EE309 Advanced Programming Techniques for EE

Lecture 13: Network programming 2

INSU YUN (윤인수)

School of Electrical Engineering, KAIST

[Slides from 15-213: Introduction to Computer Systems at CMU]

Today

Material we didn't get to yesterday

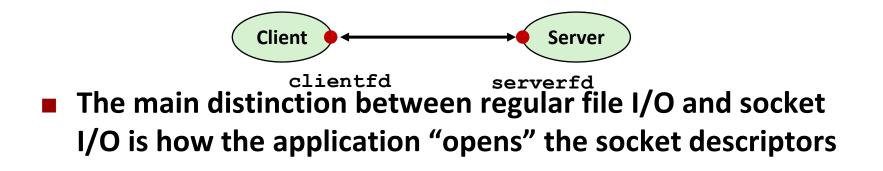
- Transmitting data using sockets
- Socket addresses
- getaddrinfo
- Setting up connections
- Application protocol example: HTTP

Sockets

What is a socket?

- To the kernel, a socket is an endpoint of communication
- To an application, a socket is a file descriptor that lets the application read/write from/to the network
- Using the FD abstraction lets you reuse code & interfaces

Clients and servers communicate with each other by reading from and writing to socket descriptors



Socket Programming Example

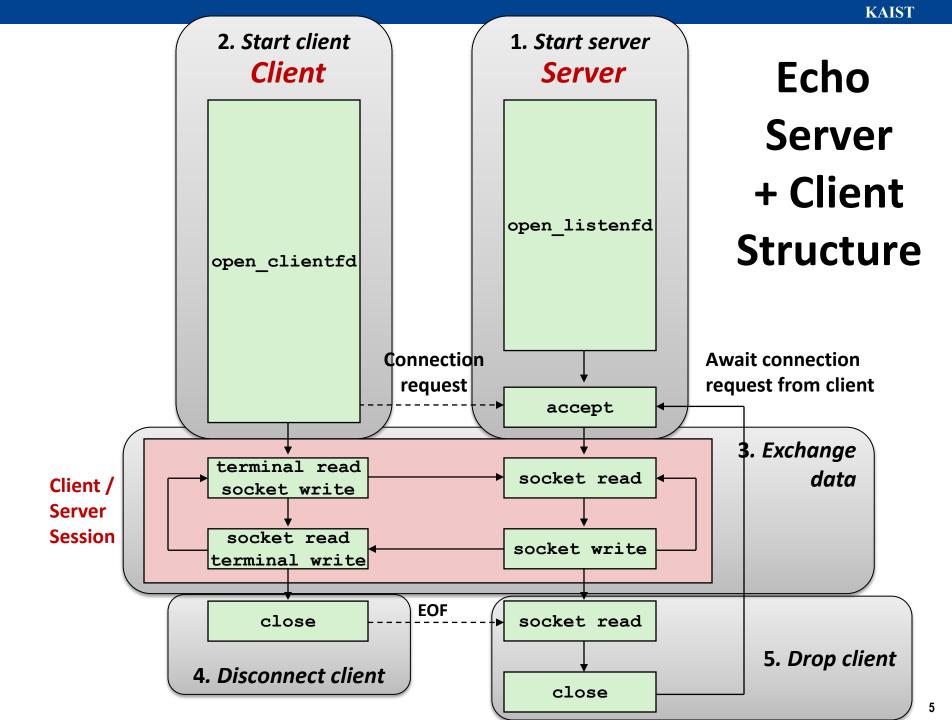
Echo server and client

Server

- Accepts connection request
- Repeats back lines as they are typed

Client

- Requests connection to server
- Repeatedly:
 - Read line from terminal
 - Send to server
 - Read reply from server
 - Print line to terminal



Recall: Unbuffered RIO Input/Output

- Same interface as Unix read and write
- **Especially useful for transferring data on network sockets**

```
#include "csapp.h"
ssize_t rio_readn(int fd, void *usrbuf, size_t n);
ssize_t rio_writen(int fd, void *usrbuf, size_t n);
```

Return: num. bytes transferred if OK, 0 on EOF (rio_readn only), -1 on error

- rio readn returns short count only if it encounters EOF
 - Only use it when you know how many bytes to read
- rio_writen never returns a short count
- Calls to rio_readn and rio_writen can be interleaved arbitrarily on the same descriptor

Recall: Buffered RIO Input Functions

 Efficiently read text lines and binary data from a file partially cached in an internal memory buffer

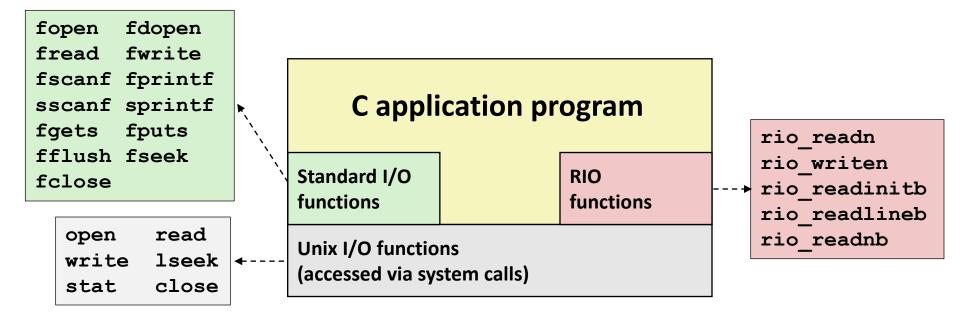
```
#include "csapp.h"
void rio_readinitb(rio_t *rp, int fd);
ssize_t rio_readlineb(rio_t *rp, void *usrbuf, size_t maxlen);
ssize_t rio_readnb(rio_t *rp, void *usrbuf, size_t n);
```

Return: num. bytes read if OK, 0 on EOF, -1 on error

- rio_readlineb reads a *text line* of up to maxlen bytes from file fd and stores the line in usrbuf
 - Especially useful for reading text lines from network sockets
- Stopping conditions
 - maxlen bytes read
 - EOF encountered
 - Newline ('\n') encountered

Today: Unix I/O, C Standard I/O, and RIO

- Two incompatible libraries building on Unix I/O
- Robust I/O (RIO): 213 special wrappers good coding practice: handles error checking, signals, and "short counts"



The RIO Package (213/CS:APP Package)

