Some Terminology

**internet**: collection of packet switching networks interconnected by routers

**(the) Internet**: “public” interconnection of networks

**end system = host**: computer that is attached to the network ↔ router; usually one network interface

**router = gateway = intermediate system**: routes packets, several interfaces

**subnetwork**: part of an internet (e.g., single Ethernet)

**firewall**: router placed between an organization’s internal internet and a connection to the external Internet, restricting packet flows to provide security.

### Internet WAN Physical Layers

<table>
<thead>
<tr>
<th>Gb/s</th>
<th>remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Giga Ethernet</td>
<td>1.25 fiber</td>
</tr>
<tr>
<td>T-3</td>
<td>0.045 fiber, TP or coax</td>
</tr>
<tr>
<td>OC-3c</td>
<td>0.155 fiber</td>
</tr>
<tr>
<td>OC-12</td>
<td>0.622 fiber</td>
</tr>
<tr>
<td>OC-48</td>
<td>2.4 fiber</td>
</tr>
<tr>
<td>OC-192</td>
<td>10 fiber</td>
</tr>
</tbody>
</table>

**Dense Wavelength Division Multiplexing**

- multiple optical $\lambda$ in single fiber
- 1.6 to 2 Tb/s per fiber
- interfaces typically 622 Mb/s to 10 Gb/s
Link-Layer Mechanisms Used

Roughly in order of popularity:

- ATM
- IP over SONET (synchronous optical network)
- frame relay
- gigabit Ethernet (with range extenders)
- T1, T3

Asynchronous Transfer Mode (ATM)

- 48-byte cells plus 5-byte header
- routing by label swapping
- virtual circuits (VCs) and paths (VPs)
- in-order delivery, but cells can be lost
- adaptation layers:
  
  AAL1    continuous bit rate (CBR); “circuit emulation”
  AAL2    multiplexed low-delay voice
  AAL3/4  data (rarely used)
  AAL5    IP packet in several cells

Frame Relay

- variable-length packets
- permanent or switched virtual circuits (PVC, SVC)
- typically, lower bandwidth (≤ 45 Mb/s)
- popular as access mechanism, corporate networks

Internet Link Layers
Wireless Access

- Industrial, Scientific, Medical (ISM) bands (unlicensed): 902–928 MHz (US only), 2.4 GHz, 5.8 GHz
- analog cellular: 800 MHz
- PCS: 1.9 GHz

Wireless Ethernet:
- 900 MHz, 2.4 GHz, or 5 GHz
- 1 or 2 Mb/s, soon 5.5 Mb/s or 11 Mb/s
- collision-based, with reservation (RTS/CTS)
- IEEE 802.11 = FH or DS

Cellular Digital Packet Data (CDPD):
- pauses in AMPS voice traffic

Wireless access

<table>
<thead>
<tr>
<th>Technology</th>
<th>band</th>
<th>mod.</th>
<th>rate</th>
<th>open range (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAM</td>
<td></td>
<td></td>
<td>8.0 k/b</td>
<td></td>
</tr>
<tr>
<td>GSM data</td>
<td>1.9 GHz</td>
<td>TDMA</td>
<td>9.6 kb/s</td>
<td></td>
</tr>
<tr>
<td>CDPD</td>
<td></td>
<td></td>
<td>19.2 kb/s</td>
<td></td>
</tr>
<tr>
<td>Metricom Ricochet</td>
<td>902-928 MHz</td>
<td>FH</td>
<td>28.8 kb/s</td>
<td>300-450</td>
</tr>
<tr>
<td>Bluetooth</td>
<td>2.4 GHz</td>
<td>FH</td>
<td>432 kb/s</td>
<td>10</td>
</tr>
<tr>
<td>802.11</td>
<td>2.4 GHz</td>
<td>DS</td>
<td>1 Mb/s</td>
<td>540</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Mb/s</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Mb/s</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.5 Mb/s</td>
<td>120</td>
</tr>
</tbody>
</table>

Internet Traffic

- 5,000-8,000 TB/month or 15.4–24.7 Gb/s
- long-distance calls: 525 GDEM or 64 Gb/s
- all the world’s telephones: 600 Gb/s
- almost all (90%?) of the traffic is TCP

Voice vs. Data Traffic
Voice vs. Data Traffic

- local vs. LANs vs. private networks
- capacity vs. traffic
- hop length of data traffic < voice
- link utilization (higher for voice)
- revenue

