EE 122: Introduction To Communication Networks

Fall 2013
Sylvia Ratnasamy

http://inst.eecs.berkeley.edu/~ee122/

Material with thanks to Jennifer Rexford, Ion Stoica, Vern Paxson, Scott Shenker, Dave Anderson and other research colleagues.
Today

- Introductions
- What is (this course on) networking about?

5 minute break

- Class policies, administrivia and roadmap
Are you in the right class?

- **Spring** offering: taught by EE faculty
  - More emphasis on link technologies, wireless, communication theory, and mathematical analysis

- **Fall** offering: taught by CS faculty
  - More emphasis on Internet architecture, protocol design and real-world practice

- Classes are different in content and style
Introductions
Teaching Assistants

- Sameer Agarwal
- Kaifei Chen
- Sangjin Han
- Gautam Kumar
- Radhika Mittal
- Andrew Or
- Aurojit Panda
- Colin Scott
- Justine Sherry
- Steve Wang

See the course website for TA office hours and sections
Sameer Agarwal

- 5th year PhD student
- research focus: data centers and databases
- TA for EE122 in Fall’10
Kaifei Chen

- 2nd year PhD student
- research focus: low power networking
Sangjin Han

- 3rd year PhD student
- research focus: high performance network software
Gautam Kumar

- Head TA
- 3\textsuperscript{rd} year PhD student
- research focus: network resource management
- TA for EE122 in Fall’12

-- awarded Outstanding GSI!
Radhika Mittal

- 2nd year PhD student
- research focus: network congestion control
Andrew Or

- Senior
- research focus: software-defined networking (SDN)
- TA for EE122 in Fall’12
Aurojit Panda

- 3rd year PhD student
- research focus: SDNv2
- TA for EE122 in Fall’12
Colin Scott

- 3rd year PhD student
- research focus: network troubleshooting
- TA for EE122 in Fall’12
Justine Sherry

- 4th year PhD student
- research focus: network services
- TA for EE122 in Fall’11
Steve Wang

- Senior
- research focus: network middleboxes
Instructor: Sylvia Ratnasamy

- Ph.D. in Computer Science from Berkeley in 2002
- Worked at Intel between 2002-2011
- Back at UCB, on the faculty since 2011
- Networking has been my research focus throughout

- **Office hours:** Thursday 10:30-11:30am in 413 Soda Hall
  - And by appointment (arrange by email)
  - *Always* happy to chat if you have a serious problem
My teaching style

- Still work-in-progress

- My first time teaching an undergrad course
  - means I’m going to botch some things
  - but I will listen to (constructive) feedback
    - “You speak too fast/slow”
    - ‘Speak LOUDER!’
    - “I can’t stay awake…”
    - “That was incomprehensible…”
What is networking about?
What is a network?

- A system of “links” that interconnect “nodes” in order to move “information” between nodes

- Yes, this is very vague
There are many different types of networks

- Internet
- Telephone network
- Transportation networks
- Cellular networks
- Supervisory control and data acquisition networks
- Optical networks
- Sensor networks

We will focus almost exclusively on the Internet
The Internet is transforming everything

- The way we do business
  - E-commerce, advertising, cloud-computing

- The way we have relationships
  - Facebook friends, E-mail, IM, virtual worlds

- The way we learn
  - Wikipedia, MOOCs, search engines

- The way we govern and view law
  - E-voting, censorship, copyright, cyber-attacks

Took the dissemination of information to the next level
The Internet is big business

- Many large and influential networking companies
  - Cisco, Broadcom, AT&T, Verizon, Akamai, Huawei, ...
  - $120B+ industry (carrier and enterprise alone)

- Networking central to most technology companies
  - Google, Facebook, Intel, HP, Dell, VMware, …
Internet research has impact

- The Internet started as a research experiment!
- 4 of 10 most cited authors work in networking
- *Many* successful companies have emerged from networking research(ers)
But why is the Internet interesting?