- RIO is a set of wrappers that provide efficient and robust I/O in apps, such as network programs that are subject to short counts
- RIO provides two different kinds of functions
 - Unbuffered input and output of binary data
 - rio_readn and rio_writen
 - Buffered input of text lines and binary data
 - rio_readlineb and rio_readnb
 - Buffered RIO routines are thread-safe and can be interleaved arbitrarily on the same descriptor
- Download from <u>http://csapp.cs.cmu.edu/3e/code.html</u>
 - → src/csapp.c and include/csapp.h

Implementation of rio_readn

```
/*
* rio readn - Robustly read n bytes (unbuffered)
*/
ssize t rio readn(int fd, void *usrbuf, size t n)
{
   size t nleft = n;
   ssize t nread;
   char *bufp = usrbuf;
   while (nleft > 0) {
      if ((nread = read(fd, bufp, nleft)) < 0) {</pre>
          if (errno == EINTR) /* Interrupted by sig handler return */
             else
             return -1; /* errno set by read() */
      }
      else if (nread == 0)
                            /* EOF */
         break;
      nleft -= nread;
      bufp += nread;
   }
                         /* Return >= 0 */
   return (n - nleft);
                                                         csapp
```

Echo Client: Main Routine

```
#include "csapp.h"
int main(int argc, char **argv)
{
    int clientfd;
    char *host, *port, buf[MAXLINE];
    rio t rio;
   host = argv[1];
   port = argv[2];
    clientfd = Open clientfd(host, port);
    Rio readinitb(&rio, clientfd);
    while (Fgets(buf, MAXLINE, stdin) != NULL) {
       Rio writen(clientfd, buf, strlen(buf));
       Rio readlineb(&rio, buf, MAXLINE);
       Fputs(buf, stdout);
    Close (clientfd);
    exit(0);
                                                  echoclient.c
```

Echo Server: echo function

- The server uses RIO to read and echo text lines until EOF (end-of-file) condition is encountered.
 - EOF condition caused by client calling close (clientfd)

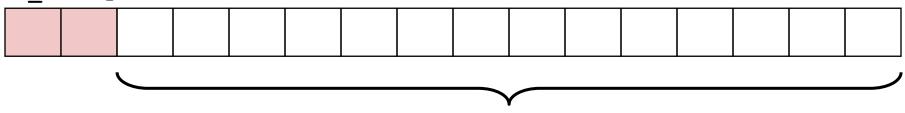
```
void echo(int connfd)
{
    size_t n;
    char buf[MAXLINE];
    rio_t rio;
    Rio_readinitb(&rio, connfd);
    while((n = Rio_readlineb(&rio, buf, MAXLINE)) != 0) {
        printf("server received %d bytes\n", (int)n);
        Rio_writen(connfd, buf, n);
    }
} echo.c
```

Socket Address Structures

- Generic socket address:
 - For address arguments to connect, bind, and accept (next lecture)
 - In C++ this would be an abstract base class
 - For casting convenience, we adopt the Stevens convention:
 typedef struct sockaddr SA;

```
struct sockaddr {
    uint16_t sa_family; /* Protocol family */
    char sa_data[14]; /* Address data */
};
```

sa family

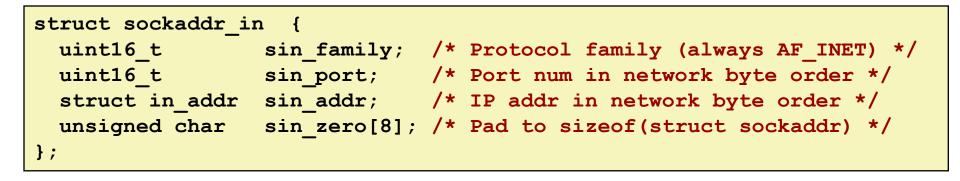


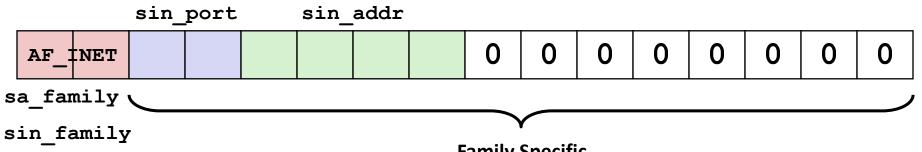
Family Specific

Socket Address Structures

Internet (IPv4) specific socket address:

Must cast (struct sockaddr_in *) to (struct sockaddr *) for functions that take socket address arguments.





Host and Service Conversion: getaddrinfo

- getaddrinfo is the modern way to convert string representations of hostnames, host addresses, ports, and service names to socket address structures.
 - Replaces obsolete gethostbyname and getservbyname funcs.

Advantages:

- Reentrant (can be safely used by threaded programs).
- Allows us to write portable protocol-independent code
 - Works with both IPv4 and IPv6

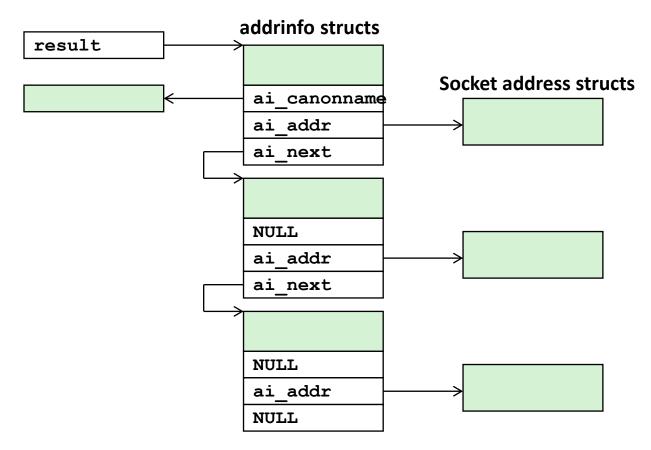
Disadvantages

- Somewhat complex
- Fortunately, a small number of usage patterns suffice in most cases.

Host and Service Conversion: getaddrinfo

- Given host and service, getaddrinfo returns result that points to a linked list of addrinfo structs, each of which points to a corresponding socket address struct, and which contains arguments for the sockets interface functions.
- Helper functions:
 - freeadderinfo frees the entire linked list.
 - gai_strerror converts error code to an error message.

Linked List Returned by getaddrinfo



addrinfo Struct

<pre>struct addrinfo {</pre>		
int	<pre>ai_flags;</pre>	/* Hints argument flags */
int	ai_family;	<pre>/* First arg to socket function */</pre>
int	<pre>ai_socktype;</pre>	<pre>/* Second arg to socket function */</pre>
int	<pre>ai_protocol;</pre>	<pre>/* Third arg to socket function */</pre>
char	<pre>*ai_canonname;</pre>	/* Canonical host name */
size_t	<pre>ai_addrlen;</pre>	<pre>/* Size of ai_addr struct */</pre>
struct sockaddr	<pre>*ai_addr;</pre>	/* Ptr to socket address structure */
struct addrinfo	<pre>*ai_next;</pre>	<pre>/* Ptr to next item in linked list */</pre>
};		

- Each addrinfo struct returned by getaddrinfo contains arguments that can be passed directly to socket function.
- Also points to a socket address struct that can be passed directly to connect and bind functions.

