15-213

"The course that gives CMU its Zip!"

Internetworking
Nov 20, 2001

Topics

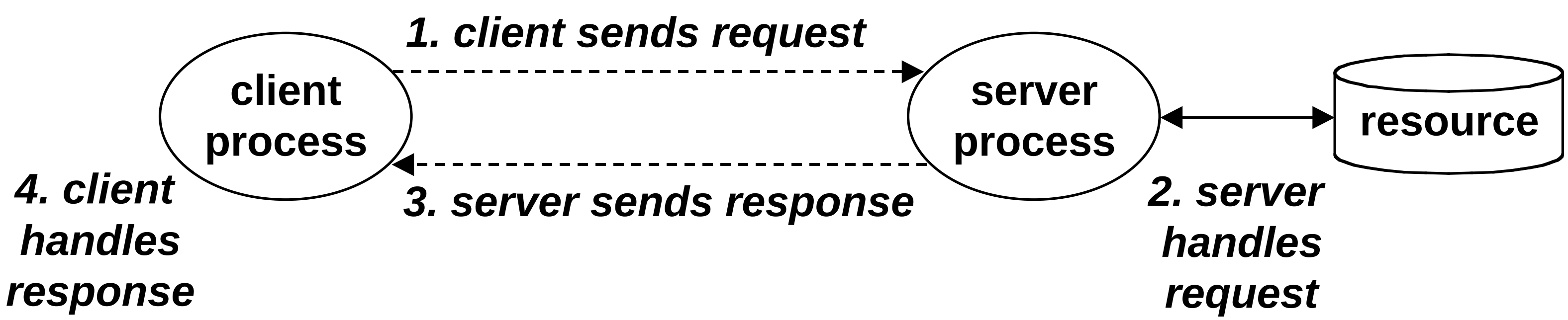
- **internets**
- **The Global IP Internet**
- **Programmer's view of the Internet**

`class25.ppt`

A client-server transaction

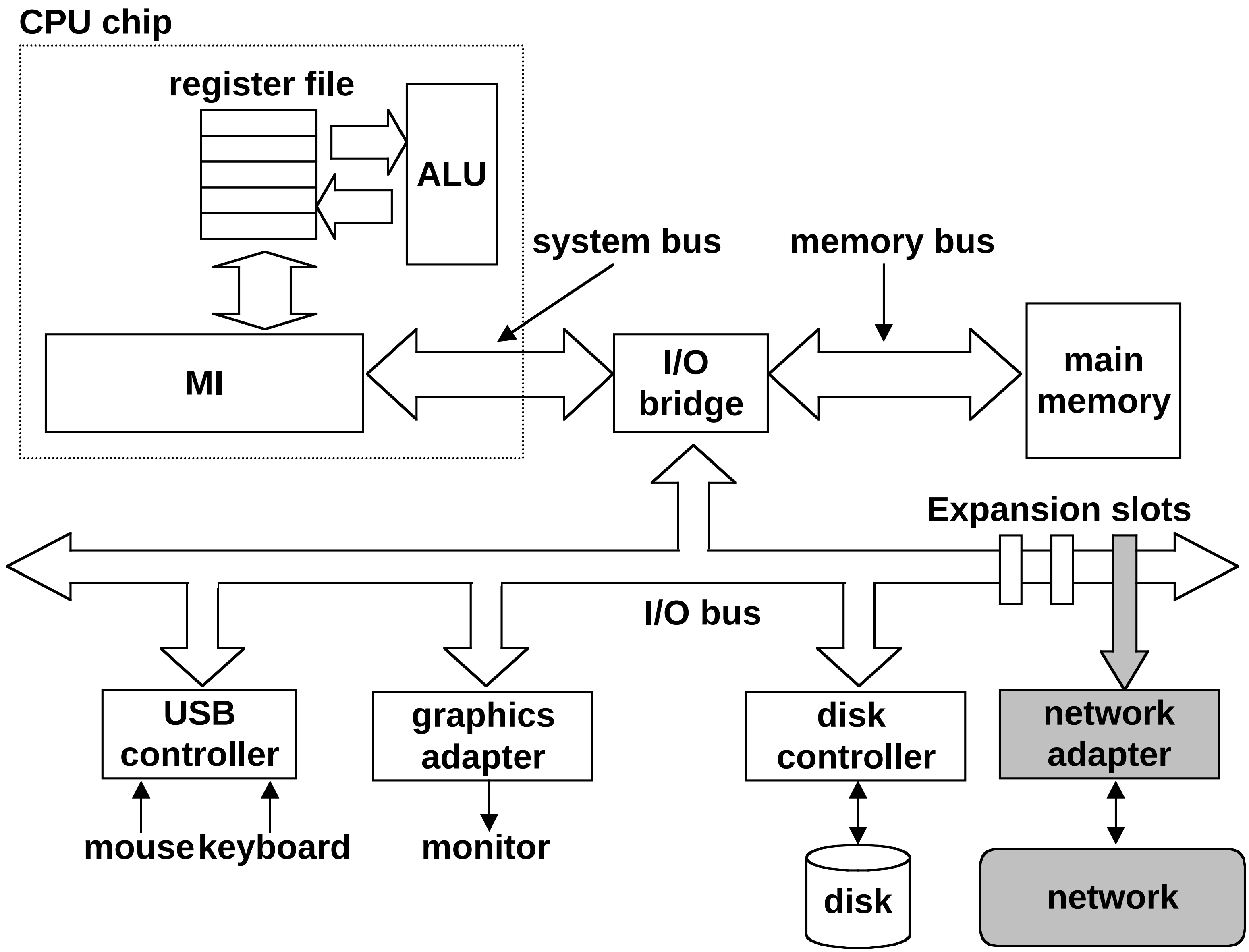
Every network application is based on the client-server model:

- a *server* process and one or more *client* processes
- server manages some *resource*.
- server provides *service* by manipulating resource for *clients*.



Note: clients and servers are processes running on hosts (can be the same or different hosts).

Hardware organization of a network host



Computer networks

A network is a hierarchical system of boxes and wires organized by geographical proximity

- LAN (local area network) spans a building or campus.
  - Ethernet is most prominent example.
- WAN (wide-area network) spans country or world.
  - typically high-speed point-to-point phone lines.

