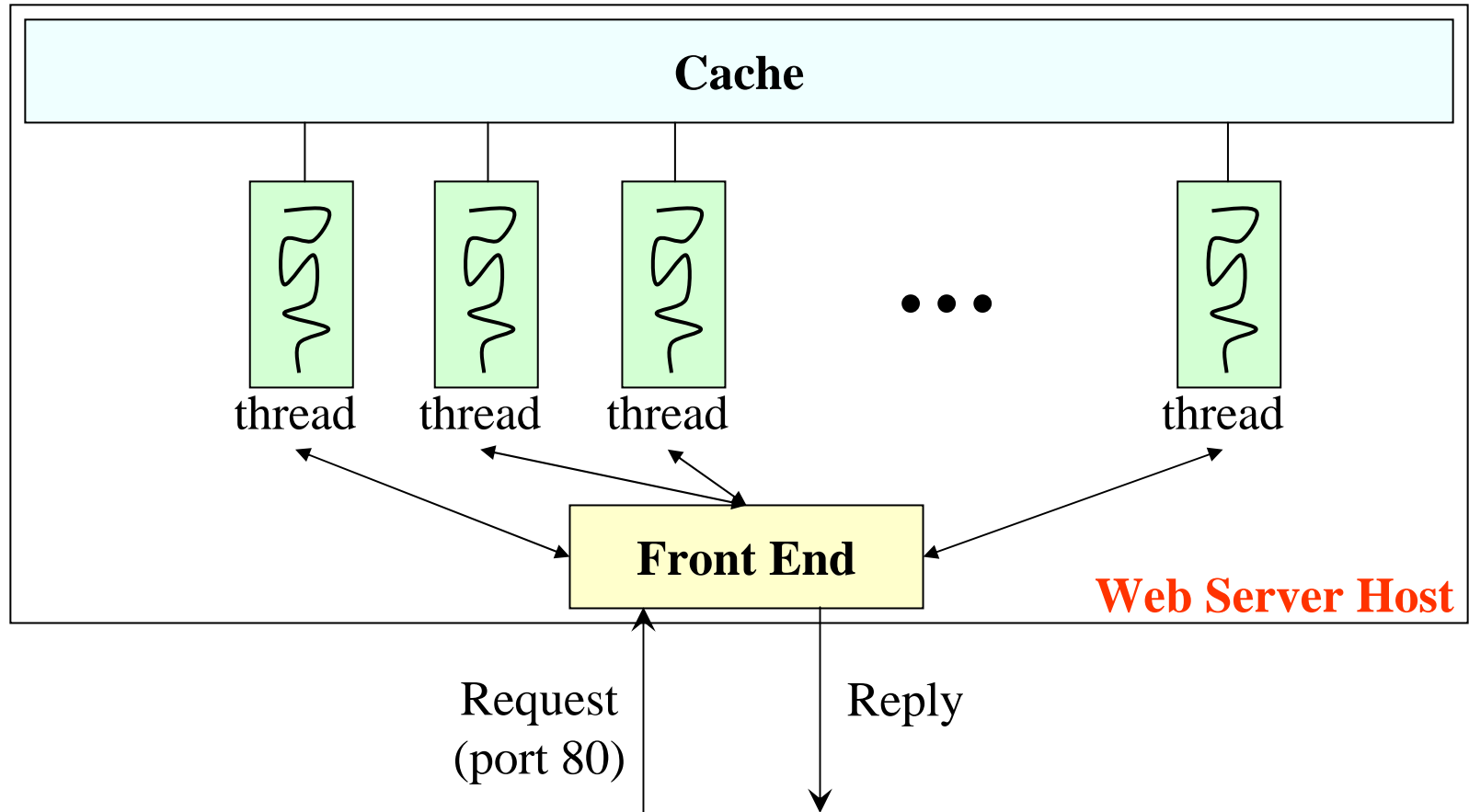
1. Accept TCP connection from browser.
2. Get the name of the file requested.
3. Read the file from disk.
4. Send the file back to client.
5. Release the TCP connection.

HTTP 1.0

HTTP 1.1: Client and server can exchange multiple messages through the same TCP connection.

Question: Name two effects of HTTP 1.1 on performance.
Hint: Think about the effect on TCP's congestion control mechanisms.

Server-side Performance Enhancements



Server Side Threads

1. Resolve name of the web page requested.
2. Authenticate client.
3. Perform access control on client.
4. Perform access control on web page.
5. Check cache.
6. Fetch requested page from disk.
7. Determine MIME type to include in response.
8. Return reply to client.
9. Enter record of transaction in log.