(socket, connect, bind to be discussed next)

Host and Service Conversion: getnameinfo

- getnameinfo is the inverse of getaddrinfo, converting a socket address to the corresponding host and service.
 - Replaces obsolete gethostbyaddr and getservbyport funcs.
 - Reentrant and protocol independent.

int getnameinfo(const SA *sa, socklen_t salen,	<pre>/* In: socket addr */</pre>
char *host, size_t hostlen,	/* Out: host */
char *serv, size_t servlen,	/* Out: service */
<pre>int flags);</pre>	<pre>/* optional flags */</pre>

Conversion Example

```
#include "csapp.h"
int main(int argc, char **argv)
Ł
    struct addrinfo *p, *listp, hints;
    char buf[MAXLINE];
    int rc, flags;
    /* Get a list of addrinfo records */
   memset(&hints, 0, sizeof(struct addrinfo));
   hints.ai family = AF INET; /* IPv4 only */
   hints.ai socktype = SOCK STREAM; /* Connections only */
    if ((rc = getaddrinfo(argv[1], NULL, &hints, &listp)) != 0) {
        fprintf(stderr, "getaddrinfo error: %s\n", gai strerror(rc));
       exit(1);
    }
                                                               hostinfo.c
```

Conversion Example (cont)

Running hostinfo

whaleshark> ./hostinfo localhost
127.0.0.1

whaleshark> ./hostinfo whaleshark.ics.cs.cmu.edu
128.2.210.175

whaleshark> ./hostinfo twitter.com
199.16.156.230
199.16.156.38
199.16.156.102
199.16.156.198

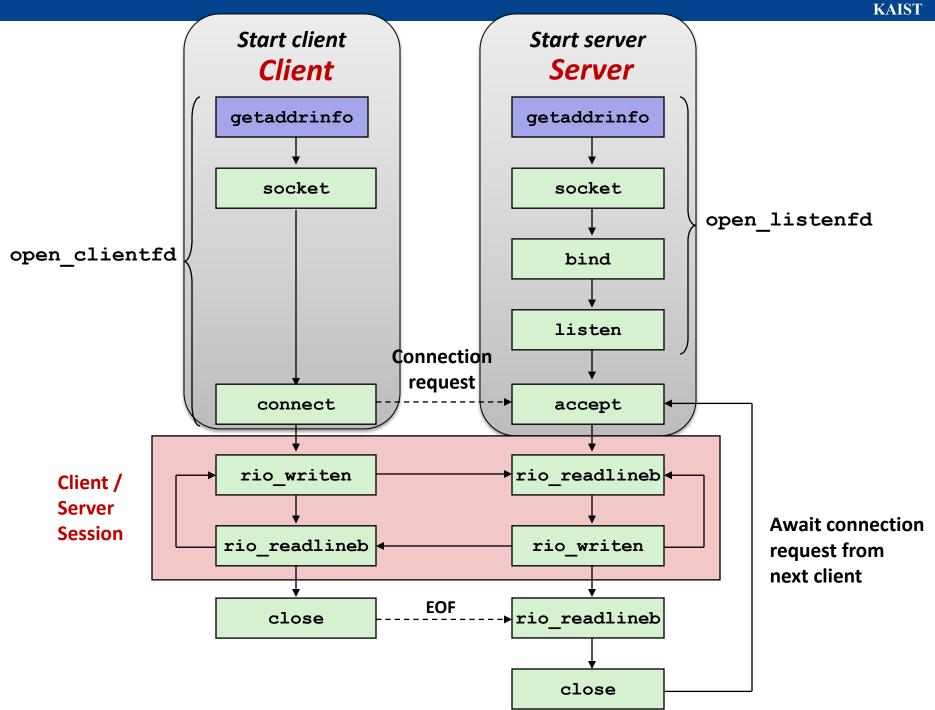
whaleshark> ./hostinfo google.com
172.217.15.110
2607:f8b0:4004:802::200e

Today

- Questions from yesterday
- Material we didn't get to yesterday
 - Transmitting data using sockets
 - Socket addresses
 - getaddrinfo

