485/585 Lecture 1

crandall@cs.unm.edu
Administrativia

- TA: Shekh Ahammed Adnan Bashir (e-mail address will be provided in class)
- Mailing list, links fixed
- Do homework 1, even though it won't be graded
- Temporary CS/B146 accounts
What's to come

- Lectures will be long and cover more than one class period
  - Lots of pauses to try things out on Linux machines
- Lab 1 will require you to know the basics of most of what we'll learn for the rest of the semester
  - Like a survey
- We'll be going back and learning things in more detail, such as routing algorithms and TCP congestion control
Empiricism

- “a theory that all knowledge originates in experience” (Merriam-Webster)
- Francis Bacon and the Royal Society
- Robert Boyle and his vacuum pump
Lab 1 will be assigned soon

- You can start reading and playing around with Wireshark, if you like:
  https://citizenlab.ca/2015/04/chinas-great-cannon/
- You'll be analyzing all 11 packet captures (or, PCAPs) linked to in that report to test the claims of the report
- Most students swim instead of sink ;-)
  - We're going to go over the basics you need to know to complete Lab 1 now...
A network of two machines

10.0.8.1 — 10.0.8.2

IP address: identifies the machine on the network. Local (10.0.0.0/8, 192.168.0.0/16, 172.16.0.0/12) vs. Internet routable (e.g., 64.106.21.143).
man man
man ifconfig
ifconfig | less
ARP: Address Resolution Protocol

- MAC (media access control) address (e.g., “c4:02:32:6b:00:00”) is supposed to be unique to the network interface
  - Also called a hardware address
What is a network interface?
About network interfaces

- All a machine needs to be a router is more than one network interface
  - Still needs a routing policy, though
- Ethernet is a shared medium, can connect more than two machines on the same network via:
  - Hub (truly shared)
  - Switch (Smart about what it broadcasts)
The word “network”

• Can refer to a network in an abstract sense
  – E.g., the Internet is a network, the network of cities connected by the U.S. Interstate Highway system, my social network of friends

• Can refer to a set of machines that use a shared medium (Ethernet) to communicate directly on the same subnet without needing a router.
ARP resolution table from http://chrissanders.org/packet-captures/
• Classless Inter-Domain Routing
• /27 has a net mask of 255.255.255.224
• 10.10.1.32/27 has 32 possible IPs on it
CIDR seems complicated, is simple once you learn the notation

- I'll assign a homework
Network of networks
IP (Internet Protocol) routing

Graphic by Danny Adams
tracepath -n uchicago.edu
route -n
arp -n
Packets

- Almost all packets we'll deal with in this class are IP (Internet Protocol) packets.
- Do not call packets “packages”.
- 1 unit of data to be routed across the network/Internet.
- Important for Lab 1: IP packets have a Time-to-Live (TTL) field.
TCP/IP

• TCP = Transport Control Protocol
• Port: a number that identifies a process or service on the remote machine
• Socket: a way for a process on one machine to communicate with a process on another machine
  - Can be identified by two port:ipaddress tuples
• TCP is connection-oriented, packets can be lost and retransmitted, delivered out of order, etc.
  - Compare to UDP, which is the User Datagram Protocol
• See http.pcap example from https://wiki.wireshark.org/SampleCaptures#HyperText_Transport_Protoocol_.28HTTP.29
screen
Ctrl+A then C to Create
Ctrl+A then N for Next
host www.cs.unm.edu
nc 64.106.20.27 80
(in other terminal...)
netstat -tpn | less
DNS maps hostnames to IPs and vice versa

- host 64.106.20.60
- host wiki.cs.unm.edu
- man dig
Where do these standards come from?

- IETF = Internet Engineering Task Force
- RFC = Request for Comments
  - MUST, MUST NOT, SHOULD, SHOULD NOT, MAY (RFC 2119)
- “The only laws on the Internet are assembly and RFCs” --Phrack 65
  - Assembly is an abstraction
  - RFCs are not always followed
    - Often ambiguous
OSI (Open Systems Interconnection) model

- Layer 1: Physical (think Ethernet, 802.11)
- Layer 2: Data Link (think ARP)
- Layer 3: Network (think IP)
- Layer 4: Transport (think TCP)
- Layer 5: Session (think NetBIOS, SOCKS)
- Layer 6: Presentation (think SSL/TLS)
- Layer 7: Application (think HTTP)
Some Wireshark stuff (http.pcap)

- You should poke around the GUI
- You should also check out tcpdump, tshark, and Python dpkt
- View::Name Resolution
- Right Click->Follow TCP Stream
- Can look in, e.g., IP header for, e.g., TTL (Time-to-live), or, e.g., TCP header for, e.g., ports
- Never completely trust abstractions
  - Can you trust the wireshark GUI?
  - Can you trust tshark raw text output?
  - Can you trust raw bits off the wire?
TCP 3-way handshake

- TCP header has flags
  - SYN is “Synchronize”, it means the sequence number has a special meaning
  - ACK is “Acknowledge”, it means the acknowledgment number has meaning
  - RST: “I have no record of such a connection”
  - Also, FIN, CWR, ECN, URG, PUSH
TCP 3-way handshake

- SYN: I'd like to open a connection with you, here's my initial sequence number (ISN)
- SYN/ACK: Okay, I acknowledge your ISN and here's mine
- I ACK your ISN

Image from Wikipedia
Great Firewall vs. Great Cannon

Reproduced from https://citizenlab.ca/2015/04/chinas-great-cannon/
Proof that the GFW exists?

- Google for “site:.edu.cn” or “site:.cn” to find web servers in China's domain name
- Use the Linux `host` command to get an IP address
- Use ip2location.com to confirm that the web server is in China
- Use “nc w.x.y.z 80” (where w.x.y.z is the IP address) to connect to the server and manually request via HTTP:
  - index.html
  - probablynotthere.html
  - falungong.html