MIME Types `/etc/mime.types`

```
# This file controls what Internet media types are sent to the client for
# given file extension(s).  Sending the correct media type to the client
# is important so they know how to handle the content of the file.
# Extra types can either be added here or by using an AddType directive
# in your config files.  For more information about Internet media types,
# please read RFC 2045, 2046, 2047, 2048, and 2077.  The Internet media type
# registry is at <http://www.iana.org/assignments/media-types/>.
# MIME type                               Extension
application/applefile
application/http
application/mac-binhex40                 hqx
application/msword                        doc
application/octet-stream                 bin dms lha lzh exe class so dll
application/pdf                          pdf
audio/basic                             au snd
audio/midi                               mid midi kar
audio/mpeg                               mpga mp2 mp3
audio/x-aiff                             aif aiff aifc
audio/x-pn-realaudio                    ram rm
audio/x-realaudio                        ra
audio/x-wav                             wav
image/bmp                               bmp
image/gif                               gif
image/jpeg                               jpeg jpg jpe
image/png                               png
video/mpeg                               mpeg mpg mpe
video/quicktime                          qt mov
```

HTTP Request Methods

Method	Description
GET	Request to read a web page
HEAD	Request to read a page's header
PUT	Request to store a web page
POST	Append to a web page
DELETE	Remove a web page
TRACE	Echo the incoming request
CONNECT	Reserved for future use
OPTIONS	Query options

HTTP Message Headers

Header	Type	Contents
User-Agent	Response	Information about browser and platform
Accept	Response	Type of pages client can handle
Host	Response	Server's DNS name
Content-Encoding	Response	How content is encoded
Content-Length	Response	Page's length in bytes
Content-Type	Response	Page's MIME Types
Cookie	Response	Send previously set cookie back to server
Last-Modified	Response	Time and date of page's last change

HTTP Request

```
>> IP PACKET *****
```

```
Next protocol = TCP, Transmission Control Protocol.
```

```
Source          = 134.82.18.89
```

```
Destination = 134.82.7.4
```

```
>> TCP PACKET *****
```

```
source          port = 1650, service = ?
```

```
destination port = 80, service = HTTP, HyperText Transfer Protocol.
```

```
PAYLOAD HEXDUMP:
```

```
0000000: 4745 5420 2f20 4854 5450 2f31 2e30 0d0a  GET / HTTP/1.0..
0000010: 5265 6665 7265 723a 2066 696c 653a 2f2f  Referer: file://
0000020: 2f44 756f 2532 3048 442f 5379 7374 656d  /Duo%20HD/System
0000030: 2532 3046 6f6c 6465 722f 5072 6566 6572  %20Folder/Prefer
0000040: 656e 6365 732f 4e65 7473 6361 7065 2532  ences/Netscape%2
0000050: 30c4 2f42 6f6f 6b6d 6172 6b73 2e68 746d  0./Bookmarks.htm
0000060: 6c0d 0a43 6f6e 6e65 6374 696f 6e3a 204b  l..Connection: K
0000070: 6565 702d 416c 6976 650d 0a55 7365 722d  eep-Alive..User-
0000080: 4167 656e 743a 204d 6f7a 696c 6c61 2f32  Agent: Mozilla/2
0000090: 2e30 3120 284d 6163 696e 746f 7368 3b20  .01 (Macintosh;
00000a0: 493b 2036 384b 290d 0a48 6f73 743a 2077  I; 68K)..Host: w
00000b0: 7777 2e62 7563 6b6e 656c 6c2e 6564 750d  ww.bucknell.edu.
00000c0: 0a41 6363 6570 743a 2069 6d61 6765 2f67  .Accept: image/g
00000d0: 6966 2c20 696d 6167 652f 782d 7862 6974  if, image/x-xbit
00000e0: 6d61 702c 2069 6d61 6765 2f6a 7065 672c  map, image/jpeg,
00000f0: 2069 6d61 6765 2f70 6a70 6567 2c20 2a2f  image/pjpeg, */
0000100: 2a0d 0a0d 0acb                                     *.....
```

