I. Announcements

II. What does ICANN (Internet Corporation for Assigned Names and Numbers) do?
   a. Draws up a contract with each registry
   b. Runs an accreditation system for registrars
   c. Oversees IP addresses (through companies)
   d. Oversees root servers
      i. There are 13 root servers on the Internet where complete address tables can be found
      ii. Root servers ultimately resolve addresses
          1. Because things change slowly, intermediate name servers cache addresses
          2. Very few address queries actually come to a root server
   e. Most root servers are anycast
      i. Have many locations from which they operate (so more than 13 in truth)
      ii. Anycast can distribute pain from DDoS attack

III. Revisiting TCP/IP
   a. IP provides an unreliable connection
      i. Every packet has source and destination information
      ii. Each packet is independent of all others
      iii. Packets may be delivered in arbitrary order
      iv. Limited size (long messages have to be subdivided)
   b. TCP
      i. Reliable 2-way byte stream built with IP
      ii. Go to a specific host and a specific port at the host
          1. SSH is port 22, SMTP (25), Telnet (23), HTTP (80), HTTPS(443)
      iii. Each TCP segment is wrapped in an IP packet and sent; has to acknowledge that it was received or it is resent
   c. Picture of TCP/IP in action
   d. How things are connected
      i. Local nets connected to ISP
      ii. These connect to regional ISPs
      iii. Then to larger ISP
          1. E.g. PAETEC, ATT, Cogentco, Level 3, ...
          iv. Traffic exchanged at Internet exchanges (peering points)
   e. Various tools that are available
      i. traceroute
      ii. speedtest
      iii. ping
      iv. various phone books
          1. who owns this site; names to numbers and vice versa
   f. Major Problem
      i. Not enough IP addresses
         1. IPv4 is 4 bytes ~ 4B addresses
      ii. On the horizon IPv6
1. 128 bits (16 bytes) long ~ 340 billion billion billion billion
2. Being adopted very slowly

IV. How big is the internet
   a. Various parameters
      i. Hosts
         1. 1969 – 4 hosts
         2. 1992 passed 1M
         3. 2013 passed 1B
      ii. Domain names
         1. Passed 100K 1994
         2. Passed 1M 1996
         3. Now more than 300M
      iii. Web sites
         1. 1 in 1990
         2. Passed 1M in 1997
         3. Passed 1B in 2014
      iv. Facebook accounts
         1. Started in 2004
         2. Passed 100M in 2008
         3. Passed 1B in 2012
         4. Passed 2b in 2017
   b. Geography
      i. 2012 distribution of hosts
         1. US 505M
         2. Japan 64.5M
         3. Brazil 26.6M
         4. Italy 25.7M
         5. China 20.6M
         6. Germany 20.0M
         7. Iraq, Guam, North Korea, Chad < 30 each
      ii. Internet users (2014)
         1. 3.2B worldwide
         2. China 626M
         3. EU 398M
         4. US 276M
         5. India 237M
         6. Japan, Brazil, Russia, Germany, Nigeria between 60 and 110M)

V. Summarizing the internet
   a. Packets vs circuits
   b. Names and addresses
   c. Routing
   d. Protocols and standards
      i. Particularly TCP/IP
   e. Layering
i. Divide system into layers each of which provides services to the one above it
ii. While calling on services from the one below it
iii. Hides complexity of details

VI. Challenges going forward with internet
   a. Privacy and security are increasingly more difficult
      i. Encryption helps to maintain privacy (http: vs https)
      ii. Security an ever increasing problem
   b. Service guarantees are hard
   c. Some resources are running low
      i. IPv4 vs IPv6 (32 bit vs 128 bit) but not widely accepted