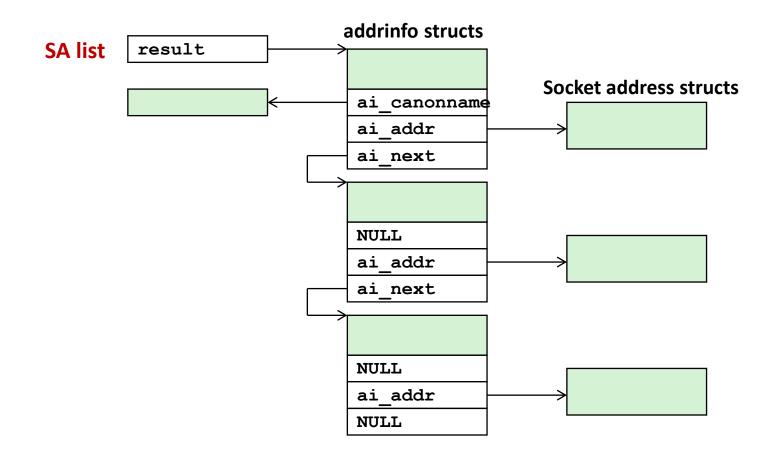
Setting up connections

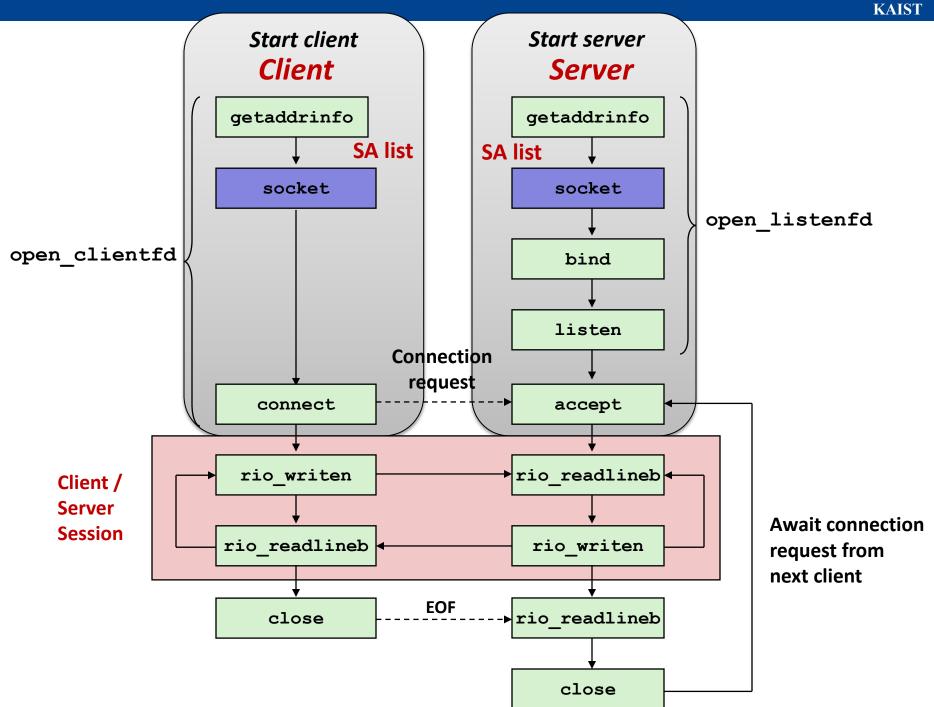
Application protocol example: HTTP



Review: getaddrinfo

 getaddrinfo converts string representations of hostnames, host addresses, ports, service names to socket address structures





Sockets Interface: socket

Clients and servers use the socket function to create a socket descriptor:

int socket(int domain, int type, int protocol)

Example:



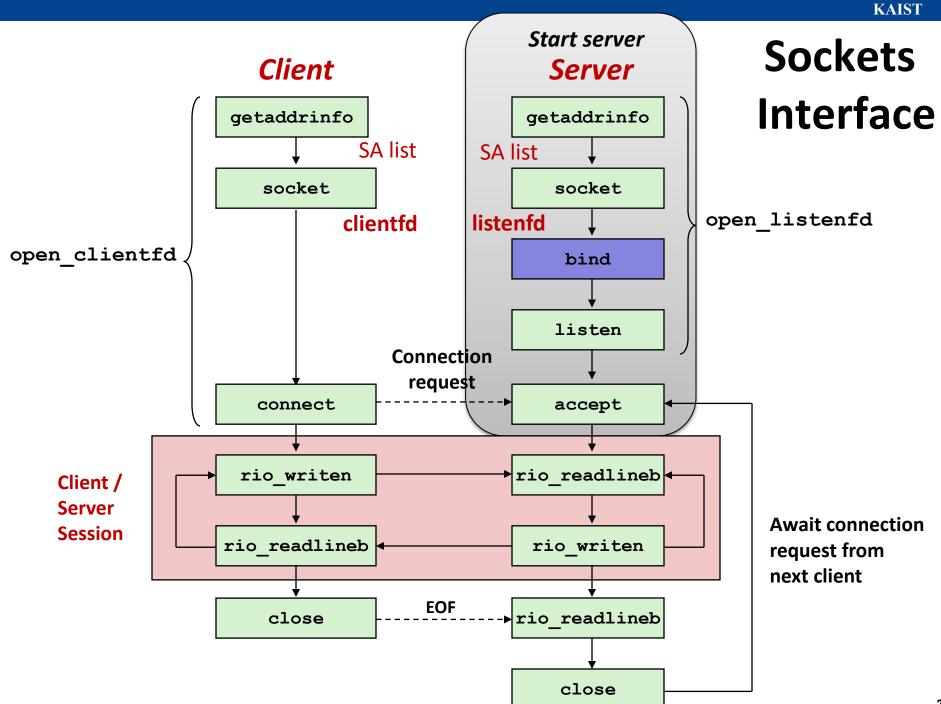
Protocol specific!

Indicates that we are using 32-bit IPV4 addresses

Indicates that the socket will be the end point of a reliable (TCP) connection

Example:

Use getaddrinfo and you don't have to know or care which protocol!



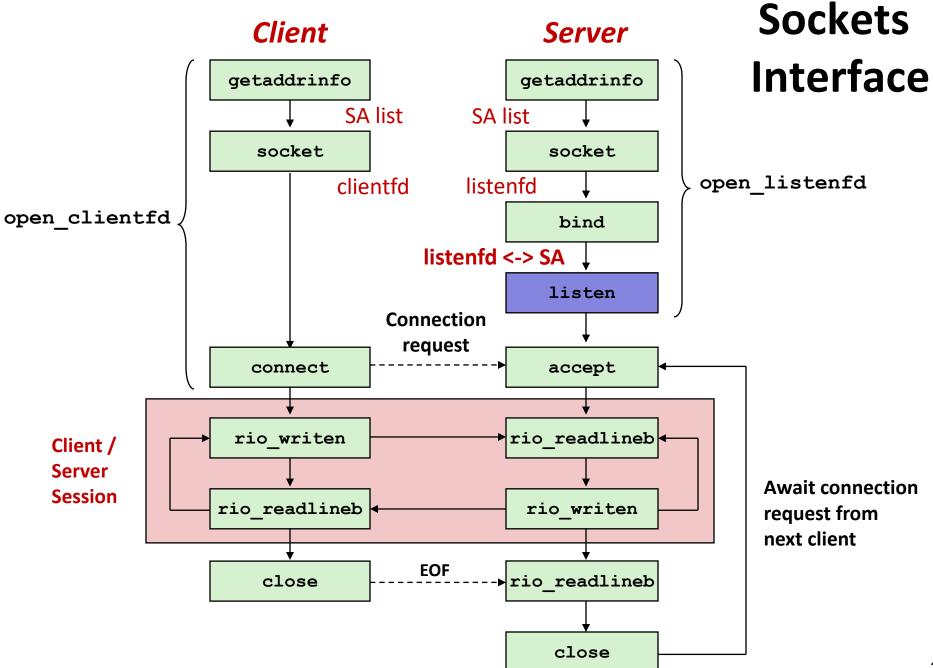
Sockets Interface: bind

A server uses bind to ask the kernel to associate the server's socket address with a socket descriptor:

int bind(int sockfd, SA *addr, socklen_t addrlen);

Our convention: typedef struct sockaddr SA;

- Process can read bytes that arrive on the connection whose endpoint is addr by reading from descriptor sockfd
- Similarly, writes to sockfd are transferred along connection whose endpoint is addr
- Best practice is to use getaddrinfo to supply the arguments addr and addrlen.