“What’s your formal model for the Internet?” -- theorists

“Aren’t you just writing software for networks” – hackers

“You don’t have performance benchmarks???” – hardware folks

“Isn’t it just another network?” – old timers at AT&T

“What’s with all these TLA protocols?” – all

“But the Internet seems to be working…” – my parents
A few defining characteristics of the Internet
A federated system

- The Internet ties together different networks
  - >18,000 ISP networks

Tied together by IP -- the “Internet Protocol” : a single common interface between users and the network and between networks
A federated system

- The Internet ties together different networks
  - >18,000 ISP networks

- A single, common interface is great for interoperability…
- …but tricky for business

- Why does this matter?
  - ease of interoperability is the Internet’s most important goal
  - practical realities of incentives, economics and real-world trust drive topology, route selection and service evolution
Tremendous scale

- 2.4 Billion users (34% of world population)
- 1 Trillion unique URLs
- 294 Billion emails sent per day
- 1 Billion smartphones
- 937 Million Facebook users
- 2 Billion YouTube videos watched per day
- Routers that switch 10 Terabits/second
- Links that carry 100 Gigabits/second
Enormous diversity and dynamic range

- Communication latency: microseconds to seconds ($10^6$)
- Bandwidth: 1Kbits/second to 100 Gigabits/second ($10^7$)
- Packet loss: 0 – 90%

- Technology: optical, wireless, satellite, copper

- **Endpoint devices**: from sensors and cell phones to datacenters and supercomputers
- **Applications**: social networking, file transfer, skype, live TV, gaming, remote medicine, backup, IM
- **Users**: the governing, governed, operators, malicious, naïve, savvy, embarrassed, paranoid, addicted, cheap …
Constant Evolution

1970s:
- 56kilobits/second “backbone” links
- <100 computers, a handful of sites in the US
- Telnet and file transfer are the “killer” applications

Today
- 100+Gigabits/second backbone links
- 5B+ devices, all over the globe
- 20M Facebook apps installed per day
Asynchronous Operation

- Fundamental constraint: speed of light

- Consider:
  - How many cycles does your 3GHz CPU in Berkeley execute before it can possibly get a response from a message it sends to a server in NY?
    - Berkeley to New York: 4,125 km
    - Traveling at 300,000 km/s: 13.75 milliseconds
    - Then back to Berkeley: 2 x 13.75 = 27.5 milliseconds
    - $3,000,000,000$ cycles/sec * 0.0275 = 84,000,000 cycles!

- Thus, communication feedback is always dated
Prone to Failure

- To send a message, all components along a path must function correctly
  - software, modem, wireless access point, firewall, links, network interface cards, switches,…
  - Including human operators

- Consider: 50 components, that work correctly 99% of time $\rightarrow$ 39.5% chance communication will fail

- Plus, recall
  - scale $\rightarrow$ lots of components
  - asynchrony $\rightarrow$ takes a long time to hear (bad) news
An Engineered System

- Constrained by what technology is practical
  - Link bandwidths
  - Switch port counts
  - Bit error rates
  - Cost
  - …
Recap: The Internet is...

- A complex federation
- Of enormous scale
- Dynamic range
- Diversity
- Constantly evolving
- Asynchronous in operation
- Failure prone
- Constrained by what’s practical to engineer
Recap: The Internet is…

- Too complex for theoretical models
- “Working code” doesn’t mean much
- Performance benchmarks are too narrow
So, what do we need?

We still don’t really know...

- No consensus on what constitutes the “correct” or “best” network design
- No consensus on “top 10 problems”
- No consensus on the right prioritization of goals

Before you flee...
What we do know

- The early Internet pioneers came up with a solution that was successful beyond all imagining
- Several enduring architectural principles and practices emerged from their work
Some key principles

- Statistical multiplexing [lecture 2]
- Packets [lecture 2]
- The network is “application neutral” [lecture 3]
- Best effort service [lecture 3]
- A layered protocol architecture [lectures: all]
- A “narrow API” at the network layer [lecture 8]
- The “end to end” design principle [lecture 8]
- Decentralization [lecture: 2, 3, 6, 8, 9, 21]
What we do know

- The early Internet pioneers came up with a solution that was successful beyond all imagining
- Several enduring architectural principles and practices emerged from their work
- But it is just one design
- And numerous cracks have emerged over time
  - want to diagnose problems but IP hides federation
  - want to block unwanted traffic but the network doesn’t authenticate
  - can’t optimize for different applications or customers
  - complex and buggy protocols
  - upgrading protocols is deeply painful
What we do know

- The early Internet pioneers came up with a solution that was successful beyond all imagining
- Several enduring architectural principles and practices emerged from their work
- But it is just one design
- And numerous cracks have emerged over time
- As have new requirements
  - Mobility, reliability, data centers, sensors, …
Hence, networking today is still debating the big questions...