Internet Names and Addresses

<table>
<thead>
<tr>
<th>proto</th>
<th>src</th>
<th>dest</th>
<th>pkts</th>
<th>bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP</td>
<td>http</td>
<td></td>
<td>35%</td>
<td>66.4%</td>
</tr>
<tr>
<td>TCP</td>
<td>http</td>
<td></td>
<td>33%</td>
<td>7%</td>
</tr>
<tr>
<td>TCP</td>
<td>nntp</td>
<td></td>
<td>1.8%</td>
<td>3.8%</td>
</tr>
<tr>
<td>TCP</td>
<td>ftp</td>
<td></td>
<td>1.4%</td>
<td>3.2%</td>
</tr>
<tr>
<td>TCP</td>
<td>smtp</td>
<td></td>
<td>1.8%</td>
<td>1.9%</td>
</tr>
<tr>
<td>TCP</td>
<td>nntp</td>
<td></td>
<td>1.3%</td>
<td>1.5%</td>
</tr>
<tr>
<td>UDP</td>
<td>dns</td>
<td>dns</td>
<td>3.1%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

April 1997, NLANR
Names, addresses, routes

Shoch (1979):

**Name** identifies what you want,

**Address** identifies where it is,

**Route** identifies a way to get there.

Saltzer (1982):

**Service and users:** time of day, routing, . . .

**Nodes:** end systems and routers

**Network attachment point:**  ≥ 1 per node ⇒ multihomed host vs. router

**Paths:** traversal of nodes and links

binding = (temporary) equivalence of two names

Internet names and addresses

<table>
<thead>
<tr>
<th>example</th>
<th>organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC address</td>
<td>flat, permanent</td>
</tr>
<tr>
<td>IP address</td>
<td>topological (mostly)</td>
</tr>
<tr>
<td>Host name</td>
<td><a href="http://www.ietf.org">www.ietf.org</a></td>
</tr>
</tbody>
</table>

host name \(\rightarrow\) many-to-many

IP address \(\rightarrow\) \(1\)-to-1 MAC address

Mappings in the Internet

<table>
<thead>
<tr>
<th>whois</th>
<th>domain name</th>
<th>owner description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDAP</td>
<td>key (name)</td>
<td>address, other info</td>
</tr>
<tr>
<td>YP</td>
<td>name</td>
<td>data item</td>
</tr>
<tr>
<td>DNS</td>
<td>host name</td>
<td>IP addresses</td>
</tr>
<tr>
<td></td>
<td>IP address</td>
<td>host name</td>
</tr>
<tr>
<td>atmarp</td>
<td>IP address</td>
<td>ATM NSAP</td>
</tr>
<tr>
<td>ARP</td>
<td>IP address</td>
<td>Ethernet address</td>
</tr>
<tr>
<td>RARP</td>
<td>MAC address</td>
<td>IP address</td>
</tr>
</tbody>
</table>

The Internet Domain Name System

We’ll talk about *name resolution* later . . .

organization administering host

organization administering subnames to left

organization type or country

lupus.fokus.gmd.de

Anywhere from two to \(\infty\) parts
Internet (IP) Addresses

Each Internet host has one or more globally unique 32-bit IP addresses, traditionally consisting of a network number and a host number:

- Originally, two-level hierarchy → \( n \)-level, changing
- An IP address identifies an *interface*, not a host!
- A host may have two or more addresses. Why?

<table>
<thead>
<tr>
<th>network</th>
<th>host</th>
<th>Class A</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Class B</td>
</tr>
<tr>
<td>110</td>
<td></td>
<td>Class C</td>
</tr>
<tr>
<td>1110</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>network</th>
<th>host</th>
<th>Class D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1110</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Internet addresses

- (almost) every *interface* has one
- But may
  - Change (dial-in)
  - Have lots (WWW servers)
  - Have none (some routers)
  - Not be globally unique
- Old: class-{A,B,C} → 2-level addressing: network, host
- New: classless interdomain routing (CIDR) → aggregation, route on prefix and mask

IP addresses

- Dotted decimal notation: 4 decimal integers, each specifying one byte of IP address:
  - Host name: lupus.fokus.gmd.de
  - 32-bit address: 1100 0000 0010 0011 1001 0101 0011 0100
  - Dotted decimal: 192.35.149.52
- Loopback: 127.0.0.1 (packets never appear on network)
- Own network (broadcast): hostid = 0; own host: netid = 0
- Directed broadcast: hostid = all ones
- Local broadcast: 255.255.255.255

CIDR: Classless Interdomain Routing

- Problem: too many networks → routing table explosion
- Problem: class C too small, class B too big (and scarce)
- Discard class boundaries → supernetting
- ISP assigns a contiguous group of \( 2^n \) class C blocks
- "Longest match routing" on masked address; e.g. 192.175.132.0/22