An *internetwork (internet)* is an interconnected set of networks.

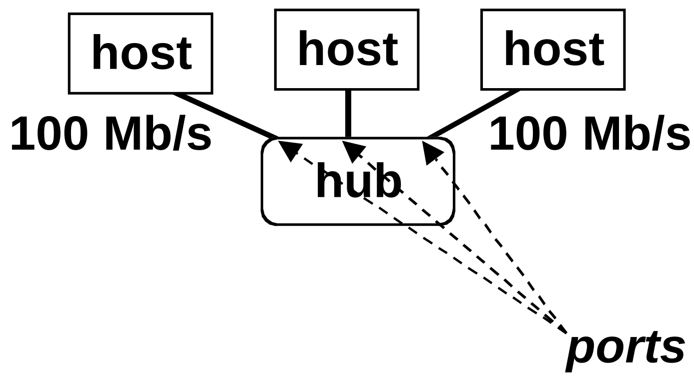
- The IP Internet is the most famous example of an internetwork.

Let's see how we would build an internetwork from the ground up.

Lowest level: Ethernet segment

Ethernet segment consists of a collection of *hosts* connected by wires (twisted pairs) to a *hub.*

Spans room or floor in a building.



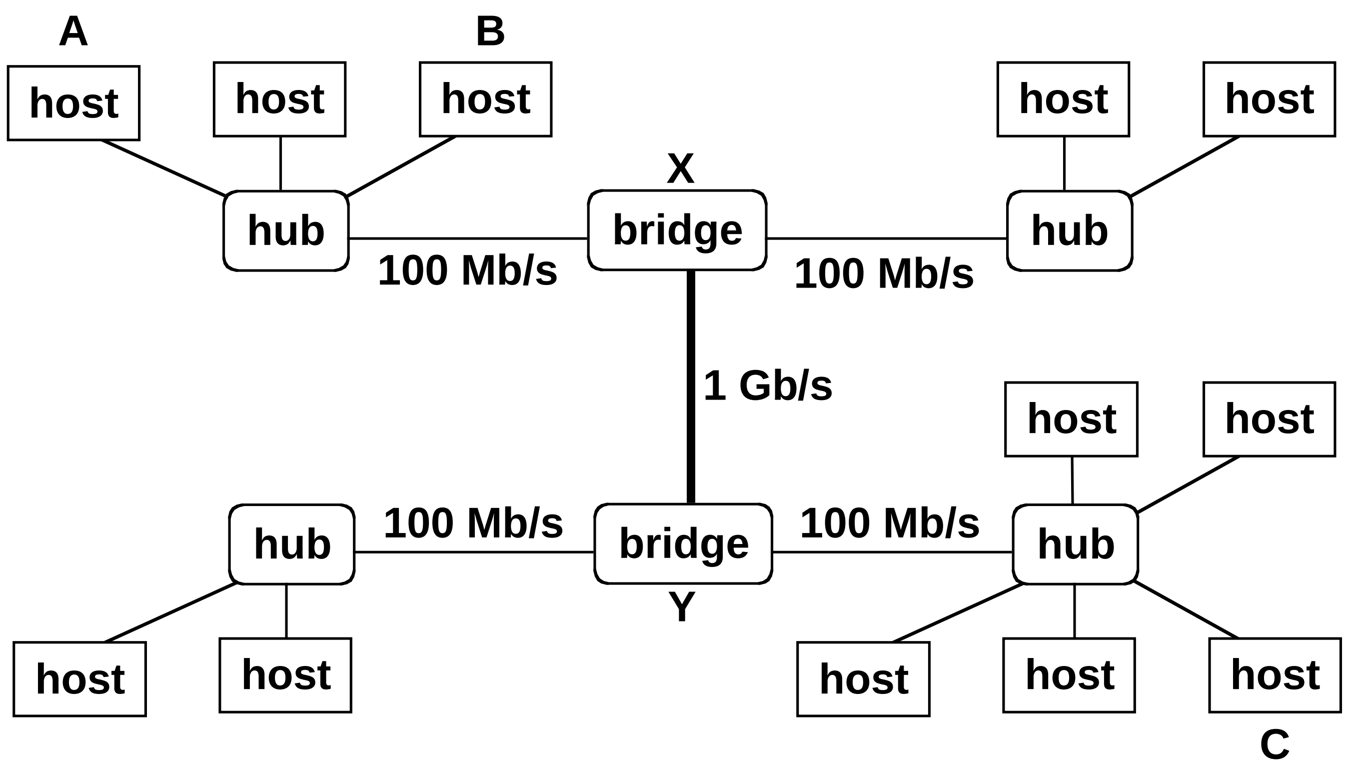
Operation

- Each Ethernet adapter has a unique 48-bit address.
- Hosts send bits to any other host in chunks called *frames.*
- Hub slavishly copies each bit from each port to every other port.
  - every host sees every bit.

Next level: Bridged Ethernet segment

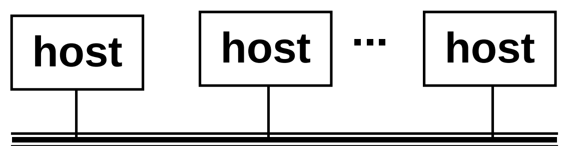
Spans building or campus.

Bridges cleverly learn which hosts are reachable from which ports and then selectively copy frames from port to port.



Conceptual view of LANs

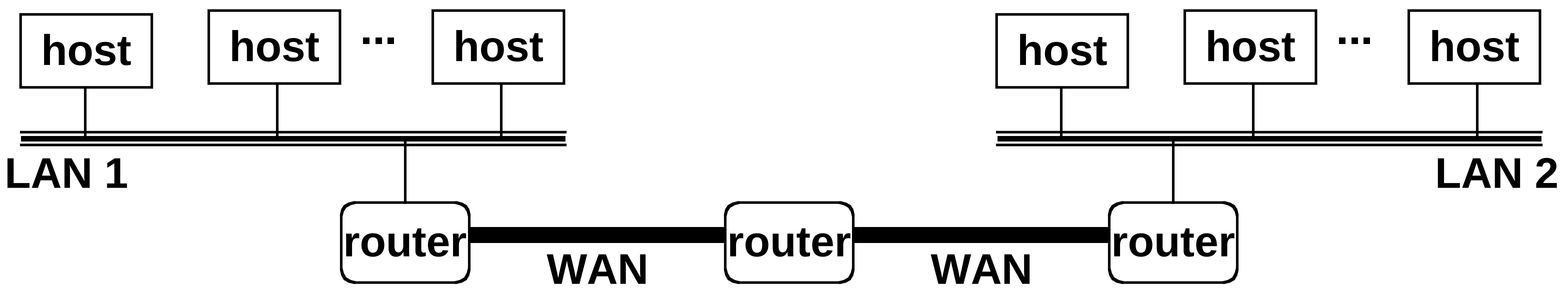
For simplicity, hubs, bridges, and wires are often shown as a collection of hosts attached to a single wire:



Next level: internets

Multiple incompatible LANs can be physically connected by specialized computers called *routers.*

The connected networks are called an *internet*.