- Packets → “circuits”
- Statistical multiplexing → “reservations”
- Protocol layers
- A “narrow waist” at the network layer
- Best-effort service
- The “end to end” design principle → “middleboxes”
- Decentralization → “centralize”
Backing up a level

- The Internet offers us a lesson in how to reason through the design of a very complex system
  - What are our goals and constraints?
  - What's the right prioritization of goals?
  - How do we decompose a problem?
  - Who does what? How?
  - What are the interfaces between components?
  - What are the tradeoffs between design options?

- In short: a lesson in how to architect a system
Network Architecture

- More about thinking rigorously than doing rigorous math
- More about understanding tradeoffs than running benchmarks
- More about practicality than optimality
What (I hope) EE 122 will teach you

- How the Internet works
- Why it works the way it does
- How to think through a complicated (networking) design problem
Let’s take a 5 minute break
Today

- Introductions
- What is (this course on) networking about?

5 minute break

- Class policies, roadmap, administrivia
Class Workload

- Three projects
- Three homeworks
- Exams:
  - midterm: October 23 in class
  - final: December 20, 3-6pm, location TBA
  - closed book, open crib sheet
Grading

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Homeworks</td>
<td>15% (3x 5% each)</td>
</tr>
<tr>
<td>3 Projects</td>
<td>40% (10+10+20)</td>
</tr>
<tr>
<td>Midterm exam</td>
<td>20%</td>
</tr>
<tr>
<td>Final exam</td>
<td>25%</td>
</tr>
</tbody>
</table>

- Course graded to mean of B
Topics we will cover

- Basic concepts [Lectures 2, 3]
  - packets, circuits, delay, loss, protocols
- How the “insides” of the Internet work [Lectures 3-8]
  - IP, DV/LS routing, BGP
- How endpoints use the network [Lectures 9-16]
  - TCP, DNS, HTTP
- Crucial lower-level technologies [Lectures 17-20]
  - Ethernet, wireless
- Important new(er) topics [Lectures 21-26]
  - management, security, datacenters
Three projects

- Project 1: Routing (in simple simulator)
- Project 2: Reliable Transport (in simple simulator)
- Project 3: TBD (will involve a “real” network)
  - Larger project, in two phases

TAs will handle all project-related questions!
Administrivia: Textbook

  - 5th Edition ok, but translate the reading assignments

- You will not be tested on material we didn’t cover in lecture or section
  - Use as a reference and a source of examples
Enrollment and wait list

- Class size is capped at 330

- Wait-listed students will be admitted as and when registered students drop the class
  - Seniors will be given priority
  - If you’re planning to drop, please do so soon!
Class communications

- Web site: http://inst.eecs.berkeley.edu/~ee122/
  - Assignments, lecture slides, announcements

- Use your instructional account to hand in homework and projects
  - accounts will be handed out next week

- Use Piazza for all other intra-class communication
  - You should all be signed up by now

- Copy Gautam (gautamk@cs) on any emails sent directly to me (sylvia@cs)
Policy on late submissions, re-grade requests, cheating

- Detailed description is on the class website

- Summary version:
  - You may submit assignments late, but to a point, and it will cost you
  - You can submit requests to re-grade a homework or a midterm, but it may cost you
  - The policy on re-grades for projects will be announced by the lead TA on the project and may vary across projects
  - Don’t cheat
Class Participation

- Ask and answer questions!!
  - it helps you understand
  - it helps others understand
  - it helps you stay awake
  - it helps me stay awake
  - it’s just more fun for all of us

- Sit towards the front

- Limit electronic access for < 90 minutes
  - you will have a 5 minute break in the middle to get online
Summary

- Learning about the Internet is
  - important
  - relevant
  - fun -- challenging and interdisciplinary problems

- For our next lecture
  - read 1.1 and 1.3 of K&R
  - make sure you are registered with the correct email addr. and on piazza
Any questions?