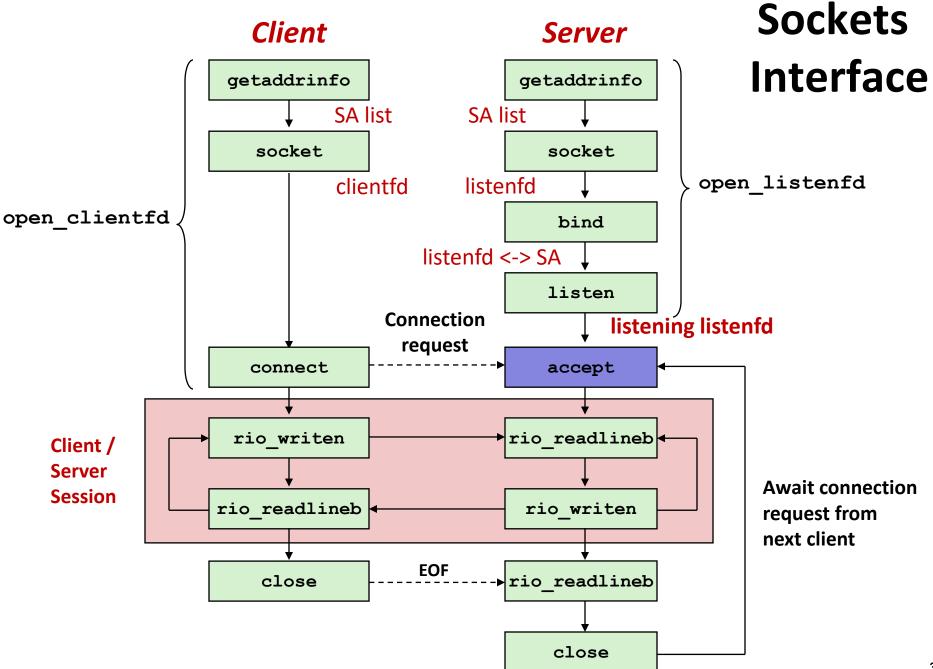


Sockets Interface: listen

- Kernel assumes that descriptor from socket function is an active socket that will be on the client end
- A server calls the listen function to tell the kernel that a descriptor will be used by a server rather than a client:

int listen(int sockfd, int backlog);

- Converts sockfd from an active socket to a *listening* socket that can accept connection requests from clients.
- backlog is a hint about the number of outstanding connection requests that the kernel should queue up before starting to refuse requests (128-ish by default)

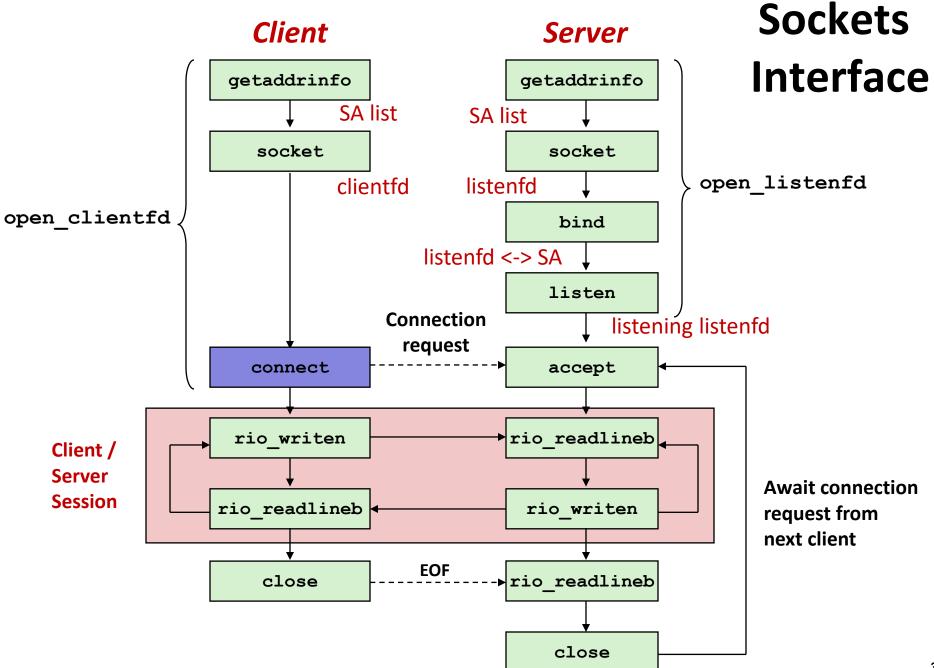


Sockets Interface: accept

 Servers wait for connection requests from clients by calling accept:

int accept(int listenfd, SA *addr, int *addrlen);

- Waits for connection request to arrive on the connection bound to listenfd, then fills in client's socket address in addr and size of the socket address in addrlen.
- Returns a connected descriptor connfd that can be used to communicate with the client via Unix I/O routines.



Sockets Interface: connect

A client establishes a connection with a server by calling connect:

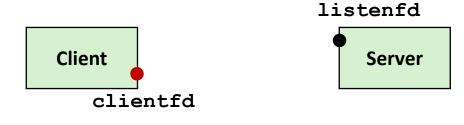
int connect(int clientfd, SA *addr, socklen_t addrlen);

- Attempts to establish a connection with server at socket address addr
 - If successful, then clientfd is now ready for reading and writing.
 - Resulting connection is characterized by socket pair

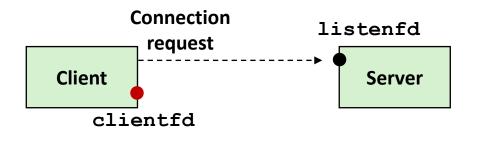
(x:y, addr.sin_addr:addr.sin_port)

- x is client address
- y is ephemeral port that uniquely identifies client process on client host
- Best practice is to use getaddrinfo to supply the arguments addr and addrlen.