LAN 1 and LAN 2 might be completely different, totally incompatible LANs (e.g., Ethernet and ATM)

The notion of an internet protocol

How is it possible to send bits across incompatible LANs and WANs?

Solution: *protocol software* running on each host and router smooths out the differences between the different networks.

Implements an *internet protocol* (i.e., set of rules) that governs how hosts and routers should cooperate when they transfer data from network to network.

- **TCP/IP is the protocol for the global IP Internet.**

What does an internet protocol do?

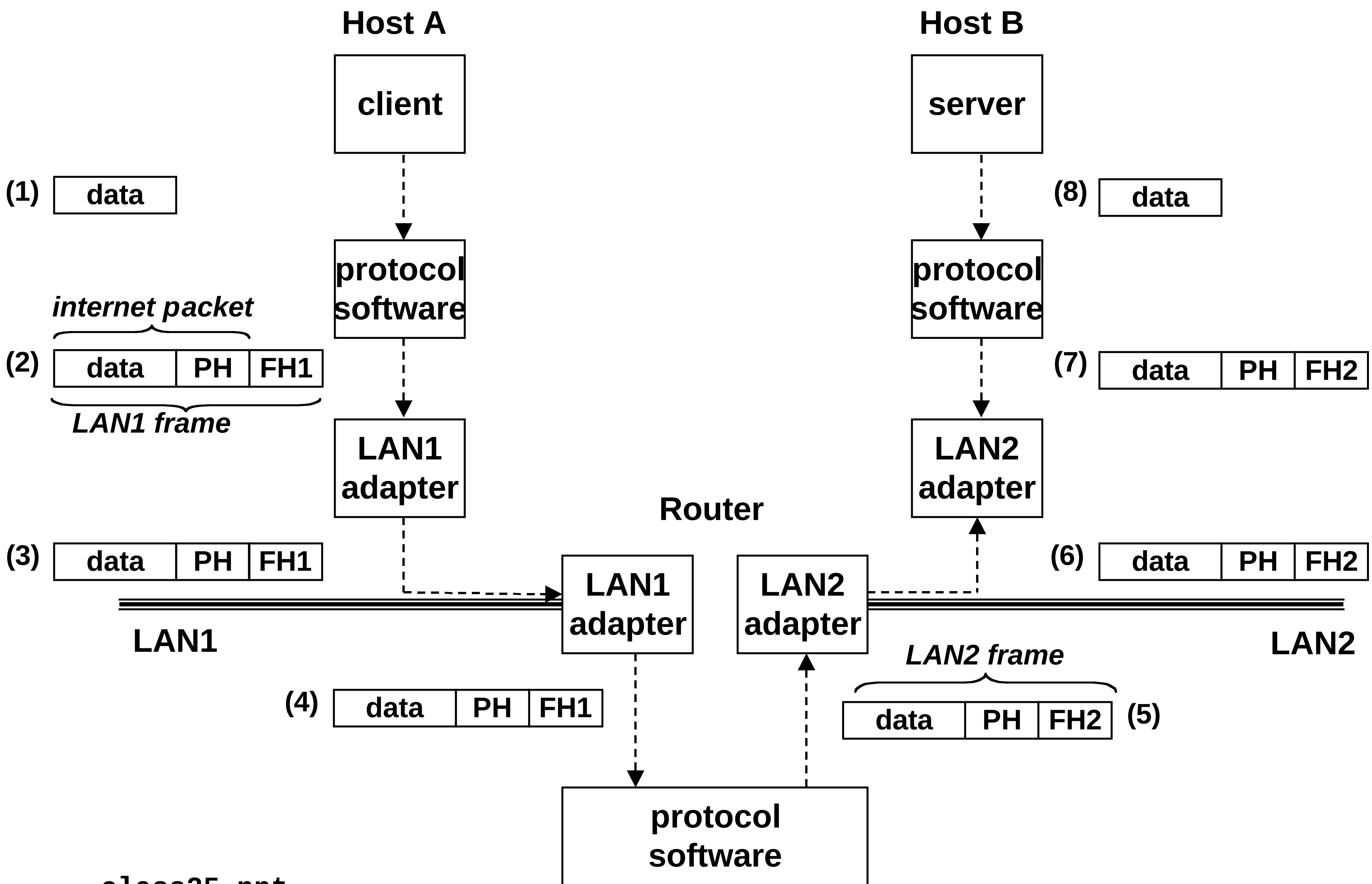
Naming scheme

- **The internet protocol defines a uniform format for host addresses.**
- **Each host (and router) is assigned at least one of these internet addresses that uniquely identifies it.**

Delivery mechanism

- **The internet protocol defines a standard transfer unit (*packet*)**
- **Packet consists of *header* and *payload***
  - **header: contains info such as packet size, source and destination addresses.**
  - **payload: contains data bits sent from source host.**

Transferring data over an internet



Other issues

We are glossing over a number of important questions:

- **What if different networks have different maximum frame sizes? (segmentation)**
- **How do routers know where to forward frames?**
- **How are routers informed when the network topology changes?**
- **What if packets get lost?**

These questions form the heart of the area of computer systems known as networking.

