CSci 4271W
Development of Secure Software Systems
Day 13: Networking overview

Stephen McCamant (he/him)
University of Minnesota, Computer Science & Engineering

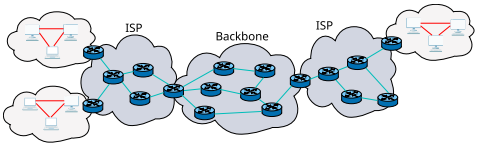
Networks and software

What happens as a result of the following program?

```
import urllib.request
with urllib.request.urlopen("http://neverssl.com/") as f:
  for line in f: print(line)
```

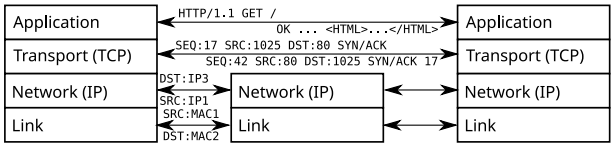
The application request is encoded into a transport connection divided over a sequence of datagrams that are sent from host to host in a series of frames...

The Internet



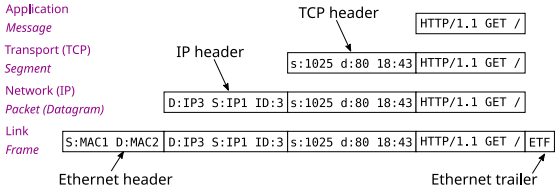
- The internet consists of end system networks connected to other networks by ISPs or autonomous systems (ASes).
- AS routing protocols determine how to reach each network
- Packets of data are forwarded between hosts via TCP/IP
- Domain names are translated to IP addresses via DNS

Internet protocol stack



- Urllib makes an application protocol (HTTP) request to `neverssl`'s web server application.
- The request is sent over a transport protocol (TCP) session
- The session consists of IP datagrams sent over the Internet
- The datagrams are sent in frames between forwarding hosts

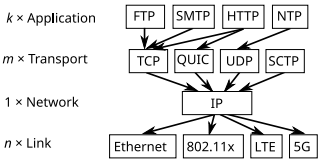
Packet formats



Data at each level are encapsulated in the appropriate units of the lower-level protocol

Internet Protocol (IP)

IP is the "narrow waist" of the TCP/IP network stack:



connectionless, unreliable, datagram delivery

IP headers

An IPv4 datagram is a header followed by data:

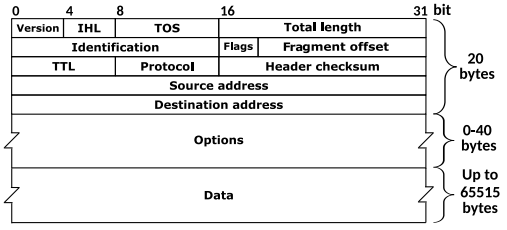
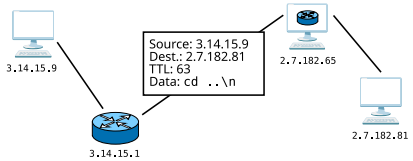


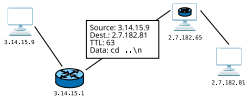
Figure: Michel Bakni via Wikimedia Commons, CC BY-SA, after RFC 791

IP forwarding



- Datagrams are forwarded based on the 4-byte IP address (16 bytes in IPv6) of the destination host.
- Typical routes are ~10-20 hops across 4-5 ISPs.

IP forwarding (cont'd)



- Possible because (a) the source knows a router, and (b) each router knows a next hop towards the destination.
- Packets too large for the link layer at some hop can be fragmented and then are reassembled at the destination.
- ICMP used for errors: TTL exceeded, host unreachable

Outline

Internet and IP

Announcements intermission

More Internet protocols

Midterm debrief, cont'd

Upcoming assignments

- Homework 3 late submissions still open
- Section drafts for project 1 due tonight on Gradescope
- Due date for full project 1 is Tuesday, March 18th
  - One-time extension to Friday is available, but must be requested by Monday the 17th.