HTTP Response

```
>> IP PACKET *****
Next protocol = TCP, Transmission Control Protocol.
Source          = 134.82.7.4
Destination     = 134.82.18.89
>> TCP PACKET *****
source          port = 80, service = HTTP, HyperText Transfer Protocol.
PAYLOAD HEXDUMP:
0000000: 3c54 4954 4c45 3e42 7563 6b6e 656c 6c20 <TITLE>Bucknell
0000010: 486f 6d65 2050 6167 653c 2f54 4954 4c45 Home Page</TITLE>
0000020: 3e0a 3c62 6f64 7920 6261 636b 6772 6f75 >.<body backgrou
0000030: 6e64 3d22 6874 7470 3a2f 2f77 7777 2e62 nd="http://www.b
0000040: 7563 6b6e 656c 6c2e 6564 752f 7465 7874 ucknell.edu/text
0000050: 7572 6573 2f62 6163 6b67 726f 756e 6422 ures/background"
0000060: 3e0a 0a3c 5441 424c 4520 424f 5244 4552 >..<TABLE BORDER
0000070: 3d35 2077 6964 7468 3d31 3030 253e 0a3c =5 width=100%>.<
0000080: 5452 3e0a 3c54 4420 7661 6c69 676e 3d6d TR>.<TD valign=m
0000090: 6964 646c 6520 616c 6967 6e3d 6365 6e74 iddle align=cent
00000a0: 6572 3e3c 494d 4720 616c 743d 2242 7563 er><IMG alt="Buc
00000b0: 6b6e 656c 6c20 6c6f 676f 2220 5352 433d knell logo" SRC=
00000c0: 222f 6769 6673 2f42 7563 6b6e 656c 6c2e "/gifs/Bucknell.
```

Images

```
>> IP PACKET *****
```

```
Source      = 134.82.7.4
```

```
Destination = 134.82.18.89
```

```
>> TCP PACKET *****
```

```
source      port = 80, service = HTTP, HyperText Transfer Protocol.
```

```
destination port = 1652, service = ?
```

```
PAYLOAD HEXDUMP:
```

0000000:	4854 5450 2f31 2e30 2032 3030 2044 6f63	HTTP/1.0 200 Doc
0000010:	756d 656e 7420 666f 6c6c 6f77 730a 4461	ument follows.Da
0000020:	7465 3a20 5475 652c 2032 3120 4d61 7920	te: Tue, 21 May
0000030:	3139 3936 2031 333a 3233 3a33 3020 474d	1996 13:23:30 GM
0000040:	540a 5365 7276 6572 3a20 4e43 5341 2f31	T.Server: NCSA/1
0000050:	2e35 0a43 6f6e 7465 6e74 2d74 7970 653a	.5.Content-type:
0000060:	2069 6d61 6765 2f67 6966 0a4c 6173 742d	image/gif.Last-
0000070:	6d6f 6469 6669 6564 3a20 4d6f 6e2c 2031	modified: Mon, 1
0000080:	3420 4175 6720 3139 3935 2031 333a 3030	4 Aug 1995 13:00
0000090:	3a30 3620 474d 540a 436f 6e74 656e 742d	:06 GMT.Content-
00000a0:	6c65 6e67 7468 3a20 3138 3838 0a0a	length: 1888..

Retrieved Image

```
>> IP PACKET *****
```

```
Next protocol = TCP, Transmission Control Protocol.
```

```
Source          = 134.82.7.4
```

```
Destination     = 134.82.18.89
```

```
>> TCP PACKET *****
```

```
source          port = 80, service = HTTP, HyperText Transfer Protocol.
```

```
destination port = 1652, service = ?
```

```
PAYLOAD HEXDUMP:
```

```
0000000: 4749 4638 3961 a800 5200 f700 00ff ffff  GIF89a..R.....
0000010: ffff ccff ff99 ffff 66ff ff33 ffff 00ff  .....f..3....
0000020: ccff ffcc ccff cc99 ffcc 66ff cc33 ffcc  .....f..3..
0000030: 00ff 99ff ff99 ccff 9999 ff99 66ff 9933  .....f..3
0000040: ff99 00ff 66ff ff66 ccff 6699 ff66 66ff  ....f..f..f..ff.
0000050: 6633 ff66 00ff 33ff ff33 ccff 3399 ff33  f3.f..3..3..3..3
0000060: 66ff 3333 ff33 00ff 00ff ff00 ccff 0099  f.33.3.....
0000070: ff00 66ff 0033 ff00 00cc ffff ccff cccc  ..f..3.....
0000080: ff99 ccff 66cc ff33 ccff 00cc ccff cccc  ....f..3.....
0000090: cccc cc99 cccc 66cc cc33 cccc 00cc 99ff  .....f..3.....
00000a0: cc99 cccc 9999 cc99 66cc 9933 cc99 00cc  .....f..3....
00000b0: 66ff cc66 cccc 6699 cc66 66cc 6633 cc66  f..f..f..ff.f3.f
00000c0: 00cc 33ff cc33 cccc 3399 cc33 66cc 3333  ..3..3..3..3f.33
00000d0: cc33 00cc 00ff cc00 cccc 0099 cc00 66cc  .3.....f.
```

Documents

▀ Static

- HTML, XHTML
- Forms
- XML and XSL
- CSS

▀ Dynamic

- Server-side: CGI, PHP, JSP, ASP.
- Client-side: JavaScript, Java applets, ActiveX.