Global IP Internet

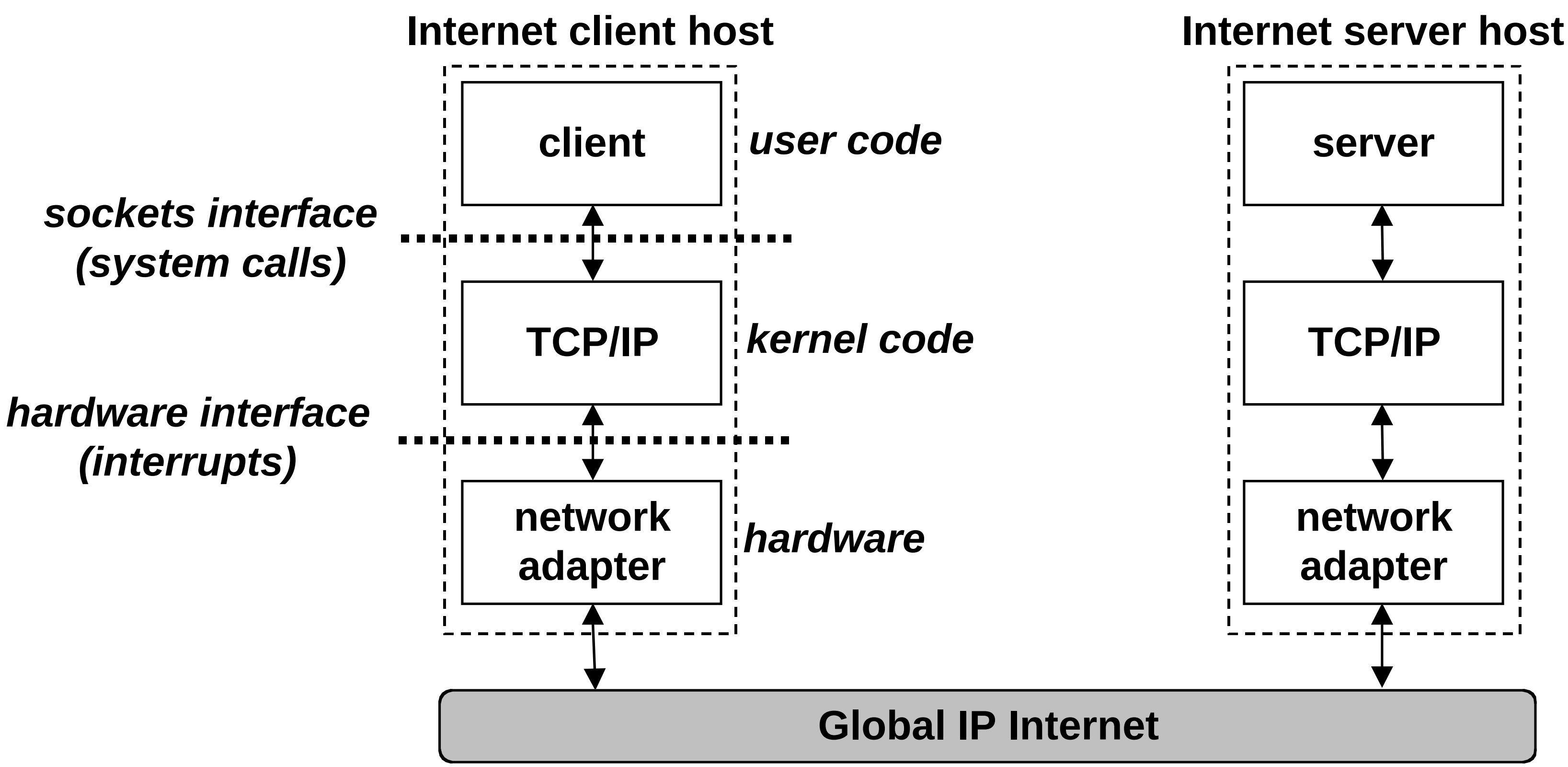
Most famous example of an internet.

Based on the TCP/IP protocol family.

- **IP (Internet protocol) :**
  - **provides basic naming scheme and unreliable delivery capability of packets (datagrams) from host-to-host.**
- **UDP (Unreliable Datagram Protocol)**
  - **uses IP to provide unreliable datagram delivery from process-to-process.**
- **TCP (Transmission Control Protocol)**
  - **uses IP to provide reliable byte streams (like files) from process-to-process.**

Accessed via a mix of Unix file I/O and functions from the Berkeley sockets interface.

Hardware and software organization of an Internet application



Basic Internet components

An Internet *backbone* is a collection of routers (nationwide or worldwide) connected by high-speed point-to-point networks.

A ***Network Access Point*** ***(NAP)*** is a router that connects multiple backbones (sometimes referred to as *peers*).

Regional networks are smaller backbones that cover smaller geographical areas (e.g., cities or states)

A ***point of presence*** ***(POP)*** is a machine that is connected to the Internet.

Internet Service Providers (ISPs) provide dial-up or direct access to POPs.