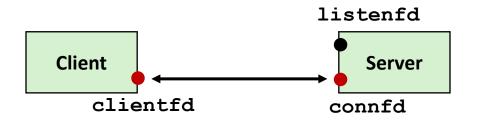
connect/accept Illustrated



1. Server blocks in accept, waiting for connection request on listening descriptor listenfd



2. Client makes connection request by calling and blocking in connect



3. Server returns connfd from accept. Client returns from connect. Connection is now established between clientfd and connfd

Connected vs. Listening Descriptors

Listening descriptor

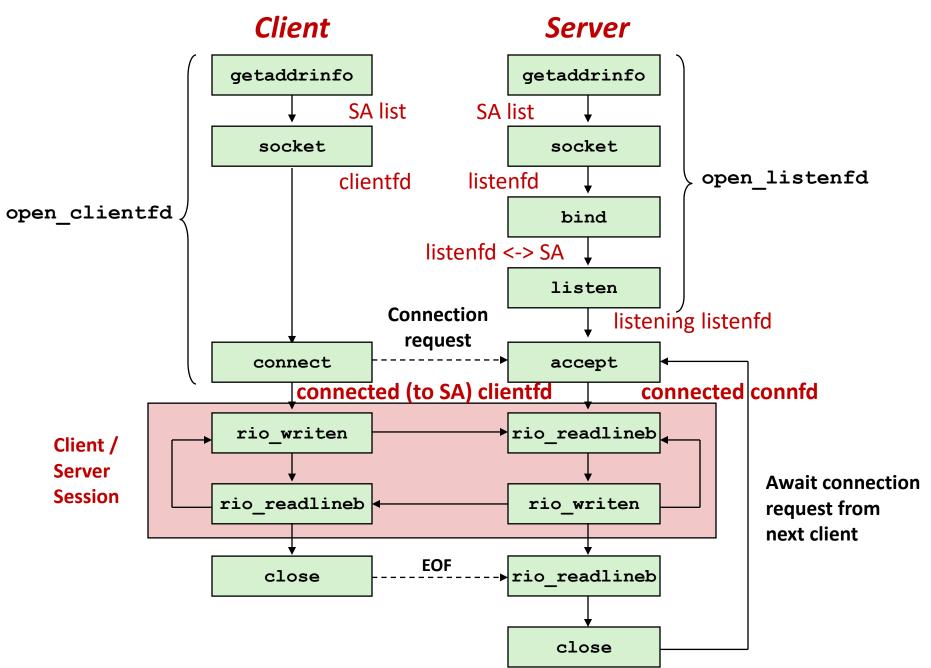
- End point for client connection <u>requests</u>
- Created once and exists for lifetime of the server

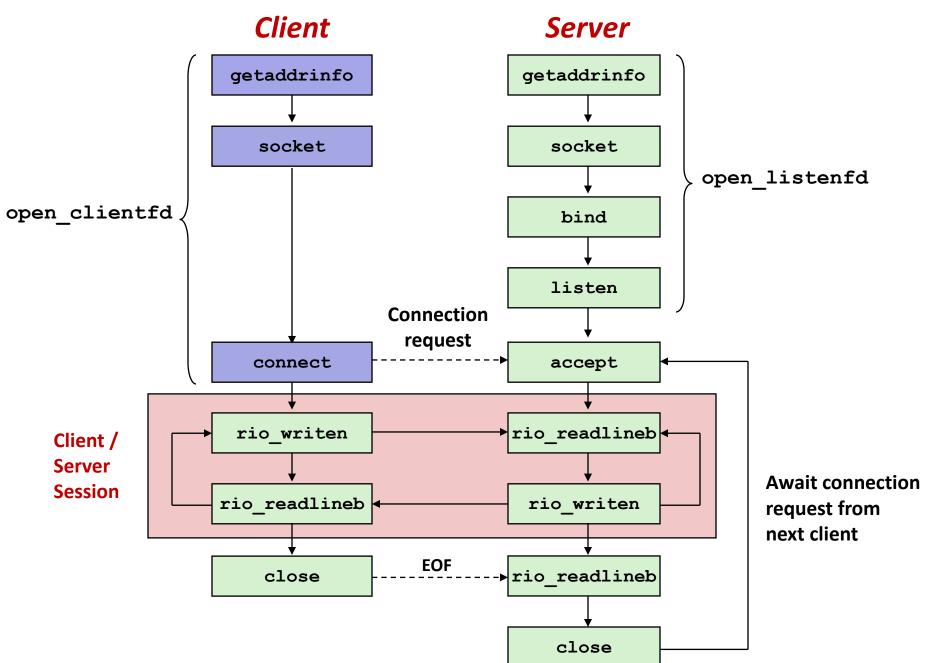
Connected descriptor

- End point of the <u>connection</u> between client and server
- A new descriptor is created each time the server accepts a connection request from a client
- Exists only as long as it takes to service client

Why the distinction?

- Allows for concurrent servers that can communicate over many client connections simultaneously
 - E.g., Each time we receive a new request, we fork a child to handle the request





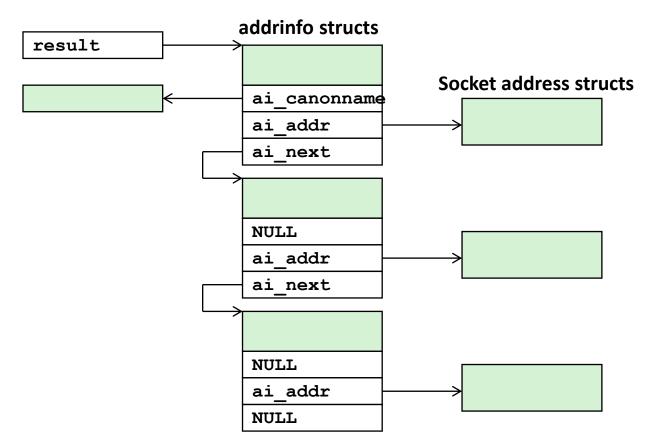
Sockets Helper: open_clientfd

Establish a connection with a server

```
int open_clientfd(char *hostname, char *port) {
    int clientfd;
    struct addrinfo hints, *listp, *p;
    /* Get a list of potential server addresses */
    memset(&hints, 0, sizeof(struct addrinfo));
    hints.ai_socktype = SOCK_STREAM; /* Open a connection */
    hints.ai_flags = AI_NUMERICSERV; /* ...using numeric port arg. */
    hints.ai_flags |= AI_ADDRCONFIG; /* Recommended for connections */
    Getaddrinfo(hostname, port, &hints, &listp);
    CSapp.C
```

AI_ADDRCONFIG means "use whichever of IPv4 and IPv6 works on this computer". Good practice for clients, not for servers.