VII. Dealing with service guarantees
   a. Cache’ing locally
   b. Cache’ing nearby
      i. Content Distribution Networks (CDNs)
         1. Content owners (e.g. media companies) pay CDNs to distribute
         2. CDNs pay edge site (e.g. ISPs, data centers) to store multiple copies
         3. You think you’re going to cnn.com but actually you’re going to a site nearby
      c. When you do load cnn
         i. 264 cookies loaded
         ii. Get input from various CDN’s
            1. href="//cdn.krxd.net"
            2. href="//ht.cdn.turner.com"
            3. href="//cdn.cnn.com"
            4. href="//www.i.cdn.cnn.com"
      iii. Input from a number of ad sites
         1. href="//www.googletagsservices.com"
         2. href="//partner.googleadservices.com"
         3. href="//aax.amazon-adsystem.com"
         4. href="//ads.rubiconproject.com"

VIII. Net neutrality – an issue in the news
   a. A decision that the FCC makes
   b. Essentially comes down to what kind of utility the internet is
      i. Title II of Communications Act of 1934 defines common carriers
   c. Do you need to provide service to everyone in the same way or can there be “fast lanes”
   d. What the Obama administration says
      i. High speed internet is a gateway to modern communication, information, entertainment and economic opportunity
      ii. So, regulate broadband and telecom as a utility a la electricity and telephone
   e. What broadband and telecom companies say
      i. This is unnecessary government meddling that will reduce incentives to invest and improve service
      ii. Everyone will lose in the long run
f. Potential consequences
   i. Public interest groups, NGOs, charities, ... say
      1. Internet will become pay-to-play technology with 2 tiers
         a. One with speedy service for big internet and media companies and affluent households
         b. One without for everyone else
      2. Will stifle competition with the big guys
         a. Right now, anyone who puts something on the internet has a chance to find a life-changing audience
   ii. Telecoms and some economists say
      1. Freedom to charge different prices for different products and services is vital to healthy markets
      2. This kind of price discrimination is the fuel of innovation and efficiency
   g. Getting rid of net neutrality allows
      i. Amazon video to bargain for higher speeds than Netflix and so get a business advantage
      ii. Companies to unbundle costs
         1. So $5/mo for hulu, $7 for month for Netflix, $10/month for google, ...
      iii. Companies to block services
         1. E.,g. VOIP
         2. ATT blocking Facetime
         3. Various telecoms blocking Skype
         4. MetroPCS tried to block all streaming besides youtube
   h. Stay tuned, the FCC votes on 12/14

IX. Now that we have the internet, what can we do with it?
   a. Build apps on top of it
      i. Picture – IP at bottom layer connecting physical layers (eg phone, fiber, Ethernet) from client to server
      ii. TCP above at client and server ends
      iii. Apps running on top of TCP
      iv. So, physical layer below packet delivery service below reliable transport service below application
   b. Applications come to `ports’’ on a machine
      i. SSH is port 22, SMTP is port 25, HTTP is port 80 or 8080, HTTPS is port 443, Telnet is port 23, Starcraft is port 6112, ...
      ii. So packets are not just directed to a machine but also to a port.
         1. connecting to Princeton.edu:80 is the same as connecting to www.princeton.edu
         2. data is wrapped (eg http wrapper in tcp wrapper in IP wrapper in Ethernet wrapper)

X. Revisiting communications
   a. Via landline, cellphone, gmail, skype call

XI. Simple advertising on the web
   a. How are ads created?
i. Google Adwords

ii. Budget $1/day

1. All countries and territories, Search network, Keywords – Dobkin, David, COS109
   a. 34+ clicks/day; 2K+ impressions
   b. Can write my own ad to direct people to my website
2. If just US 15+ clicks, 927+ impressions
3. If just New Jersey 5+ clicks, 301+ impressions
4. If just Mercer county 1+ clicks, 53+ impressions
5. If just Princeton, NJ 0 clicks, 1+ impressions

b. How do my ads get placed

i. There are various ad exchanges

ii. User uses a browser to ask for a web page

iii. Web page publisher notifies exchange that advertising space on the page is available

1. Publishers are typically portals or entertainment and news sites
2. Publisher provides information about user
   a. Past online activity
   b. Viewing and shopping habits
   c. Geographic location
   d. Demographics
   e. Probably not actual identity
3. Advertisers bid on the ad space
4. Winner’s advertisement is inserted into the page
5. Typical time 10-100 msec