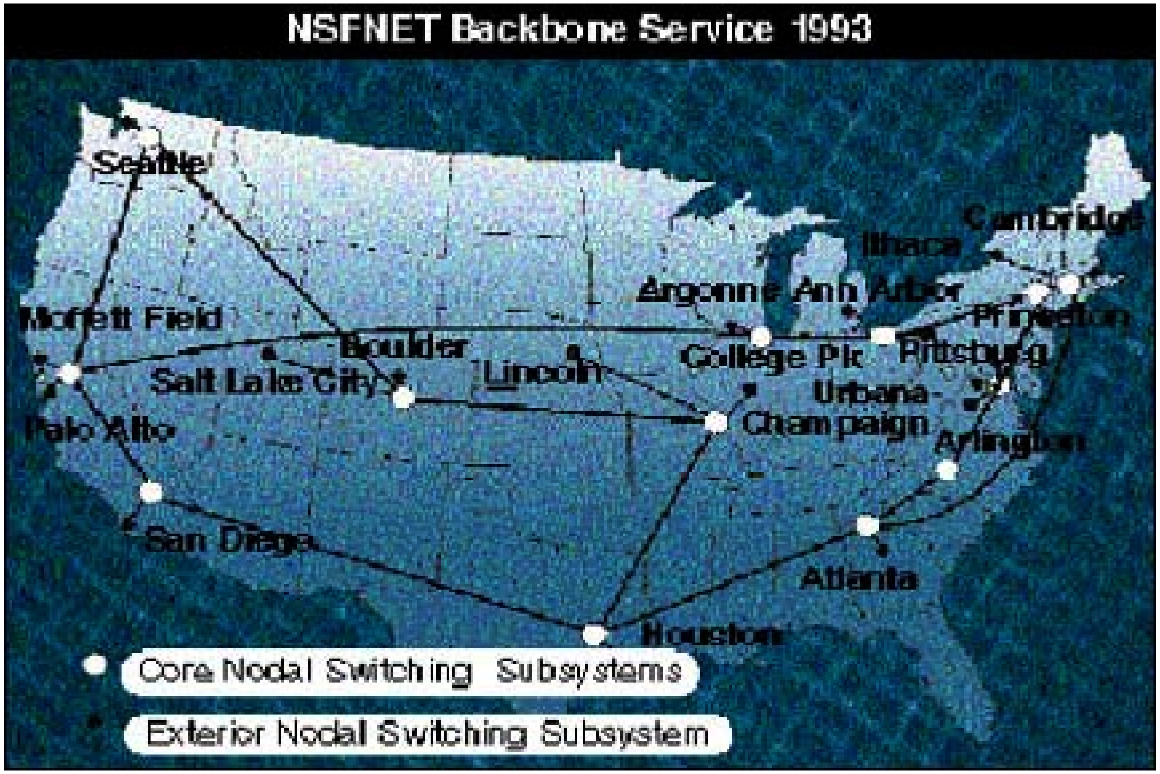
The Internet circa 1993

In 1993, the Internet consisted of one backbone (NSFNET) that connected 13 sites via 45 Mbs T3 links.

- Merit (Univ of Mich), NCSA (Illinois), Cornell Theory Center, Pittsburgh Supercomputing Center, San Diego Supercomputing Center, John von Neumann Center (Princeton), BARRNet (Palo Alto), MidNet (Lincoln, NE), WestNet (Salt Lake City), NorthwestNet (Seattle), SESQUINET (Rice), SURANET (Georgia Tech).

Connecting to the Internet involved connecting one of your routers to a router at a backbone site, or to a regional network that was already connected to the backbone.

The Internet backbone (circa 1993)



Current NAP-based Internet architecture

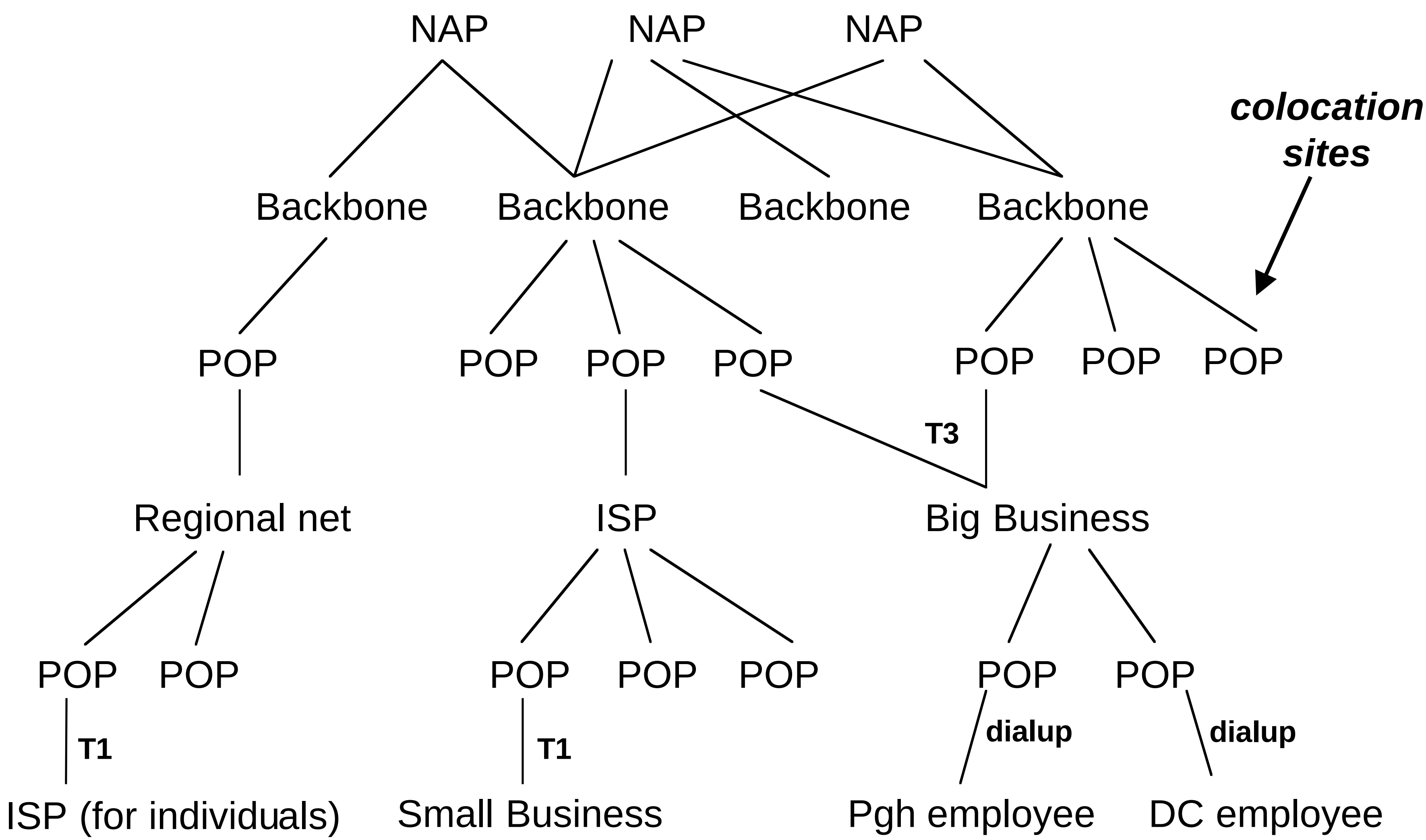
In the early 90's commercial outfits were building their own high-speed backbones, connecting to NSFNET, and selling access to their POPs to companies, ISPs, and individuals.

In 1995, NSF decommissioned NSFNET, and fostered creation of a collection of NAPs to connect the commercial backbones.