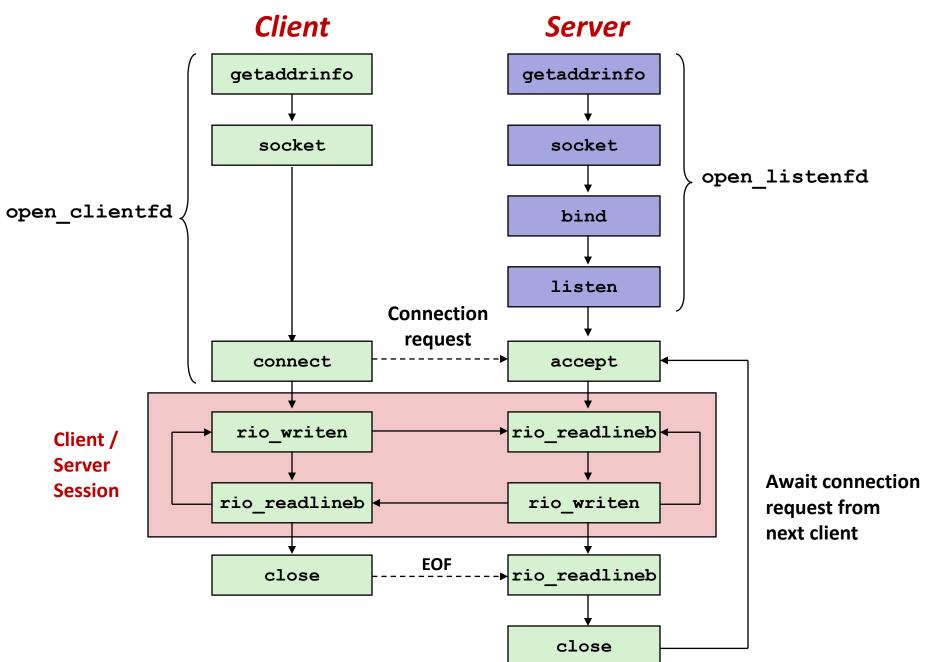
getaddrinfo



- Clients: walk this list, trying each socket address in turn, until the calls to socket and connect succeed.
- Servers: walk the list calling socket, listen, bind for all addresses, then use select to accept connections on any of them (beyond our scope)

Sockets Helper: open_clientfd (cont)

```
/* Walk the list for one that we can successfully connect to */
for (p = listp; p; p = p->ai next) {
   /* Create a socket descriptor */
    if ((clientfd = socket(p->ai family, p->ai socktype,
                           p->ai protocol)) < 0)
        continue; /* Socket failed, try the next */
    /* Connect to the server */
    if (connect(clientfd, p->ai addr, p->ai addrlen) != -1)
       break: /* Success */
    Close(clientfd); /* Connect failed, try another */
}
/* Clean up */
Freeaddrinfo(listp);
if (!p) /* All connects failed */
   return -1;
else /* The last connect succeeded */
    return clientfd;
                                                           csapp.c
```



Sockets Helper: open_listenfd

Create a listening descriptor that can be used to accept connection requests from clients.

AI_PASSIVE means "I plan to listen on this socket." AI_ADDRCONFIG normally not used for servers, but we use it for convenience

Sockets Helper: open_listenfd (cont)

```
/* Walk the list for one that we can bind to */
for (p = listp; p; p = p-ai next) {
   /* Create a socket descriptor */
    if ((listenfd = socket(p->ai family, p->ai socktype,
                           p->ai protocol)) < 0)
        continue; /* Socket failed, try the next */
   /* Eliminates "Address already in use" error from bind */
   Setsockopt(listenfd, SOL SOCKET, SO REUSEADDR,
               (const void *)&optval , sizeof(int));
    /* Bind the descriptor to the address */
    if (bind(listenfd, p->ai addr, p->ai addrlen) == 0)
       break; /* Success */
   Close(listenfd); /* Bind failed, try the next */
}
                                                         csapp.c
```

A production server would not break out of the loop on the first success. We do that for simplicity only.

Sockets Helper: open_listenfd (cont)

```
/* Clean up */
Freeaddrinfo(listp);
if (!p) /* No address worked */
   return -1;
/* Make it a listening socket ready to accept conn. requests */
if (listen(listenfd, LISTENQ) < 0) {
   Close(listenfd);
   return -1;
}
return listenfd;
CSapp.c</pre>
```

Key point: open_clientfd and open_listenfd are both independent of any particular version of IP.

Testing Servers Using telnet

- The telnet program is invaluable for testing servers that transmit ASCII strings over Internet connections
 - Our simple echo server
 - Web servers
 - Mail servers
- Usage:
 - Iinux> telnet <host> <portnumber>
 - Creates a connection with a server running on <host> and listening on port <portnumber>

Testing the Echo Server With telnet

whaleshark> ./echoserveri 15213 Connected to (MAKOSHARK.ICS.CS.CMU.EDU, 50280) server received 11 bytes server received 8 bytes

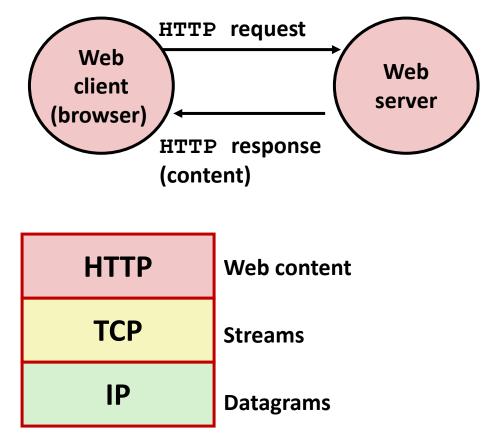
```
makoshark> telnet whaleshark.ics.cs.cmu.edu 15213
Trying 128.2.210.175...
Connected to whaleshark.ics.cs.cmu.edu (128.2.210.175).
Escape character is '^]'.
Hi there!
Hi there!
Howdy!
Howdy!
howdy!
^]
telnet> quit
Connection closed.
makoshark>
```

Today

- Questions from yesterday
- Material we didn't get to yesterday
 - Transmitting data using sockets
 - Socket addresses
 - getaddrinfo
- Setting up connections
- Application protocol example: HTTP

Web Server Basics

- Clients and servers communicate using the HyperText Transfer Protocol (HTTP)
 - Client and server establish TCP connection
 - Client requests content
 - Server responds with requested content
 - Client and server close connection (eventually)
- Current version is HTTP/2.0 but HTTP/1.1 widely used still
 - RFC 2616, June, 1999.



http://www.w3.org/Protocols/rfc2616/rfc2616.html

Web Content

Web servers return *content* to clients

 content: a sequence of bytes with an associated MIME (Multipurpose Internet Mail Extensions) type

Example MIME types

- text/html
- text/plain
- image/gif
- image/png
- image/jpeg

HTML document Unformatted text Binary image encoded in GIF format Binary image encoded in PNG format Binary image encoded in JPEG format