<table>
<thead>
<tr>
<th>address/mask</th>
<th>next hop</th>
</tr>
</thead>
<tbody>
<tr>
<td>192.175.132.0/22</td>
<td>1</td>
</tr>
<tr>
<td>192.175.133.0/23</td>
<td>2</td>
</tr>
<tr>
<td>192.175.128.0/17</td>
<td>3</td>
</tr>
</tbody>
</table>

- E.g.: All sites in Europe common prefix → only single entry in most U.S. routers
Example: ifconfig

```
ifconfig -a
le0:  flags=863<UP,BROADCAST,NOTRAILERS,RUNNING>
     inet 192.35.149.117 netmask ffffff00
     broadcast 192.35.149.0
fa0:  flags=863<UP,BROADCAST,NOTRAILERS,RUNNING>
     inet 194.94.246.72 netmask ffffff00
     broadcast 194.94.246.0
qaa0: flags=61<UP,NOTRAILERS,RUNNING>
     inet 193.175.134.117 netmask ffffff00
qaa1: flags=61<UP,NOTRAILERS,RUNNING>
     inet 129.26.216.231 netmask fffff0000
qaa2: flags=60<NOTRAILERS,RUNNING>
qaa3: flags=60<NOTRAILERS,RUNNING>
lo0:  flags=849<UP,LOOPBACK,RUNNING>
     inet 127.0.0.1 netmask ff000000
```

IP address exhaustion

As of February 2000,

- 61.1% of available address space allocated
- 49.4% of allocated address space announced
- 30.2% of available address space announced

Routing table:

- 71,717 “autonomous system” (AS) entries
- 41,256 of which are /24
Network Address (and Port) Translation (NA(P)T)

- most corporations use private address space, also residential
  - 10/8, 172.16/12, 192.168/16
- NAT translates internal $\rightarrow$ external as needed
- works for outgoing TCP connections: POP, HTTP, SMTP, Telnet
- need application layer gateway (ALG) for out-of-band protocols (ftp, SIP, RTSP, H.323, ...)
- problems:
  - controlled connections (ftp, Internet telephony, media-on-demand)
  - UDP services (streaming media)
  - security – rewriting breaks IPsec
- suggestion: Realm-Specific IP (RSIP) makes host aware of mapping

Problems with IP Addresses

- if a host moves from one network to another, its IP address changes
- currently, mostly assigned without regards to topology $\rightarrow$ too many networks $\Rightarrow$ CIDR
- limited space $\Rightarrow$ IPv6
- class thresholds: class C net grows beyond 254 hosts
- hard to change: hidden in lots of places
- multihomed host: path taken to host depends on destination address

Multihoming

- = one “stub” network, multiple providers
- options:
  1. global prefix $\Rightarrow$ aggregation $\downarrow$
  2. divide network $\Rightarrow$ no redundancy
  3. multiple addresses $\Rightarrow$ applications need to try several, address space use $\uparrow$

Mobility and Renumbering

- renumber if immediate or up-stream provider changes
- mobility: change network attachment point
- mobility = renumbering: network “location” changes
- IP address as location $\Rightarrow$ keep address, break aggregation
- renumbering is hard: configuration files, transition
- IP address as identifier $\Rightarrow$ break connections
### Subnetting

- large organizations: multiple LANs with single IP network address
- subdivide “host” part of network address ➜ subnetting

```
<table>
<thead>
<tr>
<th>Network 150.17.1.0 (mask ffffff00)</th>
</tr>
</thead>
<tbody>
<tr>
<td>150.17.1.1</td>
</tr>
<tr>
<td>150.17.1.2</td>
</tr>
<tr>
<td>150.17.1.3</td>
</tr>
<tr>
<td>150.17.2.1</td>
</tr>
<tr>
<td>150.17.2.2</td>
</tr>
<tr>
<td>150.17.2.3</td>
</tr>
</tbody>
</table>
```

netmask: 0xFF FF 00 00

- <256 subnets
- 254 nodes

### How does a packet get to the server?

E.g., web page from http://www.cs.umass.edu:

- get host name www.columbia.edu from URL;
- DNS: translate to IP address 128.59.35.60
- is it on local network? no ➜ find local router
- local router sends to Internet
- Internet routes to Columbia network router (128.59.?.?)
- Columbia router routes to web server

### Peeking inside a packet

- IP (network)
- TCP (transport)
- http (application)