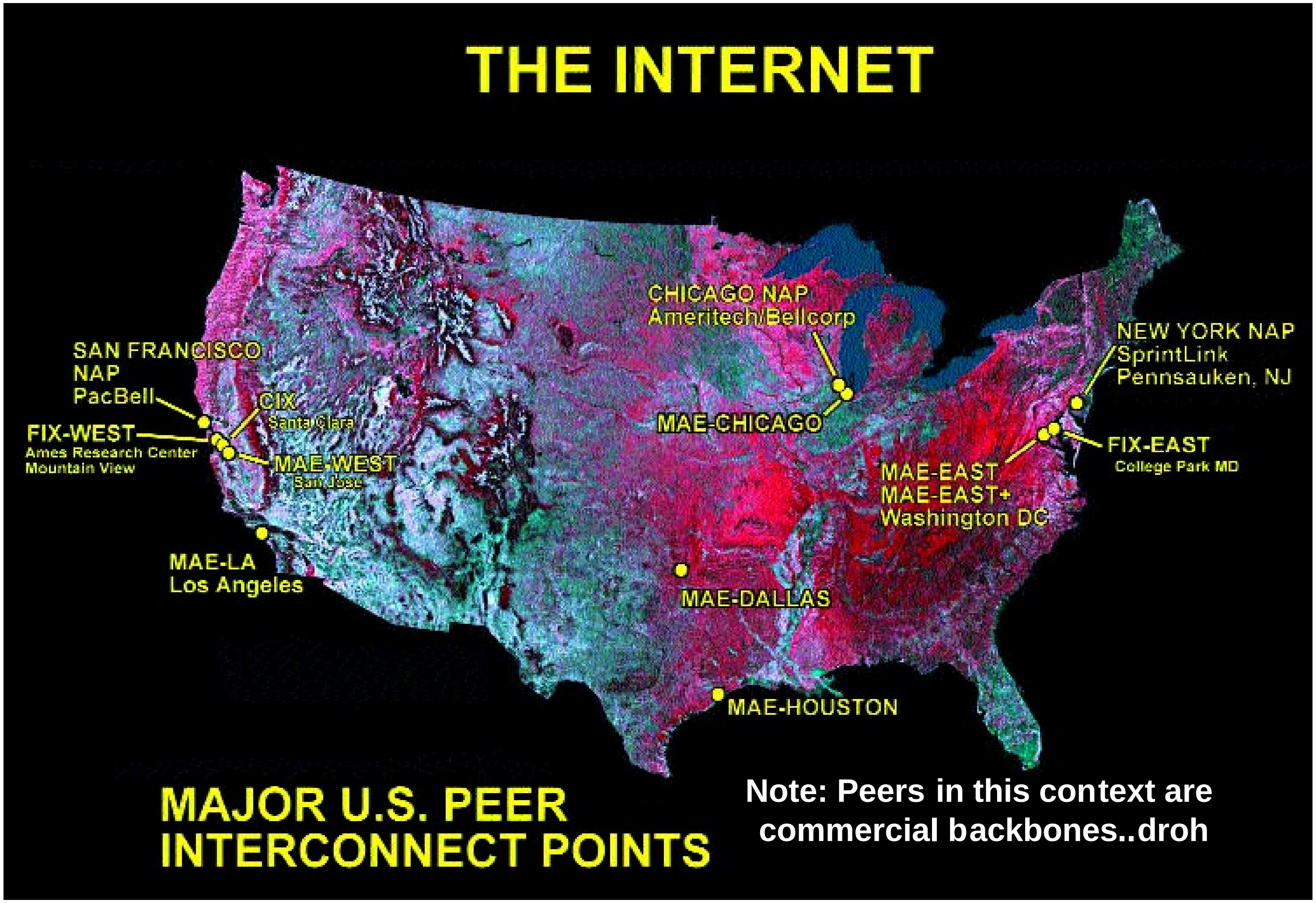
Currently in the US there are about 50 commercial backbones connected by ~12 NAPs (peering points).

Similar architecture worldwide connects national networks to the Internet.