You can find the complete list of MIME types at: http://www.iana.org/assignments/media-types/media-types.xhtml

Static and Dynamic Content

The content returned in HTTP responses can be either static or dynamic

- Static content: content stored in files and retrieved in response to an HTTP request
 - Examples: HTML files, images, audio clips, Javascript programs
 - Request identifies which content file
- Dynamic content: content produced on-the-fly in response to an HTTP request
 - Example: content produced by a program executed by the server on behalf of the client
 - Request identifies file containing executable code

Web content associated with a file that is managed by the server

URLs and how clients and servers use them

- Unique name for a file: URL (Universal Resource Locator)
- Example URL: http://www.cmu.edu:80/index.html
- Clients use prefix (http://www.cmu.edu:80) to infer:
 - What kind (protocol) of server to contact (HTTP)
 - Where the server is (www.cmu.edu)
 - What port it is listening on (80)
- Servers use suffix (/index.html) to:
 - Determine if request is for static or dynamic content.
 - No hard and fast rules for this
 - One convention: executables reside in **cgi-bin** directory
 - Find file on file system
 - Initial "/" in suffix denotes home directory for requested content.
 - Minimal suffix is "/", which server expands to configured default filename (usually, index.html)

HTTP Request Example

GET /	HTTP/1.1
Host:	www.cmu.edu

Client: request line Client: required HTTP/1.1 header Client: blank line terminates headers

- HTTP standard requires that each text line end with $\\r\n''$
- Blank line (``\r\n") terminates request and response headers

HTTP Requests

 HTTP request is a *request line*, followed by zero or more request headers

Request line: <method> <uri> <version>

- <method> is one of GET, POST, OPTIONS, HEAD, PUT,
 DELETE, or TRACE
- <uri>is typically URL for proxies, URL suffix for servers
 - A URL is a type of URI (Uniform Resource Identifier)
 - See <u>http://www.ietf.org/rfc/rfc2396.txt</u>
- **version>** is HTTP version of request (HTTP/1.0 or HTTP/1.1)

Request headers: <header name>: <header data>

Provide additional information to the server

HTTP Responses

 HTTP response is a response line followed by zero or more response headers, possibly followed by content, with blank line ("\r\n") separating headers from content.

Response line:

- <version> <status code> <status msg>
- <version> is HTTP version of the response
- <status code> is numeric status
- <status msg> is corresponding English text
 - 200 OK Request was handled without error
 - **301 Moved** Provide alternate URL
 - **404** Not found Server couldn't find the file

Response headers: <header name>: <header data>

- Provide additional information about response
- **Content-Type:** MIME type of content in response body
- Content-Length: Length of content in response body

Example HTTP Transaction

```
whaleshark> telnet www.cmu.edu 80
                                         Client: open connection to server
Trying 128.2.42.52...
                                         Telnet prints 3 lines to terminal
Connected to WWW-CMU-PROD-VIP.ANDREW.cmu.edu.
Escape character is '^]'.
GET / HTTP/1.1
                                         Client: request line
Host: www.cmu.edu
                                         Client: required HTTP/1.1 header
                                         Client: blank line terminates headers
HTTP/1.1 301 Moved Permanently
                                         Server: response line
Date: Wed, 05 Nov 2014 17:05:11 GMT
                                         Server: followed by 5 response headers
Server: Apache/1.3.42 (Unix)
                                         Server: this is an Apache server
Location: http://www.cmu.edu/index.shtml Server: page has moved here
Transfer-Encoding: chunked
                                         Server: response body will be chunked
Content-Type: text/html; charset=...
                                         Server: expect HTML in response body
                                         Server: empty line terminates headers
                                         Server: first line in response body
15c
<HTML><HEAD>
                                         Server: start of HTML content
</BODY></HTML>
                                         Server: end of HTML content
                                         Server: last line in response body
0
Connection closed by foreign host.
                                         Server: closes connection
```

- **HTTP** standard requires that each text line end with " $r\n''$
- Blank line (``\r\n") terminates request and response headers

Example HTTP Transaction, Take 2

```
whaleshark> telnet www.cmu.edu 80
                                         Client: open connection to server
Trying 128.2.42.52...
                                         Telnet prints 3 lines to terminal
Connected to WWW-CMU-PROD-VIP.ANDREW.cmu.edu.
Escape character is '^]'.
GET /index.shtml HTTP/1.1
                                         Client: request line
Host: www.cmu.edu
                                         Client: required HTTP/1.1 header
                                         Client: blank line terminates headers
HTTP/1.1 200 OK
                                         Server: response line
Date: Wed, 05 Nov 2014 17:37:26 GMT
                                         Server: followed by 4 response headers
Server: Apache/1.3.42 (Unix)
Transfer-Encoding: chunked
Content-Type: text/html; charset=...
                                         Server: empty line terminates headers
1000
                                         Server: begin response body
<html ..>
                                         Server: first line of HTML content
</html>
                                         Server: end response body
0
Connection closed by foreign host.
                                         Server: close connection
```

Example HTTP(S) Transaction, Take 3

whaleshark> openssl s_client www.cs.cmu.edu:443
CONNECTED(00000005)

•••

Certificate chain

••

Server certificate

```
----BEGIN CERTIFICATE-----
```

MIIGDjCCBPagAwIBAgIRAMiF7LBPDoySilnNoU+mp+gwDQYJKoZIhvcNAQELBQAw djELMAkGA1UEBhMCVVMxCzAJBgNVBAgTAk1JMRIwEAYDVQQHEwlBbm4gQXJib3Ix EjAQBgNVBAoTCUludGVybmV0MjERMA8GA1UECxMISW5Db21tb24xHzAdBgNVBAMT wkWkvDVBBCwKXrShVxQNsj6J

```
----END CERTIFICATE----
```

```
subject=/C=US/postalCode=15213/ST=PA/L=Pittsburgh/street=5000 Forbes
Ave/O=Carnegie Mellon University/OU=School of Computer
Science/CN=www.cs.cmu.edu issuer=/C=US/ST=MI/L=Ann
Arbor/O=Internet2/OU=InCommon/CN=InCommon RSA Server CA
SSL handshake has read 6274 bytes and written 483 bytes
...
>GET / HTTP/1.0
HTTP/1.1 200 OK
Date: Tue, 12 Nov 2019 04:22:15 GMT
Server: Apache/2.4.10 (Ubuntu)
Set-Cookie: SHIBLOCATION=scsweb; path=/; domain=.cs.cmu.edu
```

```
... HTML Content Continues Below ...
```