#### IP Forwarding

```
get destination IP address D
if network(D) == directly attached network {
    ARP: D -> MAC address
    put in link layer frame
    forward
} else {
    foreach entry in routing table {
        if (D & subnet mask) == network(entry) {
            get next hop address N
            ARP: N -> MAC address
            put in link layer frame
            forward
        }
    }
}

⇒ IP source/destination remains same, MAC changes
```
IP Forwarding

GMD Fokus Network
Network Layer: IPv4 and IPv6

- unreliable datagram ➔ misorder, loose, duplicate
- 32-bit (IPv6: 128 bit) globally unique addresses
- no checksum on payload
- allow fragmentation of large packets into MTU-sized frames
- 20 (IPv6: 40) byte header
- IP multicast: receiver group with anonymous membership

IPv4 Service Model

datagram: each packet is independent of all others
best effort: packet may arrive or not after some time

- independent packets
- unreliable
- might be reordered (rare), delayed, duplicated, . . .
- but: minimal service on top of anything (see RFC 1149)
- only header checksum
IPv4 Header

RFC 791

IPv4

version: always 4
TOS (type of service): precedence (3 bits) and “minimize delay”, “maximize throughput”, “maximize reliability”, “minimize cost” bits ➥ rarely used
identifier: identifier, different for each packet from host
TTL: time to live field; initialized to 64; decremented at each router ➥ drop if TTL = 0 (prevent loops!)
protocol: next higher protocol (TCP: 6, UDP: 17)
header checksum: add together 16-bit words using one’s complement ➥ optimized for software

IP Fragmentation and Reassembly

data link protocol may limit packets < 65,536 bytes ➥ transport layer packet may be too big to send in single IP packet
**IP Fragmentation and Reassembly**

➤ split TPDU into fragments

- each fragment becomes its own IP packet (routers don’t care)
- each fragment has same identifier, source, destination address
- fragment offset field gives offset of data from start of original packet
- *more fragments (MF)* flag of 0 if last (or only) fragment of packet
- fragments reassembled only at final destination
- routers must handle at least 576 bytes
- *do not fragment* bit prevents fragmentation ➤ drop + error message
- avoid multiple fragmentation (1500 → 620) ➤ MTU discovery

**IP Options**

Extend functionality of IP without carrying useless information:

- security and handling restrictions for military
- determine route (source route)
- record route
- record route and timestamps

(rarely used ↔ not all routers support them)

**IP Record Route Option**

- source creates empty list of ≤ 9 IP addresses
- option: length, pointer, list of IP addresses
- routers note outgoing interface in list
- . . . and bump pointer

**IP Source Route Option**

- source determines path taken by packet (≤ 9 hops)
- *loose*: any number of hops in between
- *strict*: every hop; if not directly connected, discard
- same format as record route option
- router overwrites with address of outgoing interface
- must be copied to fragments
- destination should reverse route for return packets
- not too popular ➤ router performance ↓
ICMP

- used to communicate network-level error conditions and info to IP/TCP/UDP entities or user processes
- often considered part of the IP layer, but
  - IP demultiplexes up to ICMP using IP protocol field
  - ICMP messages sent within IP datagram
- ICMP contents always contain IP header and first 8 bytes of IP contents that caused ICMP error message to be generated

<table>
<thead>
<tr>
<th>20-byte standard IP header</th>
<th>8 bit ICMP type</th>
<th>8 bit ICMP code</th>
<th>16-bit checksum</th>
<th>contents of ICMP msg</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=0 time=276 ms</td>
<td>0</td>
<td>0</td>
<td>209</td>
<td>echo reply (to a ping)</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=1 time=281 ms</td>
<td>3</td>
<td>0</td>
<td>209</td>
<td>destination network unreachable</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=2 time=276 ms</td>
<td>3</td>
<td>1</td>
<td>209</td>
<td>destination host unreachable</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=3 time=281 ms</td>
<td>3</td>
<td>2</td>
<td>209</td>
<td>destination protocol unreachable</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=4 time=276 ms</td>
<td>3</td>
<td>3</td>
<td>209</td>
<td>destination port unreachable</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=5 time=281 ms</td>
<td>3</td>
<td>4</td>
<td>209</td>
<td>fragmentation needed and DF set</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=6 time=276 ms</td>
<td>3</td>
<td>6</td>
<td>209</td>
<td>destination network unknown</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=7 time=281 ms</td>
<td>3</td>
<td>7</td>
<td>209</td>
<td>destination host unknown</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=8 time=276 ms</td>
<td>3</td>
<td>8</td>
<td>209</td>
<td>other reasons</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=9 time=281 ms</td>
<td>4</td>
<td>0</td>
<td>209</td>
<td>source quench (slow down)</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=10 time=276 ms</td>
<td>5</td>
<td>1</td>
<td>209</td>
<td>redirect message to host</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=11 time=281 ms</td>
<td>8</td>
<td>0</td>
<td>209</td>
<td>echo request (ping)</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=12 time=276 ms</td>
<td>9</td>
<td>0</td>
<td>209</td>
<td>IS-IS router advertisement (new)</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=13 time=281 ms</td>
<td>10</td>
<td>0</td>
<td>209</td>
<td>ES-IS router discovery (new)</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=14 time=276 ms</td>
<td>11</td>
<td>0</td>
<td>209</td>
<td>time exceeded = TTL zero</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=15 time=281 ms</td>
<td>12</td>
<td>0</td>
<td>209</td>
<td>IP header bad</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=16 time=276 ms</td>
<td>17</td>
<td>0</td>
<td>209</td>
<td>address (subnet) mask request</td>
</tr>
<tr>
<td>4 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=17 time=281 ms</td>
<td>18</td>
<td>0</td>
<td>209</td>
<td>address (subnet) mask reply</td>
</tr>
</tbody>
</table>