Internet connection hierarchy

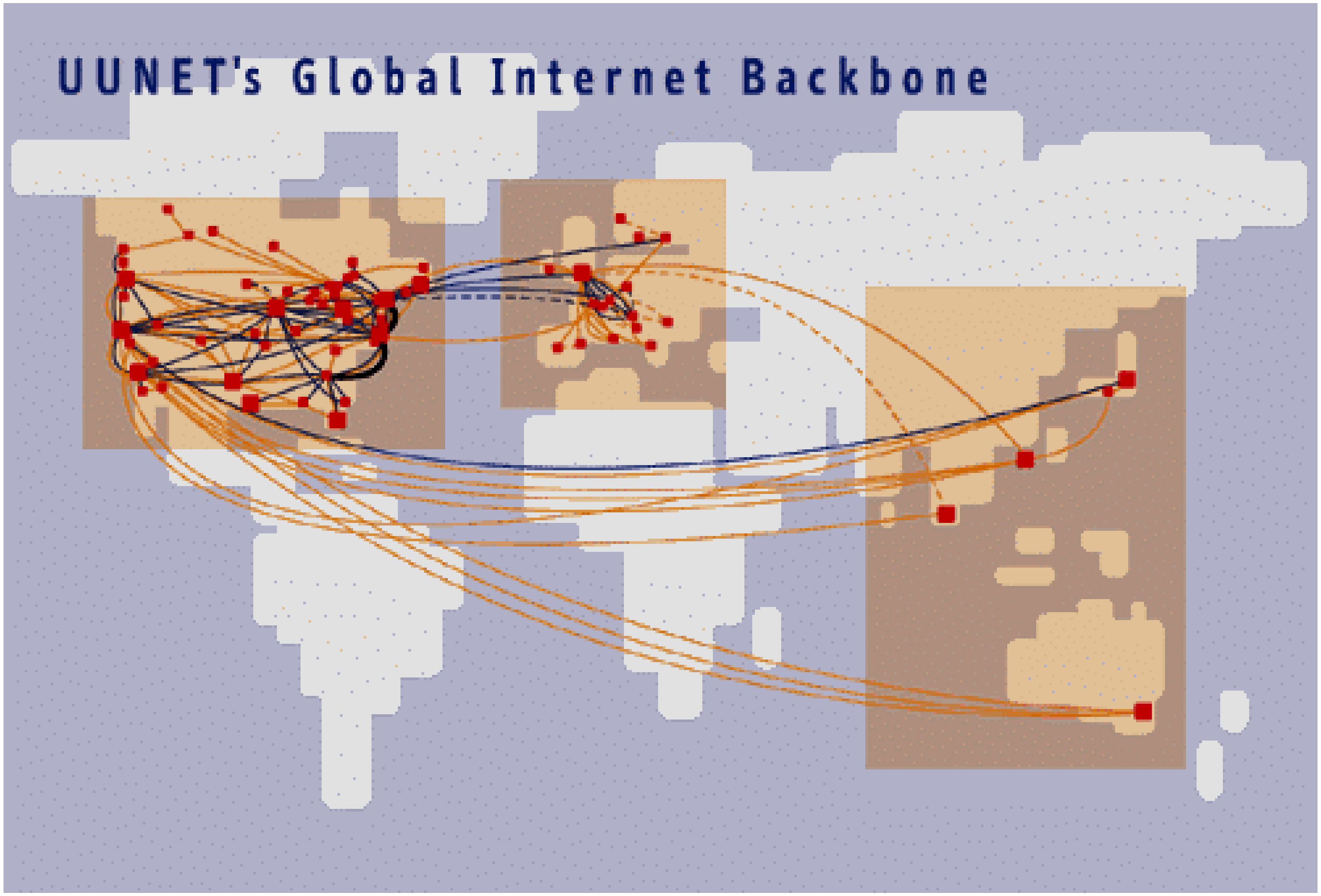


Network access points (NAPs)



Source: Boardwatch.com

MCI/WorldCom/UUNET Global Backbone



Source: Boardwatch.com

Programmer's view of the Internet

Hosts are mapped to a set of 32-bit *IP addresses*.

- **128.2.203.179**

The set of IP addresses is mapped to a set of identifiers called Internet *domain names*.

- **128.2.203.179 is mapped to www.cs.cmu.edu**

A process on one host communicates with a process on another host over a *connection.*

1. IP Addresses

32-bit IP addresses are stored in an *IP address struct*

- IP addresses are always stored in memory in network byte order (big-endian byte order)

```
/* Internet address structure */
struct in_addr {
    unsigned int s_addr; /* network byte order (big-endian) */
};
```

Handy network byte-order functions:

`htonl`: convert `long int` from host to network byte order.
`htons`: convert `short int` from host to network byte order.
`ntohl`: convert `long int` from network to host byte order.
`ntohs`: convert `short int` from network to host byte order.

Dotted decimal notation

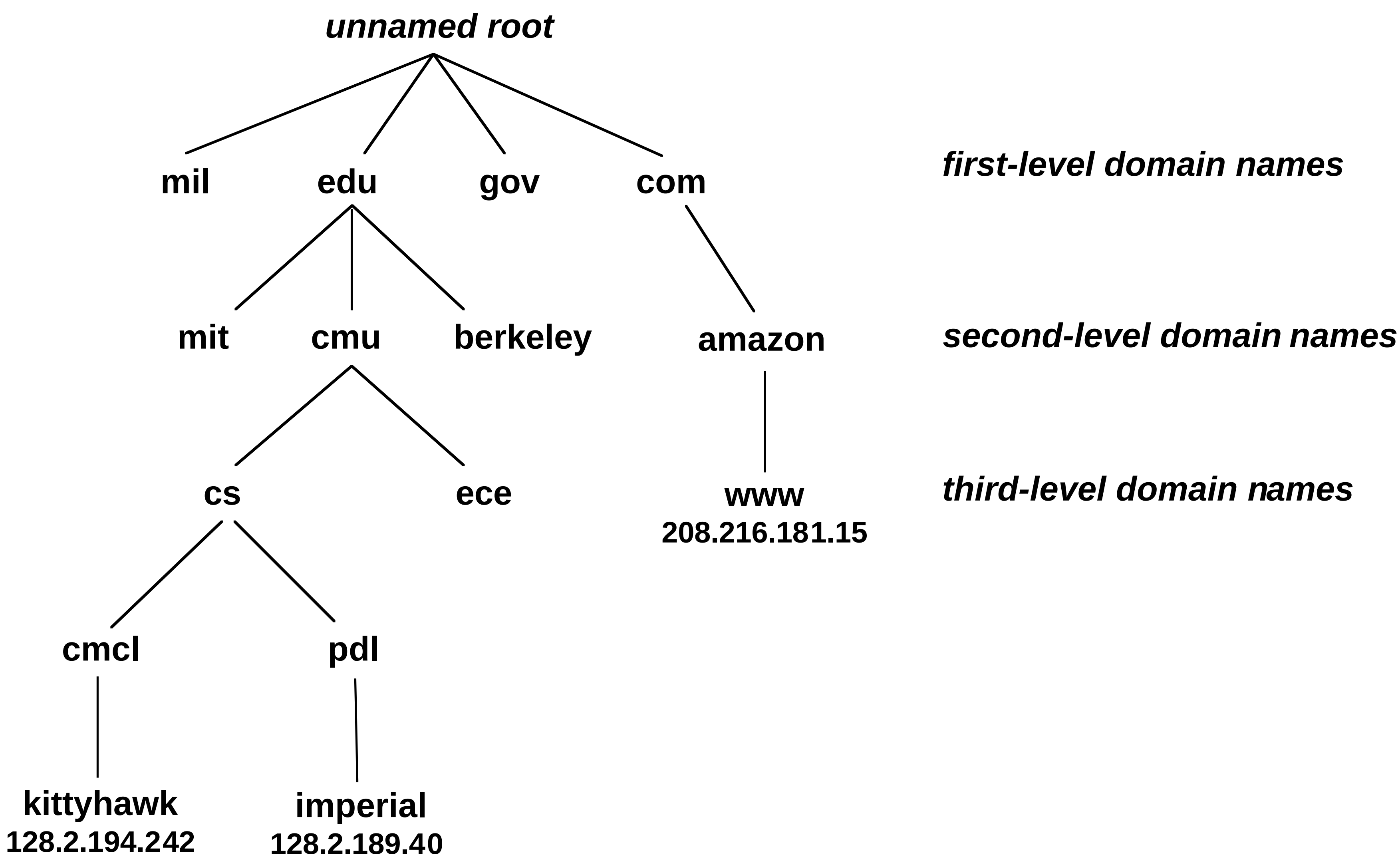
By convention, each by in a 32-bit IP address is represented by its decimal value and separated by a period

– IP address `0x8002C2F2 = 128.2.194.242`

Functions for converting between binary IP addresses and dotted decimal strings:

- `inet_aton`: converts a dotted decimal string to an IP address in network byte order.
- `inet_ntoa`: converts an IP address in network by order to its corresponding dotted decimal string.
- "n" denotes network representation. "a" denotes application representation.