Outline

Internet and IP

Announcements intermission

More Internet protocols

Midterm debrief, cont'd

User Datagram Protocol (UDP)

- UDP is a transport protocol that provides application (de)multiplexing via 16-bit port numbers.
- Each application sends and accepts datagrams from its own port number, which can be requested or assigned.
- The destination port specifies the remote application, and the source port provides a return address.
- UDP is stateless and unreliable: data may be delivered, or not, in any order, without notification to the sender.

Transmission Control Protocol (TCP)

- TCP is a reliable transport protocol. A TCP connection is defined by the 4-tuple (src:port, dst:port).
- TCP senders break up a message into packets ("segments") with sequence numbers.
- TCP receivers ACKnowledge receipt of packets in order.
- Dropped packets aren't ACKed, causing timeouts. The sender re-transmits NACKed packets after the timeout.
- Timeouts signal network congestion, which prompts senders to (voluntarily) reduce their sending rate.

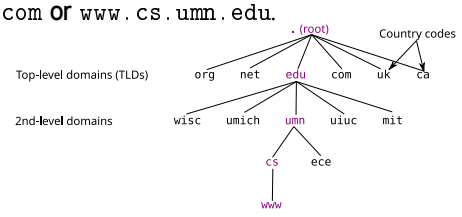
Common application protocols

HTTP	TCP	80	(unencrypted) web
DNS	U/T	53	Domain Name Service
SMTP	TCP	25	email sending
FTP	TCP	21, 20	File Transfer Protocol
SSH	TCP	22	Secure Shell
Telnet	TCP	23	(unencrypted) remote login
NTP	UDP	123	Network Time Protocol
IMAP	TCP	143	Internet Message Access Protocol
HTTPS	TCP	443	(secure) web

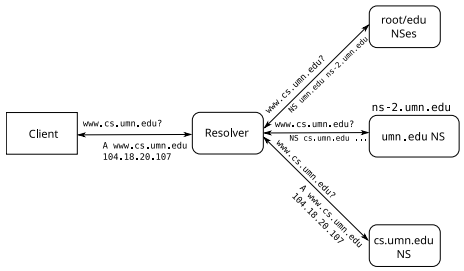
For more complete lists, see `/etc/services` or the IANA

Domain Name Service (DNS)

DNS is the protocol used to map between IP addresses and hierarchical/semantic "domain names" like `xkcd.com` or `www.cs.umn.edu`.



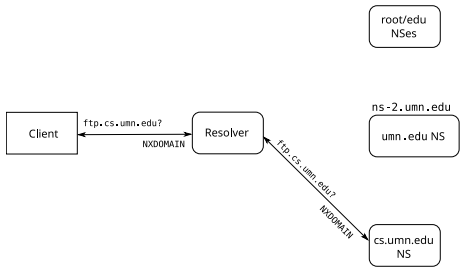
Resolving domain names



DNS caching

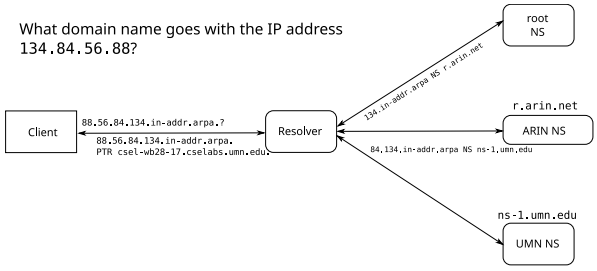
- DNS responses are usually cached, to avoid latency of repeated lookups.
- Recursive resolvers cache intermediate results.
- Negative results are also cached.
- Cached data times out according to a TTL (expiration date) in the record.

Cached DNS lookup



Reverse DNS

What domain name goes with the IP address 134.84.56.88?



Routing protocols

- Local routing: ARP (address resolution protocol) uses flooding to find hosts (get their link-level addresses) on a local network
- Intra-ISP routing: OSPF/IS-IS use link state broadcast and shortest-path trees to find next hops; iBGP uses path vectors
- Inter-ISP routing: BGP uses AS-based path vectors

Outline

Internet and IP

Announcements intermission

More Internet protocols

Midterm debrief, cont'd

Q3: memory corruption

(Code shown outside slides)