ping

- checks if host is reachable, alive
- uses ICMP echo request/reply
- copy packet data request — reply

traceroute

- allows to follow path taken by packet
- send UDP to unlikely port; ‘time exceeded’ and ‘port unreachable’ ICMP replies
- can use source route (-g), but often doesn’t work
**ARP: IP address → MAC address**

- for broadcast networks like Ethernet, token ring, ...
- if MAC address unknown, send ARP request and hold on to packet
- ARP request → broadcast: sender IP, MAC; target IP, MAC
- *all* machines update their cache ⇒ efficiency, allow change of interface
- ARP reply → requestor: reverse source/target; fill in source MAC
- directly on Ethernet, *not* IP!
- cache ARP replies; drop after 20 minutes

---

**ARP example**

```
arp -a
```

<table>
<thead>
<tr>
<th>Device</th>
<th>IP Address</th>
<th>Mask</th>
<th>Flags</th>
<th>Phys Addr</th>
</tr>
</thead>
<tbody>
<tr>
<td>le0</td>
<td>hamlet 255.255.255.255</td>
<td>08:00:09:70:7d:16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>le0</td>
<td>gaia 255.255.255.255</td>
<td>08:00:20:20:07:03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>le0</td>
<td>pern 255.255.255.255</td>
<td>08:00:20:75:3c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>le0</td>
<td>kite 255.255.255.255</td>
<td>08:00:09:92:04:01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>le0</td>
<td>condor 255.255.255.255</td>
<td>08:00:20:1c:95:ed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**RARP: MAC → IP address**

- determine IP address at boot for diskless workstations
- remember: MAC address is unique and permanent
- host broadcasts RARP request (with its own MAC address)
- RARP server responds with reply
- allows third-party queries
- want several servers for reliability

---

**Proxy ARP**

- extend network: router fronts for H3, H4
- router answers ARP requests for H3, H4 from H1, H2 with its *own* hardware address
- assumes trusting relationship
- only needs to be added to single router
- only works for broadcast networks
Transport Layer: UDP and TCP

- UDP service = IP service + checksum + ports
- TCP service = UDP service + flow control + congestion control + sequenced, reliable byte stream
- TCP for multimedia:
  - loss recovery delay (RTT + $\epsilon$)
  - windowed flow/congestion control ➞ variable bandwidth
  - no multicast

Internet Domain Names

The Internet Domain Name System (DNS)

- hierarchical, dot-separated names
- ➞ multi-level delegation
- by country and by type of organization
- needs to be overhauled (59% of all domains = .com!)

Global top-level domains (gTLDs):

- 2 letters: countries
- 3 letters: independent of geography (except edu, gov, mil)
Example

server 128.9.0.107
Default Server: b.root-servers.net
Address: 128.9.0.107

> erlang.cs.columbia.edu
Server: b.root-servers.net
Address: 128.9.0.107

Name: erlang.cs.columbia.edu
Served by:
- CUNIXD.CC.COLUMBIA.edu
  128.59.35.142
  COLUMBIA.edu
- DNS2.ITD.UMICH.edu
  141.211.125.17
  COLUMBIA.edu

Domain Name Resolution

- hierarchy of redundant servers with time-limited cache
- each server knows the 13 root servers a.root-servers.net
- each root server knows gTLDs and refers queries to those
- each domain has ≥ 2 servers, often widely distributed
- also: mailbox translation
- almost a distributed database