2. Internet Domain Names



Domain Naming System (DNS)

The Internet maintains a mapping between IP addresses and domain names in a distributed database called DNS.

Conceptually, we can think of the DNS database as being millions of *host entry structures*:

```
/* DNS host entry structure */
struct hostent {
   char  *h_name;       /* official domain name of host */
   char  **h_aliases;   /* null-terminated array of domain names */
   int   h_addrtype;    /* host address type (AF_INET) */
   int   h_length;      /* length of an address, in bytes */
   char  **h_addr_list; /* null-terminated array of in_addr structs*/
};
```

Functions for retrieving host entries from DNS:

- `gethostbyname`: query key is a DNS domain name
- `gethostbyaddr`: query key is a an IP address

Properties of DNS host entries

Each host entry is an equivalence class of domain names and IP addresses.

Each host has a locally defined domain name `localhost` which always maps to the *loopback address* `127.0.0.1`

Different kinds of mappings are possible:

- **Simple case: 1-1 mapping between domain name and IP address:**
  - `kittyhawk.cmcl.cs.cmu.edu` maps to `128.2.194.242`
- **Multiple domain names mapped to the same IP address:**
  - `eecs.mit.edu` and `cs.mit.edu` both map to `18.62.1.6`
- **Multiple domain names mapped to multiple IP addresses:**
  - `aol.com` and `www.aol.com` map to three different IP addresses
- **Some valid domain name don't map to any IP address:**
  - for example: cmcl.cs.cmu.edu

hostname: a program that queries DNS

```
int main(int argc, char **argv) { /* argv[1] is a domain name
    char **pp;                      * or dotted decimal IP addr */
    struct in_addr addr;
    struct hostent *hostp;

    if (inet_aton(argv[1], &addr) != 0)
        hostp = Gethostbyaddr((const char *)&addr, sizeof(addr),
                AF_INET);
    else
        hostp = Gethostbyname(argv[1]);
    printf("official hostname: %s\n", hostp->h_name);

    for (pp = hostp->h_aliases; *pp != NULL; pp++)
        printf("alias: %s\n", *pp);

    for (pp = hostp->h_addr_list; *pp != NULL; pp++) {
        addr.s_addr = *((unsigned int *)*pp);
        printf("address: %s\n", inet_ntoa(addr));
    }
}
```

3. Internet connections

Clients and servers communicate by sending streams of bytes of *connections*:

- **point-to-point, full-duplex, and reliable.**

***A socket* is an endpoint of a connection**

- **Socket address is an `IPaddress:port` pair**

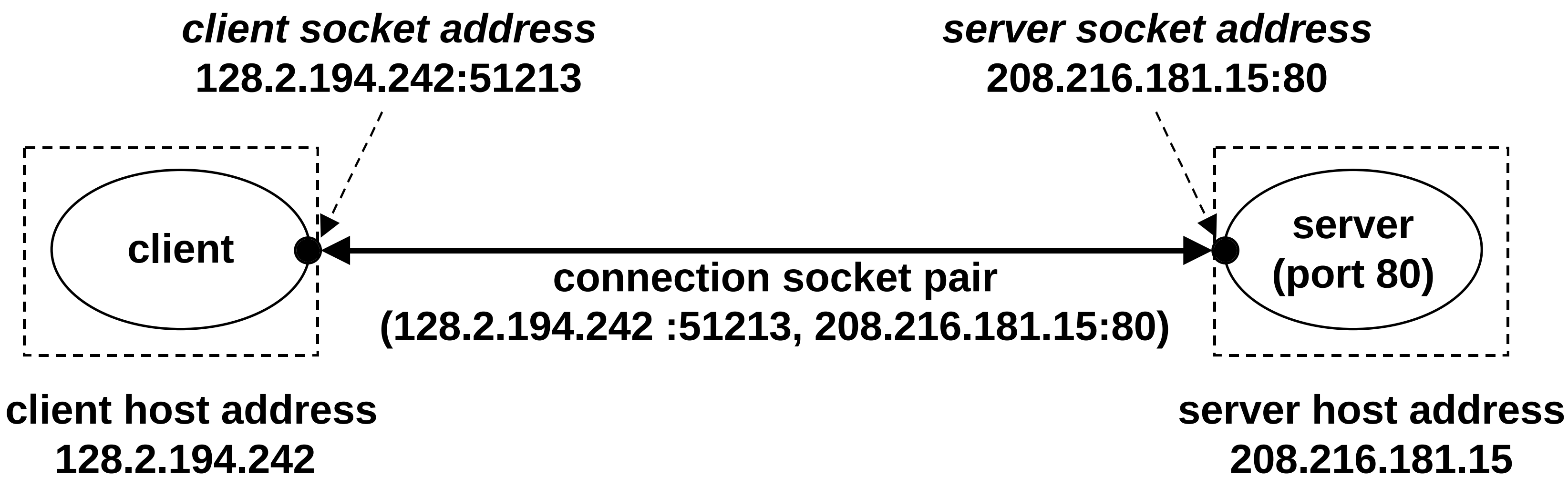
A *port* is a 16-bit integer that identifies a process:

- **ephemeral port: assigned automatically on client when client makes a connection request**
- **well-known port: associated with some service provided by a server (e.g., port 80 is associated with Web servers)**

A connection is uniquely identified by the socket addresses of its endpoints (*socket pair*)

- **`(cliaddr:cliport, servaddr:servport)`**

Putting it all together: Anatomy of an Internet connection



Next time

How to use the sockets interface to program the Internet.